Investigation of Remote Sensing for Assessing and Monitoring the *Araucaria* forest region of Brazil

Yeda Maria Malheiros de Oliveira

Green College

and

Oxford Forestry Institute

Department of Plant Sciences

University of Oxford

Thesis submitted for the degree of Doctor of Philosophy at the University of Oxford

Oxford, UK
Michaelmas Term, 1999

Investigation of Remote Sensing for Assessing and Monitoring the *Araucaria* forest region of Brazil

Yeda Maria Malheiros de Oliveira

Green College

and

Oxford Forestry Institute

Department of Plant Sciences

University of Oxford

Thesis submitted for the degree of Doctor of Philosophy at the University of Oxford

Oxford, UK
Michaelmas Term, 1999

"Subtle is the Lord ..."
Albert Einstein

"The rhythm of destruction of the forests took such extension, that each inhabitant of Paraná has as obligation to labour with all the means for that the destruction of the forests is impeded by force of the law.
Reinhard Maak"

This work is dedicated to the memory of Maria Grillo Malheiros Pinto, Adelar de Oliveira e Regina Malheiros de Oliveira. I also dedicate it to my sister Yara and my brother Adelar Jr.

ACKNOWLEDGEMENTS

" . . . for with God all things are possible" (Mark 10:27).

Foremost I wish to express my gratitude to my supervisors Professor Jeffery Burley, Dr Mike Packer and Dr Terence Dawson for their inspirational commitment to a vision. I was privileged to be associated to this very noble "team". Their guidance was fundamental to the realisation of my DPhil programme.

In addition, I would like to thank the following people and institutions:

The Brazilian Ministry of Agriculture/EMBRAPA-Florestas kindly agreed to release me from my official duties in order to come to England and provided me with financial support throughout this study period.

My fellow colleagues at EMBRAPA-Florestas have continuously provided data, information, friendship and encouragement. I am particularly indebted to Carlos A. Silva, Carlos A. Mazza, M.Cristina Mazza, Edilson Oliveira, Amilton Baggio and Antonio Carpanezzi. I wish to thank Vera Eifler especially, for her dedication, faith and understanding. My colleagues at EMBRAPA-Florestas and Sede provided me with administrative assistance and useful information during my study in Oxford. My friend and internal supervisor at EMBRAPA, Dr Vitor Hoeflich, supported and guided me from the beginning.

My friends at INPE; Thelma Krug, Flavio Ponzoni and João R. dos Santos provided me with invaluable material (Landsat imagery), guidance and incentive during the "hard work" phase at S.J.dos Campos. I am grateful to Claudio Barbosa for his patience during my SPRING macro language training stage. He always gave of his time to help me. My special thanks to Claudia Linhares de Sousa, for her understanding and friendship. She made my life more colourful.

I do not believe it was a coincidental that my long standing friend Luis Uchoa became the director at Irati National Forest: it was inspirational. His promotion and subsequent assistance, as well as that of Trajano Gracia came at a time when I most needed help.

I thank my colleagues and staff of the Plant Sciences Department at the Oxford University for their help with administrative queries and for their patience, to Liz Pearce for her encouragement in some difficult moments. I am also indebted to my friends at Green College, especially to Mrs.J.Davies.

My attorneys and friends, Ana Coelho and Valdemar Borba advise and managed my funds skillfully.

I wish to thank Gracie A. Maximiano, a selfless angel who came from Brazil just when I needed her most. I thank her for being such an example of altruism. I am very grateful to my friend Sherrill Conroy for her kindness, encouragement and corrections to my English. I am in debt to Gilson Geronasso, Joesio Siqueira, Jefferson Mendes and Roberto Hosokawa who generously gave of their time, sharing information data without thought of recompense.

To my family, who I know will be with me always. Last but not least, I would like to thank Janet Steele who on a daily basis was there for me.

I finally want to apologise for not mentioning personally so many friends who have encouraged me and prayed for my success.

ABSTRACT

Southern Brazil has experienced a rapid deterioration of its forest resources during the last century, mainly due to the impacts of exploitative land use practices. Mixed Ombrophilous Forest with *Araucaria angustifolia* (Bertoloni) O. Kuntze (Veloso and Goes-Filho), is typically dominated by the species *Araucaria angustifolia,* also known as paraná-pine. Although the species is predominant, *Araucaria* forest also supports a complex, variable and regional ecosystem composed of many species, some of which are endemic to this forest type.

The present study focused on an area of the *Araucaria* forest biome located within the first plateau of Paraná State. It includes the Irati National Forest (INF), the only protected area with *Araucaria* forest in Paraná State.

The main goal of the thesis was to determine the relationship between vegetation and satellite remote sensing for assessing and monitoring *Araucaria* forest and ecosystems associated with the *Araucaria* forest region. The reflective properties and spectral signatures of the four main forest types in the INF (*Araucaria* forest, *Araucaria angustifolia* plantation, *Pinus elliottii* plantation and *Pinus taeda* plantation) were analysed and characterised. It was noted that spectral responses of *A. angustifolia* plantation and *P.taeda* plantation are similar. Standard Principal Component Analysis (PCA) revealed that the spectral data space could be reduced to three eingenvectors. The discrimination of *A.angustifolia* was achieved using the second component of the PCA using near infrared and mid infrared wavebands. Linear Mixing Modelling was undertaken to detect and identify *A.angustifolia* at sub-pixel level. The shade image was sensitive to forestry structure (density) and species differentiation. Use of the Normalised Difference Vegetation Index was successful in the biomass change detection using image differencing. These techniques were then utilised in an image differencing process, resulting in a proposed alternative monitoring system to detect forest changes in the *Araucaria* forest region.

TABLE OF CONTENTS

Abstract v

List of tables x

List of figures xii

Chapter 1: INTRODUCTION 1
1.1 *Araucaria* forest status and problems 1
1.2 The study area 3
1.3 Remote sensing of forest resources 3
1.3.1 State of art of RS investigation of *Araucaria* forest 4
1.4 Aims of this study 5
1.5 Structure of thesis 6

Chapter 2: THE ENVIRONMENT OF THE STUDY AREA 8
2.1 Introduction 8
2.2 The genus *Araucaria* 8
2.3 *Araucaria* forest region 9
2.4 Conservation Units 11
2.4.1 National Forest 11
2.4.1.1 Irati National Forest (INF) 12
2.4.1.1.1 Geology, geomorphology, soils and climate 13
2.4.1.1.2 Vegetation 13
2.4.1.2 INF Adjacent area 14
2.5 Fieldwork – Forest Inventory 15
2.5.1 *Pine* and *Araucaria angustifolia* plantation 16
2.5.2 *Araucaria* forest 16
2.6 Characterisation of Araucaria forest at INF 18

Chapter 3: REMOTE SENSING INSTRUMENTATION AND PROCESSING TECHNIQUES 20
3.1 Introduction 20
3.2 Remote sensing instrumentation 20
3.2.1 Landsat description 20
3.2.2 Orbital images, topographic maps and ancillary data 22
3.2.3 Brazilian system for receiving satellite data 23
3.3. Factors affecting spectral characteristics of remote sensing data 24
3.3.1 Sensor-related effects 24
3.3.1.1 Sensor calibration coefficients 25
3.3.2 Scene-related effects 26
3.3.2.1 Atmosphere 26
3.3.2.2 View and illumination geometry 29
3.3.2.3 Topographic effects 30
3.4 Processing methods 31
3.4.1 Image restoration 32
3.4.1.1 Geometric restoration 33
3.4.1.1.1 Geometric restoration by resampling 34
3.4.1.1.2 Ground control points 37
3.4.1.1.3 Error Assessment 39

3.4.1.2 Geometric co-registration (registration of the image-to-image) 40
3.4.1.3 Radiometric restoration 41
3.4.1.3.1 Radiometric correction 42
3.4.1.3.1.1 Converting DN to radiance 42
3.4.1.3.1.2 Converting radiance to reflectance 44
3.4.1.3.2 Calibration 45
3.4.1.3.2.1 Absolute calibration 46
3.4.1.3.2.2 Relative calibration 47
3.4.1.3.2.2.1 Radiometric rectification 48
3.5 Summary 53

Chapter 4: SPECTRAL CHARACTERISTICS OF THE REPRESENTATIVE FOREST TYPES 56
4.1 Introduction 56
4.2 Physical and physiological properties of plants (Reflective properties of tree foliage) 57
4.2.1 Sources of energy 57
4.2.2 Spectral characteristics of natural systems 59
4.3 Water, Soil and Vegetation 60
4.3.1 Water 60
4.3.2 Soils 61
4.3.3 Vegetation 62
4.3.4 Spectral characteristics of vegetation 63
4.4 Spectral responses of the forest types in INF 63
4.4.1 Pine species – spectral responses and structural variables relationships – Experiment 1 64
4.4.1.1 Generalization about the genus in Brazil 65
4.4.1.2 Fieldwork 65
4.4.1.3 Imagery information 66
4.4.1.4 Location of the sample units 66
4.4.1.5 Analysis of ground data – relationship between forest plot parameters 67
4.4.1.6 Spectral and structural relationship 67
4.4.1.6.1 Samples units 67
4.4.1.6.2 Spectral responses 69
4.4.1.6.2.1 Reflectance values of TM7 70
4.4.1.6.2.2 Reflectance values of TM5 70
4.4.1.6.2.3 Reflectance values of bands 1, 2, 3 and 4 structural factors 71
4.4.1.6.3 Regression analysis 71
4.4.1.6.4 Summary of results and conclusions for Experiment 1 72
4.4.2 *Araucaria angustifolia* plantation – spectral responses and structural variables relationships – Experiment 2 73
4.4.2.1 Generalization about the genus in plantations 73
4.4.2.2 Field work and imagery information 74
4.4.2.3 Location of the sample units 74
4.4.2.4 Ground data – relationships between forest plot parameters 75
4.4.2.5 Spectral and structural relationships 75
4.4.3 *Araucaria* forest – spectral responses and species composition–Experiment 3 76
4.4.3.1 General comments on the biome 76
4.4.3.2 Field work and imagery information 77
4.4.3.3 Location of the sample units 78
4.4.3.4 The K/d ratio 78
4.4.3.5 Ground data – relationships between forest plot parameters 79
4.4.3.6 Crow competition – a density indicator 80
4.4.3.7 Testing the equality of population means 82
4.4.3.7.1 Analysis of variance for basal area, crown diameter and spectral responses 83
4.5 Typical spectral reflectance curves in the study area 85
4.6. Conclusions 87

Chapter 5: THE CONTRIBUTION OF IMAGE TRANSFORMATIONS: DIMENSIONALITY REDUCTION OF THE FEATURE SPACE 89
5.1 Introduction 89
5.2 Vegetation Indices (VIs) approach 89
5.2.1 The RATIO 90
5.2.2 The NDVI 92
5.3 The orthogonal transformations 93
5.3.1 Principal Components Analysis (PCA) 93
5.3.1.1 Standard PCA – Experiment 1 97
5.3.1.2 Selective PCA – Experiment 2 102
5.3.1.2.1 Approach 1 when high correlation exists between bands 102
5.3.1.2.2 Approach 2 when low correlation exists 104
5.3.2 Tasselled Cap (Kauth-Thomas) transformations 108
5.4 Conclusions 110

Chapter 6: SPECTRAL MIXTURE MODELLING – REMOTE SENSING METHODS TO DETECT AND IDENTIFY TARGETS AT THE SUB-PIXEL SCALE 113
6.1 Introduction 113
6.2 The mixture concept 113
6.3 Methods for deriving spectral information 114
6.3.1 Spectral Mixture Modeling (SMM) 115
6.3.1.1 Intimate models 115
6.3.1.2 Macroscopic models 116
6.3.1.2.1 Spectral Mixture Analysis (SMA) 116
6.3.1.2.2 The Linear Mixing Model (LMM) approach – basic concepts 116
6.4 Constrained versus unconstrained approaches 117
6.4.1 Singular-value decomposition (SVD) 118
6.4.2 Alternative approaches 118
6.4.3 Constrained and unconstrained least-squares approaches 118
6.4.3.1 The Constrained Least-Square (CLS) method 119
6.4.3.2 The Weighted Least-Squares (WLS) method 120
6.5 Methods for estimating proportions 121
6.5.1 Facets of the least-squares approach 122
6.6 Error assessment 123
6.6.1 Visual analysis 123
6.6.2 Calculation of the mean error 123
6.6.3 Computation of fraction overflow 124
6.7 Applications of SMM in vegetation studies 124
6.8 The selection of end-members 125
6.9 The pure pixel (end-member) and training data 128
6.9.1 Sets of data 129
6.9.2 Point data 129
6.10 Experiments description 130
6.10.1 Experiment 1 – Linear Mixing Model: three candidates as end-members 130
6.10.1.1 The selection of end-members 131
6.10.1.1.1 The end-members vegetation and shade 133
6.10.1.2 The estimation of proportions 134
6.10.1.3 The resultant images 135
6.10.1.4 The evaluation of the model fitness 139
6.10.1.4.1 The overflow images (OI's) 140
6.10.1.4.2 The error images (EI's) 141
6.10.2 Experiment 2: the inclusion of the end-member soil in the model 142
6.11 Conclusions 148

Chapter 7: CHANGE DETECTION IN *ARAUCARIA* FOREST 151
7.1 Introduction 151
7.2 The impact of mis-registration on change detection 152
7.3 Detecting positive and negative changes of biomass 153
7.3.1 Image differencing 153
7.3.1.1 Simple image differencing – Overlay 154
7.3.1.2 Simple image differencing using NDVI 156
7.3.1.3 Simple image differencing with thresholding 158
7.3.1.4 Linear regression differencing 162
7.3.1.5 Normalised Image Differencing (NID) 166
7.3.1.6 Agreement analysis 168
7.4 Categorizing losses of vegetation 170
7.4.1 Image differencing classification 170
7.4.1.1 Data and methods 171
7.4.1.2 Sample size and classification scheme 172
7.5 Alternative process of change detection for monitoring purposes 175
7.5.1 Step 1 – Preparing the differencing image 176
7.5.2 Step 2 – Segmenting the differencing image 177
7.5.3 Step 3 – Classifying the segmented image 179
7.5.4 Supervised classification 181
7.6 Agreement analysis 182
7.7 Conclusions 185

Chapter 8: CONCLUSIONS AND RECOMMENDATIONS 190
8.1 Introduction 190
8.2 Conclusions 191
8.3 Future research directions and recommendations 201
8.3.1 Change detection and pos-classification comparison methods 203

References 205

Appendix 2.1 Paraná - Forest Remnants and Ecosystems associated to the Mata Atlântica domain – period 1990-95 218

Appendix 2.2 GPS – Global Positioning System 219

Appendix 2.3 List of native forest species inventoried at the Irati National Forest – INF 221

Appendix 3.1 Software SPRING 223

Appendix 4.1 Descriptive statistics of Experiment 1 225

Appendix 4.2 Descriptive statistics of Experiment 3 230

Appendix 5.1 List of vegetation indices calculated in preliminary stages of work 233

Appendix 6.1 Fraction images obtained using the Unconstrained Least-Square method - model with three end-members 234

Appendix 7.1 Classification Error Matrix 235

Appendix 7.2 *Araucaria angustifolia* and *Araucaria* forest reserves 237

TABLES

Chapter 2

Table 2.1 *Araucaria* forest: Original geographic distribution 9
Table 2.2 Devastation sequence of *Araucaria* forest in Paraná State 10
Table 2.3 Vegetation of the INF 13
Table 2.4 Forest plantations composition 14
Table 2.5 Summary of the sampling statistical analysis 17
Table 2.6 Summary of Basal Area (BA, in m^2) and number of trees (Trees) (typical values per hectare (ha^{-1}) for trees with more than 10 or 20 cm BDH inventoried in different areas with *Araucaria* forest 19

Chapter 3

Table 3.1 Some applications of Landsat Thematic Mapper (TM), 30 m resolution 22
Table 3.2 Temporal set used in the study 23
Table 3.3 Criteria for selecting the data set 23
Table 3.4 Applications using DN or reflectance values 24
Table 3.5 Sequence followed by SPRING in the rectification process 39
Table 3.6 Values of L_{max} and L_{min} used in the conversion to radiance 44
Table 3.7 Solar Elevation Angle (in degrees) at the time where the images were taken 45
Table 3.8 TM Solar exoatmospheric spectral irradiances at mean earth-sun distance used in the conversion to reflectance..... 45
Table 3.9 Reflectance means* of bright and dark control sets before rectification, and the difference of the means of the rectified image for the reference image (1987) and subject image (1997) 52

Chapter 4

Table 4.1 Limits of electromagnetic spectrum 58
Table 4.2 Main forest features in study area 64
Table 4.3 Location of the sample units – ANOVA applied to the different data sets 67
Table 4.4 Analysis of Variance for TM5 69
Table 4.5 Linear regression between TM5 and Volume 72
Table 4.6 Descriptive statistics of field data (*Araucaria angustifolia*) collected from stands within forest management units 75
Table 4.7 Correlation between forest stand parameters and spectral responses (Pearson) 76
Table 4.8 Descriptive statistics of field data collected from patches by sampling..... 79
Table 4.9 Matrix of correlation 81
Table 4.10 Summary of Tukey's pairwise comparisons for BA_A, CD_P and TM5..... 84
Table 4.11 Typical spectral values (reflectance * 255) for the analysed forest types 86

Chapter 5

Table 5.1 Correlation Matrix for the six reflective bands (1987 set) 98
Table 5.2a Variance-covariance matrix for the INF Landsat TM, for the 1987 set 98
Table 5.2b Principal components loadings (eigenvector matrix) for the six principal components of INF TM imagery, based on the covariance matrix 99
Table 5.3 Principal components loadings (eigenvector matrix) for selective PCA (based on the covariance matrix) of INF TM imagery – highly correlated bands 103

Table 5.4 Correlation coefficients, percent of variance mapped to PC1 and PC2, loadings (eigenvectors) for each pair of TM bands – spectral contrast for INF 104
Table 5.5 Coefficients for the tasselled cap functions for TM bands 1 to 5 and 7 108

Chapter 6

Table 6.1 Possible outcomes for the System equation solution 120
Table 6.2 Mean reflectance values for each component, mean error for each band and total error for the model considering the different approaches for unmixing pixels..... 134

Chapter 7

Table 7.1 Basic statistics of the NDVI87, NDVI97 and the difference images (Dif.image, Dif.imageR and Dif.imageN)..... 160
Table 7.2 Cross tabulation between each two thematic layers..... 169
Table 7.3 Cross tabulation between each two image differencing classification approaches..... 183
Table 7.4 Agreement among the three image differencing classification approaches..... 183

APPENDICES

Appendix 4.1

Table 4.1.1 Descriptive statistics of field data collected from stands within forest management units 225
Table 4.1.2 Analysis of Variance for Volume - *P.elliottii* (1) and *P.taeda* (2) 225
Table 4.1.3 Correlation between forest plot parameters..... 226
Table 4.1.4 ANOVA applied to the eight groups of data sets..... 226
Table 4.1.5 Correlations Coefficients of VOL (volume) and BA (Basal Area) and respective logarithms (L) against spectral reflectance and the logarithm of the spectral reflectance (Pearson) - *P.elliottii* and *P.taeda*..... 227
Table 4.1.6 Correlations (Pearson)- *P.elliottii* and *P.taeda*..... 227
Table 4.1.7 - Synthetic images correlated with volume and basal area..... 229

Appendix 7.2

Table 7.2.1 Descriptive statistics of the spectral variables for 1987/97; comparative analysis between the two dates using Matched pairs *t*-test..... 238
Table 7.2.2 Linear regression performed using information from 1987 as an independent variable and information from 1997 as a dependent variable 239

FIGURES

Chapter 2

Figure 2.1 The study area, showing Irati National Forest (INF) and adjacent areas. Irati, F. Pinheiros and T. Soares are the main urban areas in the study area 8
Figure 2.2 Original geographic distribution of Mixed Ombrophilous Forest with *Araucaria angustifolia* 9
Figure 2.3 Extent of deforestation in the Atlantic forest domain by 1990..... 10
Figure 2.4 Irati National Forest and part of study area, showing Irati town in the down left corner. Main roads are in red and main streams in blue..... 12
Figure 2.5 Main species inventoried at INF and at Reserve of Caçador..... 18

Chapter 3

Figure 3.1 Sun-synchronous orbit, the Landsat TM 5 orbit and the region where the study area is located. 21
Figure 3.2 Scheme of spatial data collection 25
Figure 3.3 Transmittance spectral of atmospheric ozone, carbon monoxide, nitrous oxide methane, and oxygen 27
Figure 3.4 Measurements of reflected solar energy influenced by atmospheric effects. 28
Figure 3.5 Solar elevation and azimuth angles 29
Figure 3.6 Different resampling methods to assign the appropriate DN to an output pixel 35
Figure 3.7 Polynomial Curve 38
Figure 3.8 The rectification process, adopted polynomial and the number of control points definition..... 38
Figure 3.9 RMS Error Tolerance 40
Figure 3.10 Schematic geometric co-registration 40
Figure 3.11 Atmosphere water vapour transmittance spectrum (left plot) and carbon dioxide transmitance spectrum (right plot) for specific conditions. The spectrum was calculated using LOWTRAN7 46
Figure 3.12 Scheme for determination of dark and bright radiometric sets for each date 50
Figure 3.13 Calculation of radiometric transform – rectification transform approach..... 50
Figure 3.14 Scheme of stages for geometric restoration and possible approaches for radiometric restoration 55

Chapter 4

Figure 4.1 Electromagnetic wave 57
Figure 4.2 Regions of the electromagnetic spectrum 58
Figure 4.3 Radiometric properties of the objects: Absorption (A), Reflection (R) and Transmission (T) 59
Figure 4.4 Interrelationships between energy sources and atmospheric absorption characteristics 60
Figure 4.5 Typical spectral reflectance curve of vegetation, soil and water 61
Figure 4.6 Differences between conifers and deciduous species spectral response 63
Figure 4.7 Main features of the INF..... 64
Figure 4.8 Central point determined by GPS (a), 3-by 3 window (b), combinations of 3-by 3 window (g and h) and combinations of 2-by 2 window (c, d, e and f) 67

Figure 4.9 (a) Boxplots representing the magnitude of the differences between the selected data sets and (b) Interval plots representing the spread of data around the mean of the variable 'Reflectance of band 5' on each plot 68
Figure 4.10 Bivariate scatterplot - Linear regression between Volume in m^3 (V) as the dependent variable and reflectance of band 5 (TM5) as the independent variable..... 72
Figure 4.11 *Araucaria* forest patches in the INF 83
Figure 4.12 Typical reflectance curves of the main forest types in the study area 86
Figure 4.13 Typical reflectance curves of vegetated and non vegetated areas (14 classes) 87

Chapter 5

Figure 5.1 (a) RATIO VI image and (b) NDVI image focusing the area of the INF. Both images have been subjected to a histogram equalisation contrast stretch ... 91
Figure 5.2 (a) Two perfectly correlated variables (x and y) and (b) two highly correlated variables (x and y) 95
Figure 5.3 Basic statistical parameters of the space attribute 96
Figure 5.4 Sub-scene of the non-standardised principal component 2 (PC2) of TM bands 1-5 and 7. A young pine stand outside the INF (1) and an older pine stand inside the INF (2) are marked. 100
Figure 5.5 Sub-scenes of the non-standardised principal component 3 (PC3) of TM bands 1-5 and 7. (a) shows (1) pine plantation, (2) *A.angustifolia* plantation and (3) *Araucaria* forest. (b) shows (1) *P.elliottii* and (2) *P.taeda* stands..... 100
Figure 5.6 Sub-scene of the study area, focusing the INF. (a) to (c) respectively show the non-standardised principal components 1-3 (PC1, PC2 and PC3) of TM bands 1-5 and 7..... 101
Figure 5.7 Colour composite of PC1 (from selective principal components of TM1 to TM3), PC1 (from selective principal components of TM5 and TM7) and TM4. Numbered are the main features of the INF. Stands of slash pine is labelled with (1), loblolly pine with (2), *Araucaria angustifolia* plantation with (3) and *Araucaria* forest with (4) 104
Figure 5.8 Sub-scene showing the presence/absence of conifers in the INF: (1) Pine plantation; (2) *A.angustifolia* plantation; (3) *Araucaria* forest; (4) alluvial soil (floodplain areas – absence of conifers); and (5) riparian areas (area with absence of conifers) 106
Figure 5.9 Series of images illustrating presence/absence of conifers..... 107
Figure 5.10 (a) greenness image; (b) brightness image; and (c) wetness image obtained from Landsat TM5 bands 1 to 5 and 7. The main features of the INF are numbered: stands of slash pine is labeled with (1), loblolly pine with (2), *Araucaria angustifolia* plantation with (3) and *Araucaria* forest with (4)..... 109

Chapter 6

Figure 6.1 Representation of mixed pixel when two different land covers are presented 114
Figure 6.2 Region that satisfies the constraints for $n = 3$ components 120
Figure 6.3 Schematic diagram for extracting end-members using sub-pixel processing. 127
Figure 6.4 Capture of spectral information in (a) *sets of data* and for (b) *point data*..... 130
Figure 6.5 Principal component 1 (PC1 - *x* axis and principal component 2 (PC2 - *y* axis.(Area around the INF) 131
Figure 6.6 Principal component 1 (PC1 - *x* axis and principal component 3 (PC3 – *y* axis. (Area around the INF) 132

Figure 6.7 Spectra of the candidates as end-members 135
Figure 6.8 Aspect of areas with *P.elliottii* (labeled as 1), *P.taeda* (labeled as 2), *Araucaria angustifolia* plantation (labeled as 3) and *Araucaria* forest (labeled as 4) in the Vt1fraction image..... 136
Figure 6.9 (a) Fraction image of Vt1 – Component *Araucaria angustifolia* – INF and adjacencies; (b) Fraction image of Vt1 – Component *Araucaria angustifolia* focusing the INF; (c) Fraction image of Vt2 – Component *Araucaria* forest; and(d) Fraction image of the Component Shade obtained using the Unconstrained Weighted Least-Square method (WLS)..... 137
Figure 6.10 (a)RGB colour composite of the faction images, where VT1 (R), Shade (G), Vt2 (B) and (b) HIS composition of the same images, where Shade (H), Vt2 (I), Vt1 (S)..... 139
Figure 6.11 Overflow images (a) OI of Vt1 – Component *Araucaria angustifolia*; (b) OI of Vt2 – Component *Araucaria* forest; and (c) OI of the Component Shade .. 140
Figure 6.12 Error images (EI's): (a) EI of TM4 and (b) EI of TM7. The images have been subjected to a histogram equalisation contrast stretch 142
Figura 6.13 a) Aspect of areas with Pelliotti (labeled as 1), Ptaeda (labeled as 2), *Araucaria angustifolia* plantation (labeled as 3) and *Araucaria* forest (labelled as 4) in the Vt1 fraction image in experiment 2 (four end-members)..... 143
Figure 6.14 (a) Fraction image of Vt1 – Component *Araucaria angustifolia*; (b) Fraction image of Vt2 – component *Araucaria* forest; and (c) Fraction image of the Component Shade; and (d) Fraction image of the Component Soil obtained using the Unconstrained Weighted Least-Square method (WLS). 144
Figure 6.15(a) Aspect of areas with *P.elliottii* (labeled as 1), *P.taeda* (labeled as 2), *Araucaria angustifolia* plantation (labeled as 3) and *Araucaria* forest (labeled as 4) in the Vt2 fraction image - experiment 1 (three end-members); (b) same identification for Vt2 fraction image - experiment 2 (four end-members)..... 145
Figure 6.16(a) Aspect of areas with *P.elliottii* (labeled as 1); *Araucaria angustifolia* plantation (labeled as 3); *Araucaria* forest (labeled as 4); stream (labeled as 5); and floodplain (labeled as 6) in the shade fraction image – experiment 1 (three end-members); (b) same identification for shade fraction image - experiment 2 (four end-members)..... 146
Figure 6.17 Overflow images (a) OI of Vt1 – Component *Araucaria angustifolia*; (b) OI of Vt2 – Component *Araucaria* forest; (c) OI of the Component Shade; and (d) OI of the Component Soil 147
Figure 6.18 Error images (EI's): (a) EI of TM3 and (b) EI of TM4. The Images have been subjected to a histogram equalisation contrast stretch 147

Chapter 7

Figure 7.1 Simple image differencing image from TM band 5 (Dif.Image5). The image has been subjected to a histogram equalisation contrast stretch for printing purposes..... 155
Figure 7.2 (a) NDVI87; (b) NDVI97; and (c) simple image differencing image from NDVI97 – NDVI87 (Dif.image). All images have been subjected to a histogram equalisation contrast stretch for printing purposes..... 157
Figure 7.3 Histogram of the grey tones of the images (a) NDVI87 and (b)NDVI97 158
Figure 7.4 (a) Histogram of Dif. Image; and (b) Schematic histogram of a difference image, indicating candidate areas to be thresholded..... 159
Figure 7.5 (a) Synthetic image (SDif.image) derived from Dif.image, presenting the three categories which represent 'negative changes' (red), 'positive changes' (light blue) and 'no change' (light pink). The boundary in navy blue represents the INF; and (b) colour composite of NDVI87(R) and NDVI97(GB) 161
Figure 7.6 Interpretation of offset and gain in the linear regression differencing 162

Figure 7.7 Graphics scatterplot of the NDVI87(X axis) and NDVI97 (Y axis)and the linear regression carried out between the independent variable (NDVI87) and the dependent variable (NDVI97) 163
Figure 7.8 (a)Simple image differencing image (Dif.imageR); and (b) synthetic image (Sdif.imageR) presenting the three categories witch represent negative changes (red), positive changes (light blue) and no change (light pink). The contours in navy blue represent the INF and tows (or villages) in the region. The image (a) has been subjected to a histogram equalisation contrast stretch for printing purposes 165
Figure 7.9 (a) Normalised image differencing image (NID) (Dif.imageN); and (b) synthetic image (Sdif.imageN) presenting the three categories which represent 'negative changes' (red), 'positive changes' (light blue) and 'no change' (light pink). The contour in navy blue represents the INF. The image (a) has been subjected to a histogram equalisation contrast stretch for printing purposes..... 167
Figure 7.10 Classified change image Sdif.images 173
Figure 7.11 Image differencing obtained from the shade image of 1987 and the shade image of 1997 (Dif.image-sh). The image has been subjected to a histogram equalisation contrast stretch for printing purposes 176
Figure 7.12 Simple differencing image Dif.image –aa..... 177
Figure 7.13 Segmentation of the Dif.image-sh and a zoomed window presenting the classes: 'severe loss' (1), 'moderate loss' (2), 'slight loss' (3), 'no-change' (4) and 'biomass gain' (5)..... 179
Figure 7.14 SDif.image-sh-u (ISOSEG classification of the segmented image Dif.image-sh) 180
Figure 7.15 SDif.image-sh-s (supervised classification using MAXVER and Dif.image-sh). 181
Figure 7.16 View of *Araucaria* forest patch inside the INF 184

APPENDICES

Appendix 7.2

Figure 7.2.1 Location of the plots in the INF. Labels 1 to 4 represent the four patches of *Araucaria* forest 240
Figure 7.2.2 (a) Location of the plots in patch 3 over the PC2 of PCA from bands 4 and 7 for 1987; (b) location of the plots in patch 3 over the PC2 of PCA from bands 4 and 7 for 1997; (c) sketch of the plot and table with information about occupation of the plot with *A.angustifolia* (d) aerial-photograph of the same patch and some the natural clusters of *A.angustifolia* in the *Araucaria* forest 241
Figure 7.2.3 Allocation of the plots in patch 4. Images from 1997 and 1998 PC2 of PCA from bands 4 and 7, sketch of the plots, statistics about the plots and the aerial-photograph showing the areas with more concentration of *A.angustifolia* 242
Figure 7.2.4 Evolution of deforestation of *Araucaria* forest in the neighbourhood of the INF: (a) aerial photograph – (05.1980); (b) TM band 5 (06.1987); and (c) TM band 5 (08.1997). The images have been subjected to a histogram equalisation contrast stretch for printing purposes. The line in red encloses the devastated area 245

Chapter 1

INTRODUCTION

The present study examines the applicability of satellite remote sensing (RS) methods for assessing and monitoring forest types in an area of the *Araucaria* forest region of Brazil.

It was initiated because *Araucaria forest* has been harvested at very high rates for so long that only a few remnants can be still considered old-growth forests. *Araucaria angustifolia* (Bert.) O.Ktze is the predominant species in the biome and is the preferred species in the harvesting process of natural areas. The majority of *Araucaria* forest patches nowadays present low abundance of its most important species and forest degradation is certainly occurring largely in the region.

The aim of the approach is to investigate: 1) if the conifer *Araucaria angustifolia* can be spectrally distinguished, using satellite remotely sensed environmental data, from the other deciduous forest types; 2) if it can be spectrally discriminated from other conifers (pines) when in pure stands; 3) if the selected forest types in the region are distinguishable at pixel level; 4) if it is possible to detect and identify targets (*A.angustifolia*) at sub-pixel level; 5) if RS methods can provide analytical procedures that are able quantitatively and qualitatively to: a) characterise the forest types; and b) to detect changes in the forest cover.

1.1 *Araucaria* forest status and problems

The loss of biodiversity is nowadays a concern for natural resource managers, conservation biologists, governments and the general public; in short, for society at large. The major concern that guided this study was the conservation of *Araucaria* forest, an endemic biome of southern Brazil. There is a recognition that human subsistence depends on vegetation either directly or indirectly for food, shelter, fuel, and health. However, the adoption of development models that are incompatible with natural resources conservation, resulting in intense and prolonged exploitation of natural resources has prevented the maintenance of sufficient reserves to ensure sustainable production and maintenance of gene stock in the *Araucaria* forest.

According to Machado *et al.* (1979), the indiscriminate logging through decades for timber extraction and expansion of agricultural frontiers reduced the original *Araucaria* forest in Paraná-State from 73,780 km^2 (1890) to only 3,166 km^2, in 1977. There is no updated information specific to the *Araucaria* forest biome. According to SOS Mata Atlântica (1998), however, the deforestation

rate between 1990 and 1995 was 0.93% a year, equivalent to a yearly average of almost 170 km^2 in the whole State.

Araucaria angustifolia gives its name to the biome *Araucaria* forest because it is the dominant species (between 15 and 40% of the forest trees). However, the forests do, also, support a complex, variable and regional ecosystem composed of many species, some of which are endemic to this forest type.

Because of the model of development adopted in the region, the remaining patches were becoming poorer (in terms of commercial tree species with emphasis to *A.angustifolia*) at an expressive rate due to shifting cultivation, uncontrolled logging and unplanned land use. To minimise the problem, in the 1960s some plantations with the species *A.angustifolia* were established. The estimate is that around 200,000 ha were planted with *Araucaria angustifolia*. Growth of individual trees has been much faster in plantations than in natural areas. However, the expansion of plantations has declined sharply to about 300 ha each year because of the successful introduction of *Pinus elliottii* Engelm. – slash pine, *Pinus taeda* L. – loblolly pine and *Eucalyptus* spp. The exotic species are less demanding and have higher growth rates (Shimizu *et al.* 1980). Preservation of the valuable genetic stock was established *ex situ* although tree improvement programmes with *Araucaria angustifolia* have remained limited.

Araucaria forest occurs naturally throughout Southern Brazil, extending from latitudes 19°30'S to 31°30'S and longitudes 41°30'W to 54°30'W, including part of the province of Misiones in Argentina. Because of the similar climatic conditions to their origin, temperate pines were introduced in *Araucaria* forest region in the late 1960's. Most were planted through the policy of fiscal incentives for the Brazilian Forest Sector (from 1966 to 1987) with the objective of supplying demands of the pulp and paper sector. Due to easy adaptation, fast growing and present developed technology, pine plantations play an important role in the regional economy. Nowadays the Brazilian forestry sector is responsible for 700,000 direct employees. Exports were around £ 2.5 billions in 1996. There are more than one million of hectares of pine plantations in Southern Brazil. The reason for the inclusion of *Pinus elliottii* and *Pinus taeda* in the present study is that frequently those forest types share the same *niche* of *Araucaria* forest. In the region, pine species are planted contiguously to *A.angustifolia* plantation stands (and/or to the natural patches of *Araucaria* forest). Previous experiences pointed to the fact that the three conifers (*A.angustifolia*, *P. elliottii* and *P.taeda* has been usually planted and managed using the same management strategy. The stands are treated to have uniform structure, being planted in regular lines and have similar initial number of trees per ha. Therefore, they might present similar spectral responses and or texture properties, eventually.

1.2 The study area

The Irati National Forest (INF) is administered by IBAMA (Brazilian Institute of the Environment and Renewable Resources). The INF central co-ordinates are latitude 25°21'S and longitude 50°35'W and it occupies an area of 3,495 ha within the *Araucaria* forest. The area was chosen because it is the only federal protected area with remnants of *Araucaria* forest in Paraná State. Actually, the INF is a Federal Productive Unit, where the management is allowed, and eventually expected. However, no selective or clear cutting has been performed in the natural forest in the last 50 years. The field plots of the Forest Inventory carried out in the INF were located by GPS (Global Positioning System) and were utilised to analyse the relationship between structural information (biotic factors) and spectral response of four main forest types: a) slash pine; b) loblolly pine; c) *Araucaria angustifolia* plantation; and d) *Araucaria* forest. Fragments of old-growth and secondary forests and small farms, where man-made forests are established contiguously to agriculture and pasture, composed the surrounding area of the INF.

1.3 Remote sensing of forest resources

Remote sensing (RS) is the general term for the field of study concerned with collecting and interpreting information about an object from a remote vantage point (USDA 1995). According to Howard (1991), RS is nowadays generally accepted as not only involving the collection of raw data, but also involving manual and automated data processing, imagery analysis and presentation of derived information.

Traditional methods for collecting land cover data were concerned with investigation of a few sampled areas. However, these methodologies have become inadequate, given the increasing need for continuous planning and management of forested areas. Entire regions rather than limited sampling areas need to be monitored at low cost and satellite remote sensing provides a good solution. It can provide up-to-date surveillance of land use and land cover patterns, and permits the derivation of thematic maps and statistics. The digital nature of satellite data makes it possible to integrate them with other thematic data in a geographic information system (GIS). Another advantage is that forest and land use mapping units can be combined with ancillary data for integrated spatial analyses. In terms of monitoring, however, the strength of RS is its ability to deal with different sources and types of information, from sources such as aerial photographs to satellite imagery, in a way that images of earlier years are compared to recent years, to measure differences in the sizes and extents of the logging or loss of forest (CCRS 1999).

Geospatial information about land use and natural resources is the backbone of effective ecosystem management plans (USDA 1995). Actually, environmental knowledge is something shared between generations. To the human ancestors, an understanding of the available natural resources meant the difference between the prosperity and the collapse of their communities. There is little doubt that this is a true statement in the present day. Natural forests are now being

altered or cleared more rapidly than any time in the past (Whitmore *et al.* 1994). Therefore, the increasing need to gather information across diverse functions, and spatial and temporal scales, has resulted in wider and more efficient use of Remote Sensing, Geographic Information Systems (GIS) and associated technologies as Global Positioning System (GPS) (USDA 1995).

1.3.1 State of the art of RS investigations of *Araucaria* forest

Three major Regional Inventories at *Araucaria* forest region have been carried out (1966, 1974 and 1978), all based on field sampling and aerial photographs. Since 1978, there have been only not integrated (local initiatives) projects done. Landsat MSS was used by Gantzel (1979), working with monochromatic images and colour diapositives. He concluded that it was possible to discriminate patches of *Araucaria* forest but, at that moment, it was not possible to differentiate *A.angustifolia* plantations from stands of *Pinus* spp. Disperati (1981) examined Landsat MSS for mapping natural stands of *A.angustifolia*, selecting two sites, Quedas do Iguaçú and Manguerinha (an Indigenous Reserve). The author could recognise patches of *Araucaria* forest in the MSS bands 6 and 7. A supervised classification was performed for mapping *Araucaria* forest and reforestation stands, using MSS bands 5 and 7 scanned and digitised by microdensitometer; the approach successfully mapped patches of *Araucaria* forest but failed discriminating reforested areas. In another experiment, a small area of *Araucaria* forest was selected, when two different crown densities were selected. Supervised and unsupervised classifications were performed and the author concluded that only dense stands of *Araucaria* forest could be mapped accurately.

The most recent effort uniting RS and *Araucaria* forest was also performed in the Manguerinha area (Lange Jr. *et al.* 1994). According to the authors, it was one of the last large reserves of *Araucaria* forest in the planet and has been subject of polemics, with many judicial actions for its possession. The authors proposed the monitoring of the area by using SPOT 2 satellite (image date:08/1988), in digital format. A digital classification using "Maximum Likelihood" strategy was performed and aerial photographs at a scale of 1:25,000 from 1980 served as ancillary data to help in the identification of training sites. Seven classes were identified in the aerial photographs, but two of them ("Capoeira/bracatinga") were mixed when the classification using the satellite image, due their similarities. Therefore, the final six identified classes were: a) Pluvial forest; b) *Araucaria* forest; c) altered forest; d) "Capoeira/bracatinga"; e) "Agriculture" f) burnt areas. Little difference in reflectance was found between the bands but apparently, the best results were presented by the near infrared region. Topographic effects were a major problem in the classification scheme and shadows were responsible for the confusion among *Araucaria and* Pluvial forests. The authors suggested that those limitations, inherent to digital classification should be overcome through the visual interpretation of the diapositives in the areas wrongly classified. They also mentioned that further analysis should be recommended testing the hypothesis that the variation in solar azimuthal angles implies in modifications in the standard of the canopy multispectral bidirectional reflectance. Finally, it was recommended sequencing the project, performing a field inventory and using higher spatial resolution products. Unfortunately, the project was discontinued.

No more research using RS or GIS has been subsequently repeated involving *Araucaria* forest. The present research reopens the debate about the necessity of monitoring changes in the *Araucaria* forest as a biome, and the disappearance of *A.angustifolia* as its most important species. RS and GIS provide potentially powerful tools to deal with this challenge. Since the 1960s researchers have warning of the alarming situation of *Araucaria* forest. The hope is that some of the tools developed in this thesis might be easily implemented, at low cost, by teams who could be responsible for implementing a systematic monitoring approach, which can identify quickly the problematic (which presents negative changes in vegetation) areas.

1.4 Aims of this study

The overall goal of this research was to develop methodology to the assessment and monitoring of *Araucaria* forest and ecosystems associated to the *Araucaria* forest region, using remote sensing.

The objectives of the research were:

(i) to analyse the reflective properties and to characterise the spectral response patterns of the four main forest types in the INF: (*Pinus elliottii* Engelm. plantation; *Pinus taeda* L. plantation; *Araucaria angustifolia* plantation; and *Araucaria* natural forest) in order to discriminate each forest type. Three main aspects were studied:

a) spectral responses and structural variables relationships;
b) spectral responses and species composition;
c) spectral characteristics of each feature and their separateness.

(ii) to assess the usefulness of image transformations in the process of forest types discrimination in a multispectral environment. The specific objectives were:

a) general cover-type discrimination;
b) coniferous versus deciduous vegetation discrimination;
c) discrimination between *A.angustifolia* plantation and the two species of the genus *Pinus* most frequently planted in the region, *P.taeda* and *P.elliottii*;
d) discrimination between *Araucaria* forest and other types of forest cover composed only of deciduous species (sub-tropical forest or riparian areas);
e) investigation about the discrimination of clusters of *A.angustifolia* in the *Araucaria* forest at pixel level.

(iii) to review computer-assisted methods for handling mixed pixels and to present an assessment of the usefulness or effectiveness of the Linear Mixing Model (LMM) to detect and identify targets at the sub-pixel level. The specific objectives were:

a) digital enhancement of phyto-physionomic characteristics of INF;
b) analysis of LMM capability for detecting the species *Araucaria angustifolia*, in plantation and natural stands;

(iv) to assess the ability and usefulness of Landsat TM imagery to detect temporal forest changes in the *Araucaria* forest region. The specific objectives were:

a) to obtain reliable field information about the past distribution (presence) of *A.angustifolia* in the protected area of INF;
b) to analyse the usefulness of Linear Mixing Modeling in the change detection process;
c) to determine the magnitude of any changes in distribution and extent of *Araucaria* forest in the period of 10 years (from 1987 to 1997) within and in areas adjacent to the area of INF;
d) to develop a method by which future changes in *Araucaria* forest as a biome and in *A.angustifolia* as a species could be monitored.

1.5 Structure of thesis

Chapter 1 presents an **introduction** to the problem, the importance of remote sensing (RS) for forestry and reviews previous efforts of using RS for *Araucaria* forest assessment. The **objectives** of the study are also summarised.

Following the Introduction, **Chapter 2** introduces the genus *Araucaria* and describes the overall geography of the region where the study sites were located. The description gives **summarised information** on the historical exploitation, anthropological aspects related to the fragmentation process, climate, geology, soils, and vegetation. This is the background information for the other chapters. The chapter also details the field data selection, collection, and aspects of the adopted inventory strategy and describes the **characterisation** of Araucaria forest.

Chapter 3 focuses on a review of remote sensing **processing procedures**. It stresses the importance of preparing the multitemporal set to be compared, dealing with atmospheric affects and illumination differences. It also presents the methodology adopted in the data **selection**. The fundamental aspects of geometric and radiometric restoration are discussed.

In **Chapter 4** criteria for separateness (or **discrimination**) of different forest types, using RS are analysed. Reflective properties of the present forest types are discussed. It is shown that *Araucaria* forest and *A.angustifolia* plantation are spectrally distinguishable from each other and from other conifer plantations such as *P.taeda* by means of the typical reflectance curves of each forest type built with data from the National Forest.

In **Chapter 5** the contribution of image transformations, based on dimensionality reduction of the feature space is analysed. The chapter analyses how **Vegetation Indices** (VIs) can improve the process of discrimination between forest covers in the study area with emphasis on *A.angustifolia* separateness; NDVI is investigated. The **orthogonal transformations** are represented by Principal Component Analysis (PCA), with the objective of enhancing and mapping the spectral differences or contrasts between the four main forest covers in the INF. Standard PCA using covariance matrix was tested and selective approaches were designed in order to explore highly correlated bands versus poorly correlated bands. The last approach involved the fraction images brightness, greenness and wetness of the Tasselled Cap (Kauth-Thomas).

Chapter 6 provides a review of computer-assisted methods for handling **mixed pixels** and an assessment of the usefulness or effectiveness of these methods for detecting *Araucaria angustifolia* as a species in natural stands of *Araucaria* forest. The chapter also focuses the application of **Linear Mixing Model** (LMM) and its fraction images derived from Landsat TM. The basic principle of the model is that the three main components in forested areas are vegetation (canopy), soil, and shadow. The objective is to verify if these generated images are capable of discriminating between genus and species and types of forested areas. The LMM is based on sub-pixel analysis and its capability of locating *A.angustifolia* as an overstorey indicator species in the natural forest is investigated. This is an essential question in the *Araucaria* forest conservation analysis and *A.angustifolia* is being considered as a target species.

Chapter 7 includes the investigation of changes on forest cover areas over a period of time (1987 and 1997) as a tool for monitoring changes (**change detection study)** in vegetation. Image differencing is the central piece of the procedure. Shade difference image derived from the LMM (analysed in **Chapter 6**) is segmented and a monitoring system is sequenced, using unsupervised classification. Supervised classification is also performed over the product of image differencing of TM band 5 and the fraction image of shade.

Finally, discussions including conclusions and recommendations for future work are in **Chapter 8**. The interactions between the different approaches are considered together with recommendations for future research investigations using new technology.

Chapter 2

THE ENVIRONMENT OF THE STUDY AREA

2.1 Introduction

The study focuses on an area within the *Araucaria* forest biome located within the first plateau of Paraná State, with the approximate co-ordinates: latitude 25°15'S to 25°30'S and longitudes 50°15'W to 50°45'W, and it covers approximately 675 km^2. It was chosen because it is a representative area of the biome and it includes the Irati National Forest (INF) which has natural stands and forest plantations and is the only protected area with remnants of *Araucaria* forest. **Figure 2.1** shows the INF and neighbouring areas.

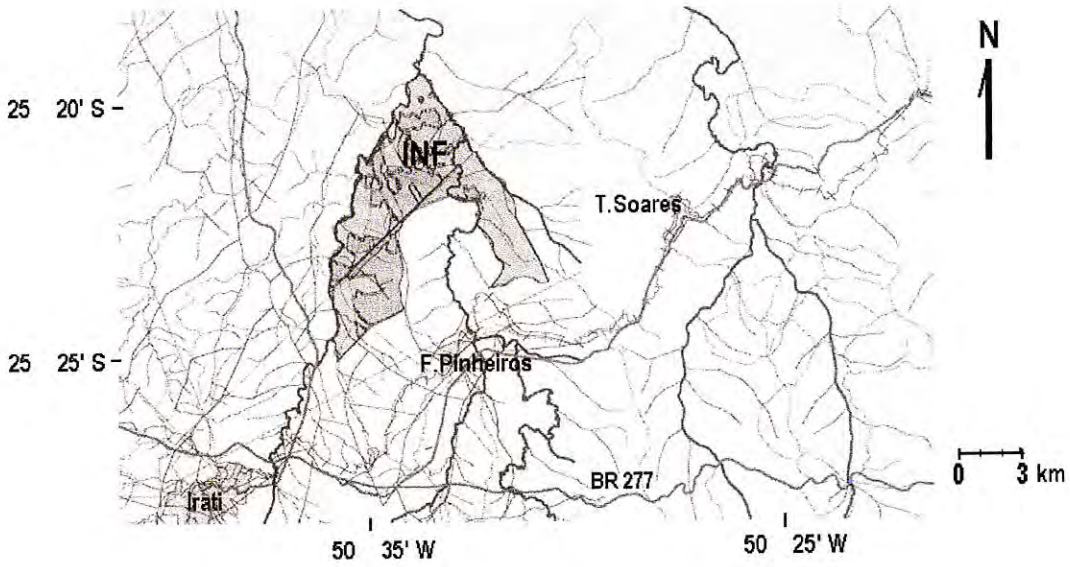


Figure 2.1 The study area showing Irati National Forest (INF) and adjacent areas. Irati, F.Pinheiros and T.Soares are the main urban areas in the study area

2.2 The genus *Araucaria*

The genus *Araucaria* is endemic to the Austral Hemisphere and is represented by 16 species that are shared among Australia, New Guinea and South America (Silva 1997). In South America, the genus occurs in the temperate latitudes (*Araucaria araucana*) and tropical (or sub-tropical) latitudes (*Araucaria angustifolia*). Southern hemisphere conifers have often been characterised as relicts or living fossils, which are gradually being replaced by angiosperms. In Brazil, it has been noticed that

A.angustifolia is not adequately regenerating. It seems that it is a generalised phenomenon that also is happening in the south-west Pacific (Veblen *et al.* 1995).

2.3 *Araucaria* forest region

Araucaria forest occurs naturally throughout southern and south-eastern Brazil, extending from latitudes 19°30'S to 31°30'S and longitudes 41°30'W to 54°30'W, including part of the province of Misiones in Argentina. The original *Araucaria* forest (**Figure 2.2** and **Table 2.1**) covered an estimated area of 200,000 km² (Maack 1950, Hueck 1953 and Golfari 1971).

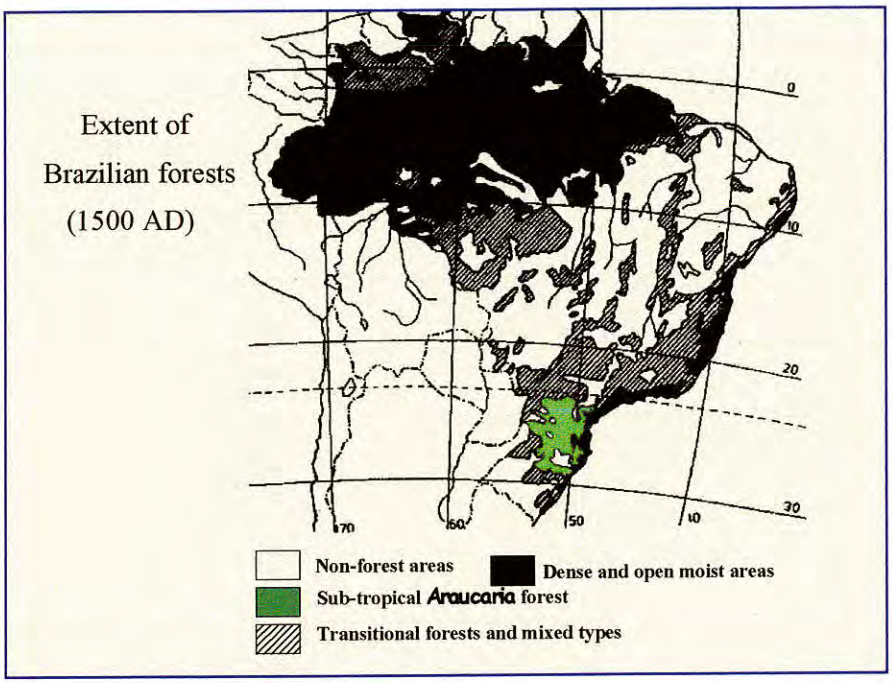


Figure 2.2 Original geographic distribution of Mixed Ombrophilous Forest with *Araucaria angustifolia*
Source: Brown Jr. *et al.* (1994)

Table 2.1 *Araucaria* Forest: Original geographic distribution

COUNTRY	STATE	AREA (km²)	MAX.ALT. (m)	MIN.ALT. (m)	% AREA
BRAZIL	Paraná (PR)	73,280	1,300	500	39.85
	S.Catarina (SC)	56,693	1,300	500	30.83
	R.G.do Sul (RS)	46,482	1,200	300	25.28
	São Paulo (SP)	5,340	1,800	700	2.90
ARGENTINA	Missiones	2,100	500	200	1.14
TOTAL		183,895	1,800	200	100.00

Sources: Maack 1950; Hueck 1953

The expansion of the agricultural frontier in the past decades has resulted in a rapid process of forest fragmentation of *the region* (**Figure 2.3** and **Table 2.2**). *Araucaria* forest is nowadays an extremely fragmented landscape and these fragments are more concentrated in the south of Paraná State (PR) and the central part of Santa Catarina State (SC). There is no updated

information about the whole region. However, **Appendix 2.1** presents the latest version map of the forest remnants and ecosystems associated to the Mata Atlântica (Atlantic Rain Forest) domain for the period 1990-95 for the Paraná State. The map actually includes *Araucaria* forest and was produced recently, using visual interpretation of Landsat TM imagery by SOS Mata Atlântica and INPE (SOS Mata Atlântica 1998).

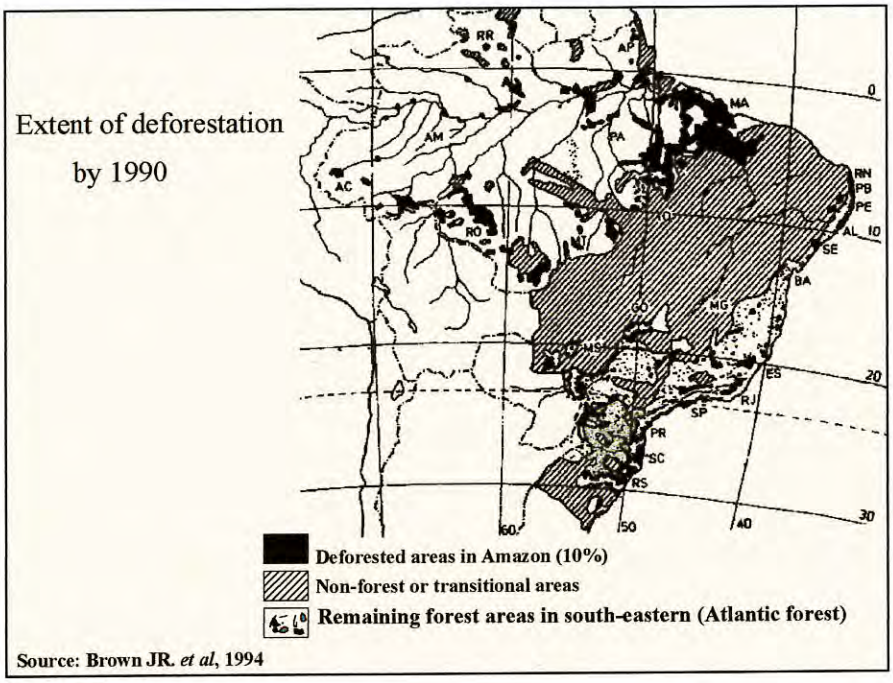


Figure 2.3 Extent of deforestation in the Atlantic forest domain by 1990.
State names abbreviated to two letters: Paraná (PR), Santa Catarina (SC) and Rio Grande do Sul (RS).
Source: Brown Jr. *et al.* (1994)

Table 2.2 Devastation sequence of *Araucaria* Forest in Paraná State

YEAR	ARAUCARIA FOREST (ha)	% REMAINING	% STATE AREA
1890	7,378,000	100.0	37.00
1930	3,958,000	54.0	20.00
1937	3,455,400	47.0	17.00
1950	2,522,400	34.0	12.50
1955	2,203,200	30.0	11.00
1960	2,043,200	28.0	10.00
1965	1,567,700	21.0	8.00
1972	433,500	6.0	2.00
1978	316,620	4.0	1.50
1988 *	151,620	2.0	0.75
1990 *	118,620	1.6	0.59
1996 *	36,120	1.0	0.20

(Sources: Maack 1968, Dillewijn 1966, FUPEF 1978 and Gubert Filho *et al.* 1991)
* Estimated by Gubert Filho (1991)

2.4 Conservation Units

The Paraná State Conservation Units represent 2.41% of the total State surface and 2.89% of the original biome surface (Jacobs 1997). Although the *Araucaria* forest is so important to the regional economy, according to the author, there are no protected areas administered by the State. The implementation of a specific Conservation Unit in old-growth remnants of *Araucaria* forest is recommended by many authors (Maack 1968; Gubert Filho *et al.* 1991; Campos 1996; Jacobs 1997; 1999) and it is a priority to create at least one.

2.4.1 National Forests

The first Brazilian Forest Code was created by the Ordinance-law 23.793 of January 23, 1934. Through it, the Federal Government received the competence of legislating on the matter of water, forests, hunting, angling and their exploitation. States would legislate only complementarily, to account for local peculiarities and supply gaps or deficiencies in the Federal legislation. A great result in direction to the protection of the forest patrimony happened with the creation of the Paraná-pine National Institute, in 1941. This Institution regulated the indiscriminate logging (and the devastation) and established norms for forest production through the limitations in the legislation of sawmill industries. In this period the first National Parks were created. The first of them was the National Park of Açungui, in Paraná State, in 1942. Many others were created in the sequence, among them the Parque Florestal (Forest Park) "Manoel Enrique of Silva, also in 1942 whose name later became Floresta Nacional de Irati (Irati National Forest – INF). These units were the only ones to have *Araucaria angustifolia* plantation for a long period. According to Zachow (1991), even with the creation of several National Parks, co-operation between the administration and conservation sectors of the forest resources was inadequate. According to the author, drastic reductions in the forest area and in the number of forest industries, caused by the expectation of forest resources depletion, was the motivation for the creation of the Brazilian Institute for Forest Development, in 1967 and later replaced by the Brazilian Institute for Environment and the Renewable Natural Resources - IBAMA, in 1989.

The National Forests are conservation units created by the Federal Government and administered by IBAMA. They are supposed to be in constant production, allowing a harmonic agreement between people and nature, through sustainable development. Nowadays there are 39 NFs (National Forests) in approximately 12.6 thousand acres distributed in the national territory as: a) North Region –24 Units; b) Northeast Region – 01 Unit; c) Southeast Region – 05 Units; and d) South Region – 09 Units. The NFs were designed to ensure a rational utilisation of the natural renewable resources, production and research, respecting the mechanisms of ecosystem support. Their goals are: a) multiple use of forest resources; b) maintenance of biological diversity; c) area protection against erosion and sedimentation; d) recuperation of degraded areas; e) development of environmental education plans; f) protection of genetic resources; g) support and diffusion of production techniques and/or a rational use of forest resources; h) protection of historical and

archaeological areas; i) development of recreation activities and ecotourism; j) development of technical-scientific activities; k) conservation of the hydro-resources; and l) demonstration of the viability of renewing natural resources.

2.4.1.1 Irati National Forest (INF)

INF lies in Município of Fernandes Pinheiro (município is a political unit corresponding to the British counties). Its central co-ordinates are latitude 25°21'S and longitude 50°35'W and it is in close proximity (5 km) to Irati, the main urbanised area of the region. The activities in the INF are: woody exploitation, seedling production, and mate-tea production. *Ilex paraguariensis* St.Hilaire (erva-mate) is a forest species that is used to prepare the mate-tea, a very common beverage in South Region. It is one of the most important species of the biome. Honey and resin are also produced but considered less important.

Previously called Parque Florestal "Manoel Enrique da Silva", INF was created in 1942 and enlarged afterwards when the area Cerro Verde was annexed. In the earlier times, it was administered by Paraná-pine National Institute. Nowadays, National Forests and National Parks are administrated by IBAMA (Brazilian Institute for Environment and Renewable Resources), which also determines the governmental policy for the Forest Sector. **Figure 2.4** presents the INF in its environment. The National Forest contours, main roads and rivers were digitised (by the author of this work as a training component) and georeferenced; the image in the background is Landsat TM band 3, from 1987, showing the proximity of the Irati town.

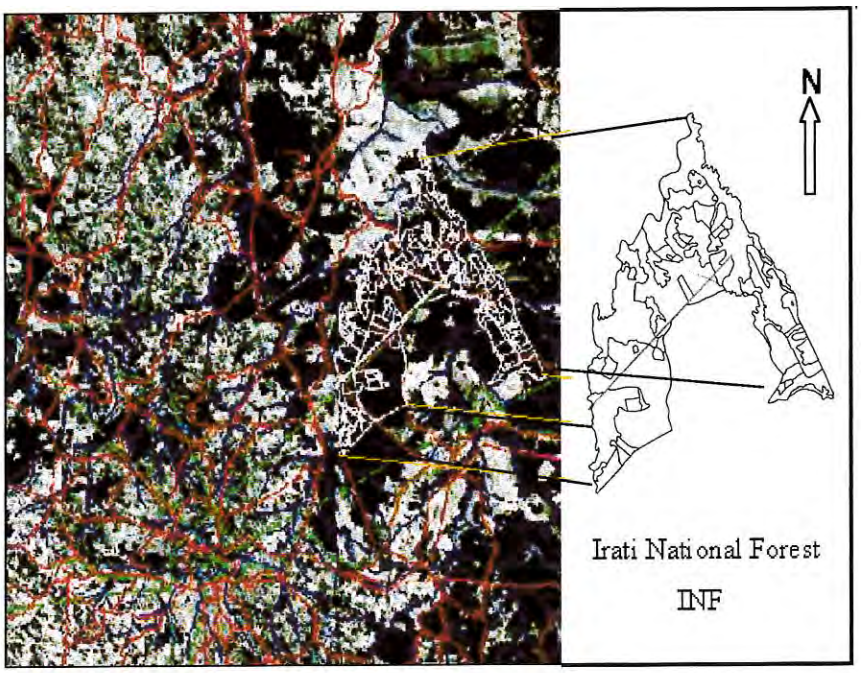


Figure 2.4 Irati National Forest and part of study area, showing Irati town in the down left corner. Main roads are in red and main streams in blue.

2.4.1.1.1 Geology, geomorphology, soils and climate

The geology in INF is characterised by the presence of sandy clayey sediments, belonging to the Palermo group (Tubarão series), from the Carboniferous age (superior layer), with great occurrence of sandstone in the northern part. The INF area, formed by several sedimentary rocks such as sandstone and stratified clay rocks, is characterised by gentle elevations, showing level differences of 100 m. The Imbituva River is at 790 m and its highest part reaches 922 m of altitude, falling to 810 m at the Coxinhos River. The region has the gentle rolling relief characteristic of highlands, where the level differences were formed by distinctive erosion. At the northern part of INF, in Cerro Verde area, the sandstone, due to the erosion process, forms eventually small valleys with steep slopes. Water breaking through the permeable sandstone, coming from the subsoil rocks with greater amounts of clay, causes some cave-in, forming small valleys typical of these landscapes.

In areas formed by sedimentary rocks, soils are deep, especially when protected by vegetation canopy. They present usually dark or grey colours at the surface horizons, due to the presence of organic matter that change to yellow, brown or red colours in the subsoil. They are formed from sandstone that presents granular texture, increasing compaction when formed from stratified clay and from diabasic rocks. Paraná-pine prefers well-structured soils, but it can be found on different kind of soils. The soils from INF are extremely acid (pH less than 5). INF does not have a meteorological station. The closest units are: a) Ponta Grossa Station (around 70 km north-east); and b) and Guarapuava Station or União da Vitória Station, both distant more than 100 km (west and south-west, respectively). All the stations indicate that the climate is Cfb type, according to Köeppen, that means well distributed rain all over the year and annual mean temperatures lower than 22°C and stable annual rainfall distribution (Goubert Filho 1991). The hottest months of the year are December and January. The coldest ones are June, July, and August, with occurrence of heavy frosts during the period.

2.4.1.1.2 Vegetation

Irati National Forest total area is 3,495 hectares within the *Araucaria* forest. It has areas with patches of natural forest and stands with forest plantations (**Tables 2.3** and **2.4**).

Table 2.3 Vegetation of the INF

Vegetation	Area (ha)	(%)
Plantations	1,308.71	37.45
Native Forest with A.angustifolia	1,272.90	36.42
Swamps	322.69	9.23
Subtropical Forest	235.19	6.73
Secondary Growth Forest (Capoeira)	184.13	5.27
Non forested areas	167.19	4.78
Research	4.19	0.12
TOTAL	3,495.00	100.00

Source: Irati National Forest Inventory – Vol.2, FUPEF (1989)

Some of the exotic species such as *Cunninghamia lanceolata, Cupressus lusitanica, Eucalyptus sp* and *Pinus pinaster* were only planted in small stands and are not going to be included in the present study. The stands of *Acacia* sp presented high mortality rate and were abandoned as experiments.

Table 2.4 Forest plantations composition

Species	Area (ha)	(%)
Araucaria angustifolia	433.13	33.09
Acacia sp	36.32	2.78
Cunninghamia lanceolata	0.37	0.03
Cupressus lusitanica	2.75	0.21
Eucalyptus sp	6.06	0.46
Pinus elliottii	677.08	51.74
Pinus taeda	96.68	7.38
Pinus pinaster	3.63	0.28
Pinus sp	2.81	0.22
Other	49.88	3.81
TOTAL	1,308.71	100.00

Source: Irati National Forest Inventory – Vol.2, FUPEF (1989)

2.4.1.2 INF adjacent area

The area selected for this study is located at the junction of several important infra-structural lines that strongly influence the economic development of the region. BR 277 (from Curitiba to Guarapuava) crosses in full the study area horizontally, which includes INF as its central part. Fragments of secondary forests and small farms, where man-made forests are established with agriculture, compose the remaining area. Farming is modernising, with a shift from subsistence-oriented farming to market-oriented farming and a number of medium-size farms; a number of extensive farms in the surrounding of Irati were established.

The region is an extremely fragmented environment. At its east-side (Fernandes Pinheiro county, to which area the INF is currently linked) there are medium and large properties with mechanised agriculture, intensive pasture activities and production of the mate-tea (*Ilex paraguariensis*). In Fernandes Pinheiro there is a large mate-tea factory-processing centre. A forest nursery that belongs to Paraná Environmental Institute (IAP) is in the neighbourhood. Surrounded by a narrow corridor, it is located an Experimental Station administrated by IAPAR (State of Paraná Research Institute) that has 1,750 ha of which 1,000 ha are native forest. Still in the same area but a little further in the north direction, planted areas from four different private enterprises are predominant. One of them has advanced fish culture, rich pasture and a farm-hotel. The Agrovila area (100 ha) can be the object of barter between Irati County and Paraná Government to be transformed into a State Forest with the IAPAR area.

The west area, smaller in extension, has the Antas River as its border. In the area, small farmers plant beans, corn, tobacco, and a milk production industry is in the neighbourhood. Nearby it can

be find a private area, which belongs to a brickyard (Olaria Van Laars), and still has an old-growth *Araucaria* forest patch and a pine plantation that belongs to an industry. In one of the corners of the INF there is a 25 ha area, which is a private property and it is almost totally deforested. Irati town has an area of 1,000 km^2 and estimated 51,000 inhabitants in 1997 (Prefeitura Municipal de Irati 1999). Distance between the urban area and the INF is 10 km.

2.5. Fieldwork – Forest Inventory

In January 1988, FUPEF (Forestry Research Foundation of Paraná) was contracted by IBAMA to carry out a Forest Inventory. The fieldwork was part of a large objective, which consisted of updating information about the renewable resources in all the Federal protected areas and test a methodology for the Continuous Forest Inventory at national level, to be carried in future. Hence, the methodology involved the installation of permanent plots instead of temporary plots. The methodology adopted is described in (IBAMA 1989). Natural stands of *Araucaria* forest were inventoried using the systematic sampling, with the sample units distributed proportionally to the size of the different patches that composed the area. Size, shape and spatial distribution of the plots is detailed in **section 4.4.3.2.** Pine and *A.angustifolia* plantation were inventoried by the stratified random sampling when the strata were the stands. The reason for different approaches were that in the natural stands the location and installation of random plots should be cost prohibitive if large number of plots were expected, due the heterogeneity of that forest type. The systematic sampling is objective and as the distribution of the sampling units is regular and all parts of the population are sampled, the precision of the parameters estimated is usually high (Phillip 1994). The restriction of randomisation can be outlined and approximate estimates of accuracy may be calculated. In the case of plantations, homogeneous areas are normally stratified by species and age. This scenario allow the researcher to choose measuring total areas that optimise the relation cost/benefit and still achieving a desired precision. Size, shape and number of plots per stands are detailed in **section 4.4.1.2** for pine plantation and in **section 4.4.2.2** for *A.angustifolia* plantation.

Another fieldwork was taken in 1998 and consisted of determining the exact position of each plot allocated in the 1987 Forest Inventory. The plots central point was recorded by using the "Global Positional System" (GPS) from GARMIN which allows precision around 1.40 m. Constraints such as accessibility of the stand, difficulties to locate the plot and weak GPS signal under canopy were frequent and were responsible for the final number of considered observations. Details about GPS and the adopted methodology are described on the **Appendix 2.2**. The following ancillary data were available: a) aerial photographs - panchromatic, covering Irati National Forest and surroundings - Date: 05/05/80 - Scale: 1:25,000 (photocopies of the original B&W photos); and b) maps: cartographic maps, in scale 1:100,000 and 1:50,000, from the Geographic Service of the Federal Government, available since 1966 and vegetation map from the Irati National Forest, made in 1986.

2.5.1. Pine and *Araucaria angustifolia* plantations

The fieldwork methodology can be summarised as: a) sampling - random sampling was adopted within the stands with stratified sampling at species and age levels. Sampling units were allocated by Cartesian System; b) sampling intensity – the number of sampling units was proportional to the area of each species and age. The minimum number of plots was 3 and the maximum 15 sampling units per age, at species level; c) sampling unit – the dimension of each plot was 20 x 30 m (600 m^2). As mentioned before, the reason for using random sampling was the homogeneity that stratification by species and age can provide. This approach diminishes the need of intense sampling without diminishing the desired accuracy. The allowed maximum sampling error (standard error) was plus or less 10% (p = 0.10) of the estimated mean volume.

The sample units were identified as permanent plots, with a small label indicating their position in the stand and their corners were painted with yellow ink. Six dominant trees were identified with a plastic tag and forked or dead trees inside each plot were identified. Measurements of GBH (girth at breast height) equal to or greater than 15 cm of diameter DBH were taken and later converted to DBH. The only reason for measuring GBH and not DBH was the non-availability of enough diameter tapes. Total heights of the 25 first trees were also measured, apart from the dominant trees; and d) volumes – taper functions and volume equation were developed from trees specially cut down by diameter classes. For both genus (Pinus and Araucaria), the individual bole volume was estimated by the Smalian method. 121 trees of *P.elliottii*, 50 trees of *P.taeda* and 114 trees of *A.angustifolia* separated into DBH (diameter at breast height) classes, were specially felled for this purpose. Eleven mathematical models were tested and the Schumacher-Hall equation was selected. It presented high multiple coefficient of determination (R^2 = 0.98 for *Araucaria angustifolia*, 0.95 for *P.elliottii* and 0.97 for *P.taeda*) and low standard estimate error in percentage (sxy% = 8.47 for *Araucaria angustifolia*, 9.98 for *P.elliottii* and 7.64 for *P.taeda*). It also presented results for F and Durbin-Watson tests highly significant. That model is expressed as:

$V = b_0 + b_1 \ln DBH \prod b_2 \ln height.$

Were V = volume, in m^3; ln = natural logarithm; DBH = diameter at breast height, in cm; and height = defined by the hypsometric relationship between diameter classes and the height (in m) of the first 25 trees in each plot.

Details about data effectively included in the study are presented in **Chapter 4, sections 4.1.1** and **4.4.2** for Pines and paraná-pine plantations respectively, where the results of the experiments about spectral responses and structural variables relationships are discussed.

2.5.2 *Araucaria* forest

For the native forest, a different strategy was adopted. The forests always present certain level of complexity and the economical aspect of the fieldwork must always be considered. In this case, the defined methodology was: a) sampling process: systematic approach, with sample units distributed proportionally to the size of the different patches that compose the area. Systematic sampling was

considered in the case of the inventory of natural areas to facilitate the location of the plots, which is less time consuming using this approach. Baselines are defined and the plots located at regular intervals. The proximity and regularity of the plots installed like transects also helps in the analysis of the curve species-area and frequency of species in the different patches. Size of plots was also different in natural stands and plantations since the homogeneity of the planted stands allows small plots to be adequate. In the case of the natural forest, the plots were distributed in strips (east-west orientation) approximately 500 m apart; b) dimension of the sample units: 15 x 100 m, located at intervals of 100 m. Using this procedure, some baselines received several units and others just one or two. Some 93 plots were installed as permanent plots. In each plot, the following information was obtained: a) identification of the species; b) all individuals with GBH (girth at breast height) equal to or greater than 60 cm (20 cm of DBH); c) timber height (bole length); d) total height of the paraná-pine trees; e) evaluation of the stem quality; f) phytosociological position; g) characteristics of the understory, access and slope. Details about the analysis are presented in **Chapter 4**, **section 4.4.3** and **Chapter 7**, **section 7.61**, **7.6.2**, and **7.6.3**.

The software used for those calculations was PLANNAT, developed by Silviconsult Engenharia. The list of identified species in the inventory is in **Appendix 2.3**. Volume was calculated by specific equations developed previously for the paraná-pine (conifer) and for all the other species (deciduous). **Table 2.5** presents the summary of the sampling statistics. The equations considered by the software were:

$V = b_0 + b_1 (DBH^2 * height)$ for paraná-pine and some Lauraceae/Meliaceae (different coefficients for each group);

$V = b_1 * ((DBH \prod 2.14581) * (height \prod 0.71915))$ for the other species.

Were V = Volume, in m^3; ln = natural logarithm; DBH = diameter at breast height in cm; and height = height, in m.

Table 2.5 Summary of the sampling statistical analysis

Estimates	Results
Mean Volume	
- m^3/sample unit	18.5
- m^3/hectare	123.1
Basal Area	
- m^2/sample unit	2.9
- m^2/hectare	19.0
Confidence Interval	
- m^3/sample unit	111.8 < x < 134.4
- m^3/hectare	168,543.6 < x < 202,689.8
Number of Trees	
- per sample unit	33
- per hectare	219
Statistical analysis	
- standard deviation (s)	68.20
- coefficient of variation (cv%)	55.41
- standard error (sx)	6.82
- error limit % (sx%)	9.15
- sample units needed	84

2.6 Characterisation of *Araucaria* forest at INF

Araucaria forest is typically dominated by the species *Araucaria angustifolia* (paraná-pine, the symbol of Paraná State). The species was, for a long period, one of the most important items of Brazilian export, providing high quality timber for general construction, framework, furniture, veneers, pulp and many other uses (Shimizu & Oliveira 1980). *Araucaria angustifolia*, the only species of the genus in Brazil, is one of only two native coniferous trees in the country [the other is *Podocarpus lambertii* Klotz. (pinho-bravo)]. It is a straight tree with almost cylindrical bole and commonly grows to heights of 30-50m with a 2m diameter at breast height (DBH). This partly explains why, until recently, *A. angustifolia* was a major resource for social and economic development in its native region through the so-called 'Pine-wood Cycle', which occurred mainly between 1930 and 1950. It may also explain why, during this period, the railway system was completed, ports were activated and industrialisation occurred (Carneiro 1979). In natural stands, the species can comprise more than 30% of the forest trees. **Figure 2.5** illustrates the dominance of the species in terms of number of trees and compares the presence of the paraná-pine and other important species at the INF and in the Reserve of Caçador, in Santa Catarina State.

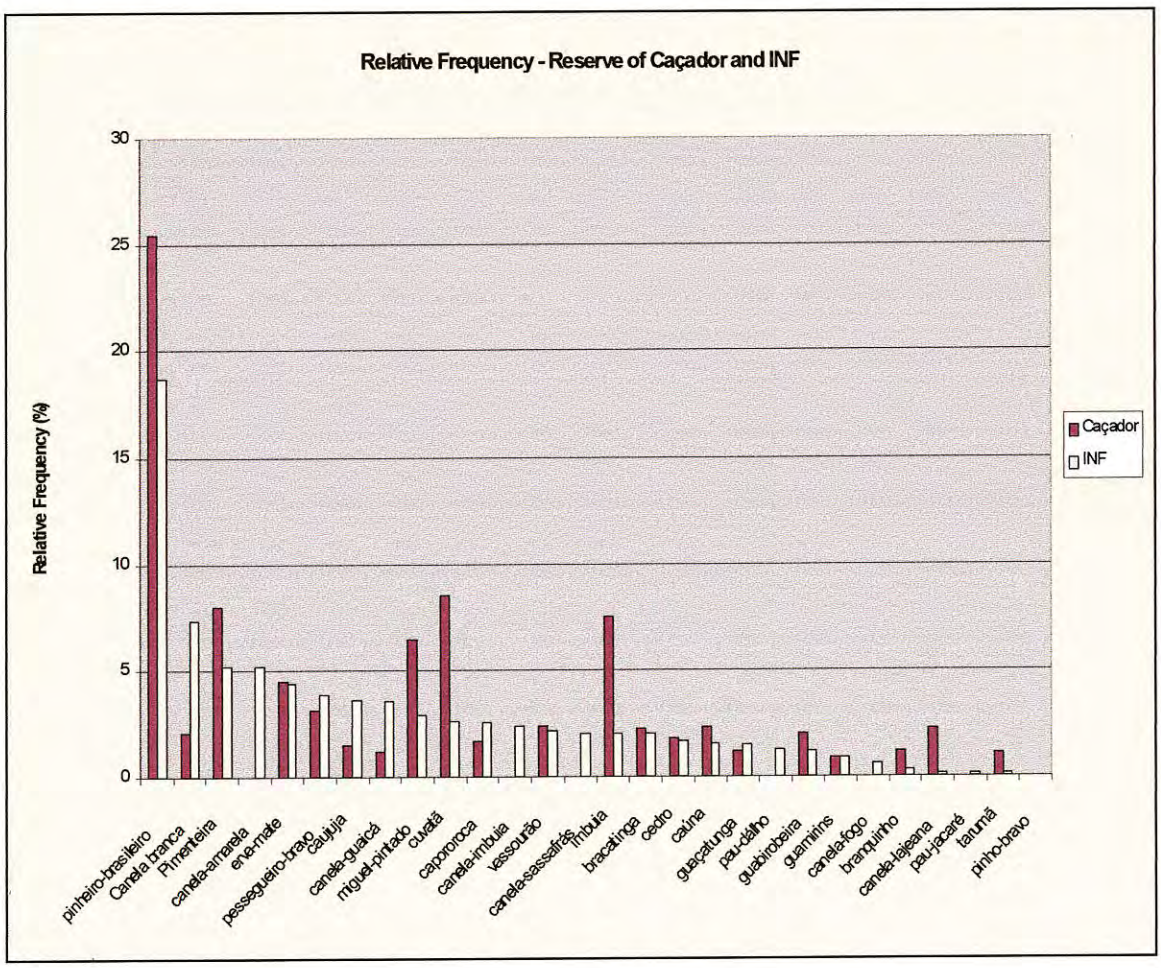


Figure 2.5 Main species inventoried at INF and at Reserve of Caçador.
Source of data for Reserve of Caçador: Silva *et al.* 1997.

Reserve of Caçador is one of the last almost untouched areas with the presence of *Araucaria* forest and presents more than 25% of the number of trees as *A.angustifolia*. At the INF it was found that the presence of *A.angustifolia* varies in the different patches, from 3% to 8% of the total number of trees in the plots with lower concentration of the species to around 40% of the total number of trees in the plots with more concentration of the species. In average, 18% of the trees are paraná-pine. Longhi (1980) found that *A. angustifolia* represented 42% of the number of trees and 50% of the basal area. Galvão *et al.* (1989) and Oliveira *et al.* (1982) found similar results.

The forests, however, also support a complex, variable and regional ecosystem composed of many species, some of which are endemic to this forest type (Klein 1963); in the field inventory, 80 different forest species were found. This result is probably only possible in when protected areas are inventoried. The majority of the private remaining areas are becoming poorer at an alarming rate due to shifting cultivation, uncontrolled logging and unplanned land use. *Ocotea porosa* (Nees.) L. Barroso (imbuia) is still relatively well represented at Reserve of Caçador but presents less than 2.5% of the number of trees at INF. *Ocotea pretiosa* (Nees) Mez. (canela-sassafrás) and *Cinnamomum sp* (pau d' alho) are not represented in the Reserve of Caçador, what intensifies the importance of the INF as a protected area. Both species are considered vulnerable by the legal Brazilian authority (IBAMA). **Table 2.6** presents a comparative summary of differences in Basal Area (BA) and number of trees in some of the last remnants of *Araucaria* forest in Southern Brazil. The numbers are not easily comparable because of differences in methodology (minimum DBH being 10 or 20 cm). However, **Table 2.6** intends to show that the INF, even not being harvested in the last 50 years presents some differences when compared with areas as the Reserve of Caçador (in Santa Catarina State) and other areas. It presents trees that are smaller in dimension than other two preserved areas (Caçador and S.J.do Triunfo), but still have number of trees that might be considered similar to those presented in those areas.

Table 2.6 Summary of Basal Area (BA, in m^2) and number of trees (Trees) (typical values per hectare (ha^{-1}) for trees with more than 10 or 20 cm DBH inventoried in different areas with *Araucaria* forest

	BA >20	BA >10	Trees > 20	Trees > 10
Caçador	35.84		222.74	
S.J.Triunfo	24.00		236.00	
Cambiju		30.10		
Palmeira		41.66		781.25
Guarapuava		19.80		433.33
Mangueirinha		51.25		804.17
Palmas		34.04		465.00
Slaviero		48.59		425.00
Cianorte		29.35		650.00
Lichinoski		86.29		587.50
Galcarneiro		28.62		550.00
Bituruna		22.40		429.17
INF	19.00		219.00	

CHAPTER 3

REMOTE SENSING INSTRUMENTATION AND PROCESSING TECHNIQUES

3.1 Introduction

The use of remote sensing for assessing and monitoring forested areas normally requires a multispectral and multi-temporal approach. Such an approach needs data from different sources or formats and the effects of all factors that can promote distortions in the quantitative analysis of these data need to be removed or, at least, minimised. First of all, a common scale is necessary and the procedure requires two types of corrections, geometric (spatial) and radiometric (spectral) restoration.

This chapter deals with approaches for removing the inherent effects of radiometric distortions from measurements of radiance on TM data and the consequent conversion of the available information to a common basis. Atmosphere, view and illumination geometry and topographic effects are also discussed as scene-related effects of RS spectral characteristics. The study indicates that image-based techniques can be used to achieve the necessary consistency. Methodological procedures have been developed in order to guarantee the necessary corrections in the satellite imagery and in order to set up an acceptable basis for an operation at monitoring system based on forest vegetation changes that occurred in the period from 1987 to 1997.

Data were obtained from satellite imagery, GPS, maps and aerial photographs. All input Landsat TM 5 images and digitised data were registered to the UTM co-ordinate system. The following section discusses the different approaches available and describes the methods used at each step. These include satellite data collection, sampling strategies, and data pre-processing.

3.2 Remote sensing instrumentation

3.2.1 Landsat description

According to Barrett and Curtis (1999) without doubt the Landsat family is the most important of all the Earth resources satellite families until now. Landsat Thematic Mapper (TM) is often used for mapping vegetation and land cover and became more important for change detection as the span

of available imagery became broader (USDA 1995). TM data from Landsat 5 (TM5) have been available since March 1984. Salient features of TM5 are:

Spatial resolution: 30 m

Spectral resolution: seven bands, three in the visible and three in the reflective infrared (IR) ranges; band 6 covers the thermal IR range at 120 m spatial resolution

Radiometric resolution: 256 levels of grey

Temporal resolution: 16 days (circular almost polar orbit)

Image swath width: 185 km

Nominal altitude: 705 km

The rotation of the Earth around the Sun, the inclination of the orbit and the orbital period are such that the satellite crosses the Equator at the same local Sun time on each orbit (Mather 1999). TM5 has an equatorial crossing time of 09:45 (local Sun time) in the descending node. The satellite travels southwards over the illuminated side of the Earth. The Earth rotates through 24.7° during a full satellite orbit. A full orbit around the Earth takes 98.9 minutes, therefore the satellite completes just over 14.5 orbits in a 24-hour period (**Figure 3.1**). The Earth is imaged between 82°N and 82°S latitude over a 16-day period.

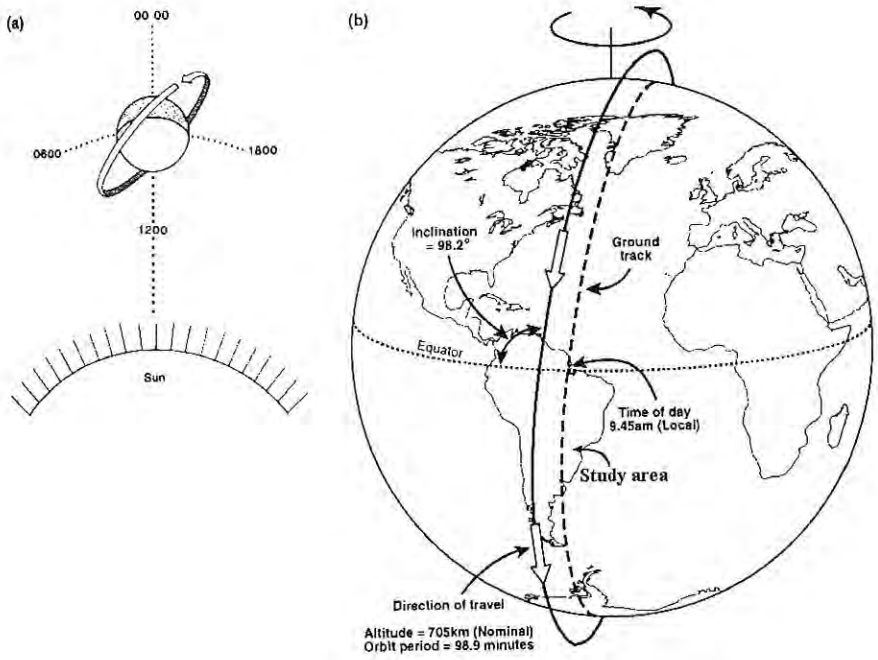


Figure 3.1 Sun-synchronous orbit, the Landsat TM 5 orbit and the region where the study area is located.
Where: (a) illustrating the principles of a Sun-synchronous orbit; and b) example of a Sun-synchronous orbit
Adapted from Mather (1999)

A TM scene is composed of around 6,000 scan lines, with 6,000 pixels per scan line. This means that in each scene there are almost 36 megabytes of pixel values per channel, except for the far fewer 2.25 megabytes in the thermal-IR channel, according to Robinson *et al.* (1995).

TM is a radiometer, with higher spatial resolution (from 57x79 m for all bands to 30x30 m – bands 1-5,7). It presents also better spectral separation (from four to seven bands) and better radiometric and geometric fidelity than the Multi Spectral Scanner (MSS), its predecessor. The improvement in the radiometric resolution, from 64 to 256 levels of grey refined the image processing quality. **Table 3.1** presents the broad applications of TM of data according to waveband.

Table 3.1 Some applications of Landsat Thematic Mapper (TM), 30 m resolution

Band	Wavelength (μm)	Application
1	0.45 – 0.52	(Coastal water mapping, useful for bathymetric mapping of shallow water, soil/vegetation differentiation, deciduous/conifer difference, and cultural feature identification)[1,2,3]; (Chlorophyll absorption)[2];
2	0.52 – 0.60	(Green reflectance by healthy vegetation, vigour assessment, and cultural feature identification; important for discriminating vegetation types)[1,2]; (Reflectance from turbid water)[2]; (Designed to measure green reflectance peak of vegetation as presented in **Figure 4.5**)[3].
3	0.63 – 0.69	((Designed to sense the chlorophyll absorption view, as presented in **Figure 4.5**)[3] and for plant species (culture feature) identification/differentiation)[1,2,3];
4	0.76 – 0.90	(Biomass surveys, water delineation, vegetation types, vigour, soil moisture discrimination)[1,2,3]; (Indicator of cell structure; complete absorption by water facilitates delineation of shorelines)[2]; (Useful for determining vegetation types)[3].
5	1.55 – 1.75	(Vegetation moisture measurement, snow/cloud difference, soil moisture measurement; penetrates through thin clouds)[1,2,3];
6	10.4 – 12.5	(Plant heat stress management, vegetation stress analysis, soil moisture discrimination, and thermal mapping applications)[1,3]; (Relative brightness temperature)[2];
7	2.08 – 2.35	(Hydrothermal mapping, mineral and rock types, vegetation moisture content)[1,3]; (Hydroxyl ion absorption)[2];

Source: {USDA (1995), Barrett and Curtis (1999)}[1]; additional information from {Campbell (1996)}[2]; additional information from {Lillesand and Kiefer (1994) }[3].

The characterisation of the different bands as presented in the literature is generalist and it is difficult to decide which ones should be more indicated in the case of the present work only based in this kind of information. Traditionally, for forest purposes, bands 3, 4 and 5 were the most used, mainly when digital data were not available. The adoption of one band from the visible region, the near-infrared band and one band from the middle-infrared region should minimise the quantity of information involved and should maximise the ability of vegetation differentiation, which was the main objective of the earlier use of RS in forestry. For the purpose of this thesis, all bands were included in the analysis but band 6, which was not available and has a different spatial resolution (120 m).

3.2.2 Orbital images, topographic maps and ancillary data

The satellite imagery used in this study comprised a temporal set (1987 and 1997 TM 5 data from scene 221/78, quadrant A). The orbital images were obtained in digital format. They were stored in CD-ROM by the National Institute for Space Research (INPE / Cachoeira Paulista) and corrected to INPE's level 4 (for details see **section 3.2.3**). The sets were obtained directly from INPE, and

were selected using a data file, which details information about the image's quality. **Table 3.2** presents a summary of the characteristics of the temporal set.

Table 3.2 Temporal set used in the study

Date	Number of bands	Path/row	Visibility level(*)	Clouds cover
20/06/87	6 (1 to 5 and 7)	1859/1204	9	Free
18/08/97	6 (1 to 5 and 7)	1869/1261	7	Free

(*) visibility level: 0 from non-visibility at all; 9 for maximum visibility.

To select the sets of images, various criteria, listed in **Table 3.3** were considered:

Table 3.3 Criteria for selecting the data sets

Criteria	Details
Date of acquisition	Seasonal land cover difference is an important variable. The obtained sets of images were taken in the winter, which frequently is the best season in terms of good visibility level and minimum cloud cover. There was a maximum difference of three months between them, which means more similarity in illumination conditions at the time of the image acquisition.
Visibility level	The image sets presented visibility level at minimum 7, which means in average 70% of visibility in the full scene. However, the visibility was 100% in the study region.
Cloud cover	Data were considered free of cloud cover, mainly in the Quadrant A

Topographic data were the IRATI (SG.22-X-C-I-4) and TEIXEIRA SOARES (SG.22-J-II-3) maps that cover the study area, both in the scale of 1:50,000, UTM-projection (zone 22-S) horizontal-datum: Córrego Alegre. The maps were copied from the original made by the Diretoria do Serviço Geográfico do Ministério da Guerra, in 1973. These maps were used in the geometric restoration process and were the source of information for the digitalisation of features like roads, main streams, urban settlements and railroads. Additional ancillary data were obtained in the front of aerial photographs from Irati National Forest (INF) and adjacent areas (Scale 1: 25000), taken on 05/05/80.

3.2.3 Brazilian system for receiving satellite data

The National Institute for Space Research (INPE) is the institution that receives and distributes satellite data in Brazil, although other private companies also sell images. A Reception Station located in Cuiabá, State of Mato Grosso do Sul, and operating since 1973 is the first component of the Brazilian System. This Station is located in the geographic centre of Latin America, which permits data acquisition for all over the Brazilian territory and part of the neighbouring countries. The other component of the Brazilian System is the data processing laboratory, located in Cachoeira Paulista, State of São Paulo (Novo 1989). Users can choose between hard copies (paper or transparencies) or digital products like CCT (computer-compatible tapes) and CD-ROMs. INPE offers three different levels of corrections: a) level 4 is a systematic correction, which means

geometric correction at rows level to remove variations in the satellite's mirrors movement and to align pixels between adjacent sweeps; b) level 5 consists of resampling in both directions (rows and columns), using the nearest neighbour algorithm and the application of one cartographic projection like Space Oblique Mercator (SOM) or Universal Transverse Mercator (UTM); and c) level 6, which utilises the same procedure as level 5, except that the resampling procedure is made with the cubic convolution algorithm.

3.3 Factors affecting spectral characteristics of remote sensing data

The information contained in Landsat digital data are represented by the digital numbers (DN) for convenience but they do not represent features such as radiance, reflectance or temperature in the terrestrial surface. They may be used directly in some applications that are based on the consistent internal differences of a scene (**Table 3.4, left column**). When other type of responses are expected (**Table 3.4, right column**), however, it is essential to convert these numbers to physical values, since differences in atmospheric conditions, illumination and/or sensor's performance have to be taken into consideration.

Table 3.4 Applications using DN or reflectance values

Digital Number (DN)	Physical Values (Reflectance)
Mono-temporal procedures	Multi-temporal procedures
Identification of features	Multi-sensor analysis
Determination (perimeter) of areas	Spectral classification
Quantification (size) of areas	Band rationing or differencing
Classification	Relationship between structural and spectral variables (biomass, volume)
Mapping	Monitoring

A satellite sensor measurement shows the intrinsic properties of the land surface but also contains the effects of interaction of several contributing factors. According to Teillet (1986) these effects can be categorised in two broad classes, called sensor-related effects and scene-related effects. However, it is a simplification of the phenomenon, because almost all of the effects are somehow inter-related.

3.3.1 Sensor-related effects

These effects are related to the upwelling radiance quantisation levels of each sensor, which are not directly comparable across spectral channels for the same sensor because of differences in the calibration gain and offset of each channel (Mispan 1997).

Some other distortions related to the sensor are: a) striping, due to detectors going out of calibration; b) random noise, due to unpredictable and unsystematic performance of the sensors or transmission of the data; and c) scan line drop out, due to signal loss from specific detectors

(Eastman 1997). Lillesand and Kiefer (1994) mentions a systematic source of distortion, called skew distortion, which occur due to eastward rotation of the earth beneath the satellite during imaging.

3.3.1.1 Sensor calibration coefficients

According to Mather (1999) sensor calibration is important in studies of change detection when comparisons of pixel values from images obtained at different times are performed. The target reflectance result is a combination of sensor calibration and atmospheric plus view angle corrections.

Normally, detectors and data systems are designed to produce a linear response to incident spectral radiance, which output is known from pre-launch calibration. The onboard calibration sources form the basis for constructing the radiometric response function by relating known radiance values on the detectors to the resulting DN's (**Figure 3.2**).

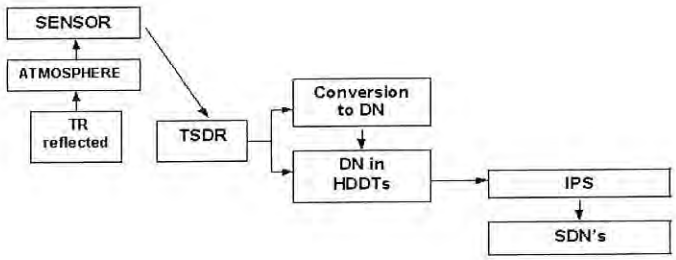


Figure 3.2 Scheme of spatial data collection

where:
Sensor = TM; TR reflected = total radiance reflected by a target; TSDR = terrestrial station for data reception; HDDts = high density digital tapes; IPS = image processing station; and SDN's = station calibrated DN's.

Gain (G, multiplicative) and offset or bias (O, additive) are the calibration coefficients that are set initially through a pre-launch calibration process. They also subsequently can be estimated over the effective lifetime of the instruments by in-flight calibration or by means of measurements made over known targets (Mispan 1997).

For any given TM channel, the relationship between radiance and DN values is a linear fit:

$$DN = G * Ln + O$$

Where:
DN = Original digital number no-calibrated; *G* = slope of response function (channel gain); *Ln* = spectral radiance measured (over the spectral bandwidth of the channel); and *O* = Intercept of response function (channel offset).

According to Lillesand and Kiefer (1994) the TM employs an internal radiometric calibration source consisting of three tungsten filament lamps, a blackbody for thermal band and a pivot mounted shutter. The shutter passes through the field of view of the instrument's detectors (16 of them per band, excluding the thermal infrared channel) each time the scan mirror changes directions. Markham *et al.* (1985) documented Landsat TM coefficients and the authors mentioned that there are differences in sensor response even for the same sensor over time. Mispan (1997) stressed that the accuracy of absolute calibration of remote sensing may be limited by the performance of the sensor over a period of time unless updated calibration coefficients are available. Moran *et al.* (1995) and Chen *et al.* (1996) also considered the sensor's degradation (due to long time operation) and mentioned that an updated calibration coefficients should used in radiometric conversion to transform digital numbers into biophysical values. Mather (1999) mentioned that an alternative to the absolute calibration procedures for the determination of reflectance factors is to use one of the images in a multi-temporal image set as a reference and adjust the other images according to some statistical relationship with the reference image.

3.3.2 Scene-related effects

Scene-related factors influence the spectral response and sometimes are responsible for misinterpretation when temporal data sets are superimposed. As Mispan (1997) mentioned, the existence of a difference between two data sets does not always mean that there has been a change in land cover, because other sources of variation in the data can happen. According to the author, the main factors that perturb the signal are: a) the atmosphere; b) illumination and viewing geometry at the time of data acquisition; and c) the target itself.

3.3.2.1 Atmosphere

According to Campbell (1996), any sensor that observes the Earth's surface using visible or near-visible radiation will record a mixture of two kinds of brightness: a) the brightness due to the reflectance of the Earth's surface; and b) the brightness of the atmosphere. Therefore, atmosphere has a significant effect on the intensity and spectral composition of the energy recorded by remote sensing. All energy reaching the remote sensing instrument must pass through a portion of the Earth's atmosphere and during this process there is an alteration in intensity of Sun's energy and wavelength by particles and gases (Campbell 1996). Consequently, the image is degraded in quality and the accuracy of interpretations may be affected. Several physical processes can modify the solar energy values, including: a) scattering; b) absorption; and c) refraction.

According to CSES (1996), there are approximately thirty gases in the atmosphere, but only seven of them produce observable features in the 0.4-2.5 μm (TM bands 1 to 7). They are water vapour (H_2O), carbon dioxide (CO_2), ozone (O_3), nitrous oxide (N_2O), carbon monoxide (CO), methane (CH_4) and oxygen (O_2). Atmosphere and aerosols scatter solar radiation; the molecular scattering (also called Rayleigh scattering) occurs when atmospheric particles have diameters that are small relative to the wavelength of radiation. It decreases rapidly ($\sim\lambda^{-4}$) with increasing wavelength.

Above 1 μm, the effect is negligible. Aerosols are the main unknown and most variable atmospheric component in the estimation and correction of the atmospheric effect (Mispan 1997; CSES 1996). The aerosol scattering also decreases with increasing wavelength, but not as rapidly as the Rayleigh scattering. Near 1 μm, the aerosol scattering contributes to approximately 10% of the total radiance measured by a downward looking satellite over land areas under typical atmospheric conditions (CSES 1996). Absorption depends on the amount of the absorbing gases in the atmosphere, mainly water vapour (H_2O) and ozone (O_3). Although the effect of gaseous absorption is usually minimised by choosing spectral bands in the atmospheric windows (regions where there are little atmospheric gaseous absorption), some of gaseous absorption still remains (Mispan 1997). Absorption reduces the amount of energy in a given wavelength whereas scattering redistributes the energy by changing its direction. The atmospheric effect is wavelength dependent and varies with time and space. This is particularly important for change detection studies since the existence of a difference between multi-temporal data sets does not always means that there has been a change in surface feature, but the difference may be due to the atmospheric effects (Mispan 1997). Robinove (1982) stated that analysts must recognise that, without an atmospheric correction to the data, the analysis is over the Earth-atmosphere system, rather than the Earth alone. If the correction is not possible, the analyst must assume that the atmosphere is uniform over the area in the entire Landsat image or that it is identical during two Landsat passes. **Figure 3.3** presents the transmittance spectra of some atmospheric constituents:

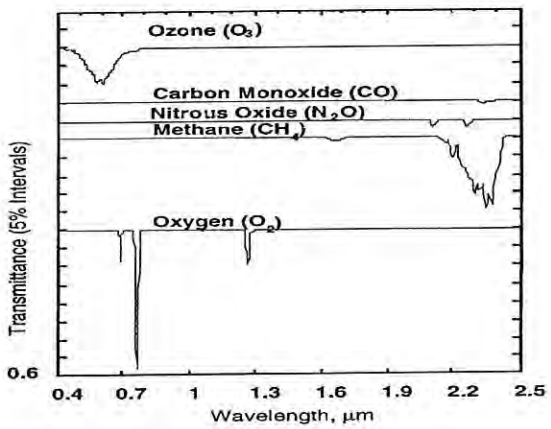


Figure 3.3 Transmittance spectra of atmospheric ozone, carbon monoxide, nitrous oxide methane, and oxygen.
Source: CSES (1996)

Atmosphere affects the radiance in two ways: a) it attenuates the energy illuminating (and being reflected by an object; and b) it acts as a reflector itself, adding a scattered, extraneous path radiance to the signal (**Figure 3.4**) detected by the sensor (Lillesand and Kiefer 1994). Mathematically, the total radiance can be expressed as:

$$L_{tot} = \frac{\rho ET}{\pi} + L_p$$

where:
L_{tot} = total spectral radiance measured by sensor; ρ = reflectance of object; E = irradiance on object, incoming energy; T = transmission of atmosphere; and L_p = path radiance, from the atmosphere and not from the object

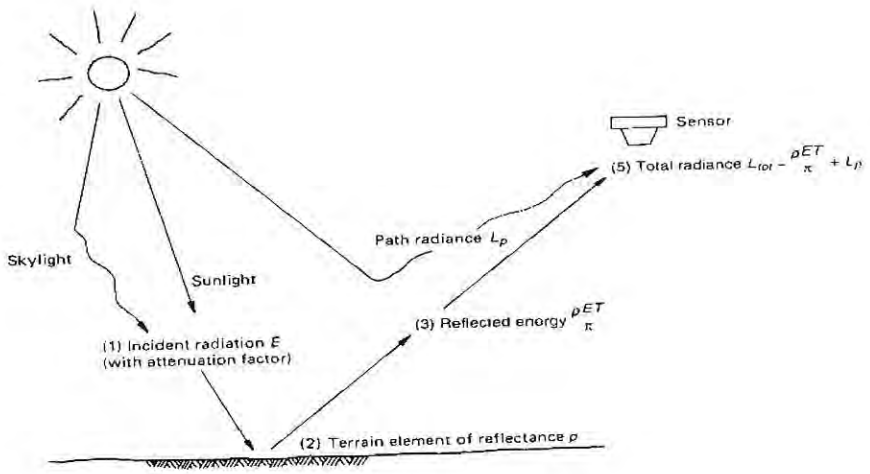


Figure 3.4 Measurements of reflected solar energy influenced by atmospheric effects.

A target having reflectance ρ reflects attenuated sunlight and skylight (E). The attenuated radiance reflected from the target ($\rho ET / \pi$) combines with the path radiance ($L\rho$) to form the total radiance (L_{tot}) recorded by sensor. Source (Lillesand and Kiefer 1994)

Retrieval of reflectance from satellite sensor digital count requires knowledge of the atmospheric conditions and the sensor absolute calibration (Moran *et al.* 1995). According to Mather (1999), atmospheric corrections might be necessary to: a) compute a ratio of the values in two bands of a multispectral image; b) relate upwelling radiance from a surface to some property of that surface; and c) compare results or ground measurements made at one time with results achieved at a later date. Atmospheric particles and molecules cause scattering effects on the transmission of energy, preferentially to shorter wavelengths (particularly the blue wavelength). The effect of haze is usually a relatively uniform elevation in spectral values in the visible bands of energy.

According to Mispan (1997), atmospheric correction algorithms involve the derivation of the optical characteristics of the atmosphere and the application of these characteristics in the correction algorithms. Perhaps atmospheric correction is the subject that has received more attention from researchers. Therefore, many atmospheric radiative transfer codes that may estimate the effects of atmospheric scattering, its characteristics and processes to attenuate them have been suggested in the literature. Those characteristics can be: a) obtained from simultaneous *in situ* atmospheric

measurement; b) obtained from radiometric measurements over calibrated ground areas (Slater 1987); c) obtained from modelling calculations; and d) estimated by using special features of the target and the atmosphere.

Most of these codes were developed on behalf of vegetation indices, especially NDVI (Normalised Difference Vegetation Index). According to Liu *et al.* (1995) NDVI's equation has a simple, open loop structure, which renders it susceptible to large sources of error and uncertainty over variable atmospheric and canopy background conditions. The authors developed a modified NDVI (MDVI) equation by adding: a) a soil and atmospheric noise feedback loop; and b) an atmospheric noise compensation forward loop. Tucker *et al.* (1986) demonstrated the measurable link between atmospheric CO_2 drawdown and terrestrial NDVI dynamics and suggested that there may be quantitative relationships between multi-temporal satellite data and atmospheric CO_2 drawdown. Myneni *et al.* (1994a) developed vegetation/atmosphere radiative transfer method to study atmospheric effects in spectral vegetation indices (noted NDVI, ARVI, SAVI and SARVI). Myneni *et al.* (1994b) proposed a simple linear model to address the influence of pixel heterogeneity, background, atmospheric, and bi-directional effects on the relationship between FAPAR (fraction of incident photosynthetically active radiation) and NDVI.

3.3.2.2 View and illumination geometry

As mentioned before, the spectral reflectance curve of a feature is not constant over time nor is it the same for the whole area covered by that feature. It will be affected by many factors, like the soil status and specificity of the target, but also by the azimuth and elevation angles of the Sun (**Figure 3.5**) and the look angle of the sensor (Mather 1999). The topographic position of the target and the state of atmosphere will influence the reflectance of the feature. Therefore, this topic is related to: a) the solar elevation angle, b) the view angle of the sensor and c) the slope and aspect angles of the target.

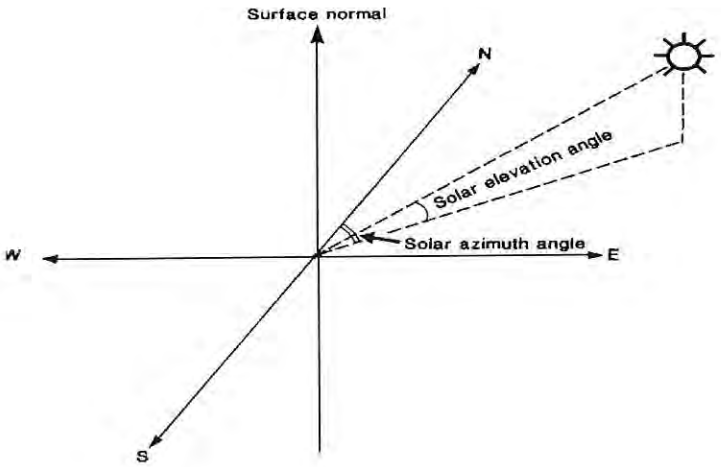


Figure 3.5 Solar elevation and azimuth angles
Source: (Mather 1999)

As stressed by Mather (1999), a satellite sensor operating in the visible and near-infrared spectral regions does not observe and detect all the reflected energy from a ground target over an entire hemisphere, but it records the reflected energy that is returned at a particular angle. The bi-directional reflectance distribution function (BRDF) describes the magnitude of the upwelling radiance in terms of illumination angle and the view angle of the sensor. Ni (1999) presented an analytical model for bi-directional reflectance over a semi-arid landscape and commented that the model had potential for surface parameter retrieval for the study of biosphere-atmosphere interaction. Royer (1985) mentioned that the bi-directional reflectance of an object tends to be higher in the anti-solar view direction and lower in the forward scatter view direction where shadowing is prevalent. Here another key word is presented; shadowing is very important to vegetation studies and according to Mather (1999) will also be dependant upon the solar elevation angle. The author mentioned that for example, shadow effects in row crops would be greater at low Sun elevation angles than at high angles. However, the phenomenon depends on the row orientation.

The presence of atmosphere alters the expected spectral response of any target, since irradiance will be reduced as the atmospheric path length (which is the distance that the incoming energy travels through the atmosphere) increases (Mather 1999). Elevation is measured upwards from the horizontal plane. Azimuth is measured clockwise from north. The zenith angle is measured from the surface angle, and equals 90° minus the elevation angle, in degrees. The path length will increase as the solar elevation angle decreases, and so the degree of atmospheric interference will increase.

With respect to view angle, Landsat satellite has a fixed nadir view angle. It will control the upward path length and a nadir view angle will be influenced less by atmospheric interactions than would an off-nadir view. Mispan (1997) summarised the results of many studies on the effect of view angles mentioning that the variations are significant and asymmetric about the nadir and are also influenced by the solar zenith angle. Mispan (1997) stressed the importance of the topic mentioning that for multi-temporal analysis, the effect of sensor viewing angle on the surface reflectance have to be adjusted before a quantitative comparison can be made. The author also mentioned that avoiding the effect of view angle in multi-temporal analysis will lead to uncertainty as the variation between data sets are not only due to different state of surface features but also from different view angle.

3.3.2.3 Topographic effects

According to Eastman (1997) topographic effect is defined simply as the difference in radiance values from inclined surfaces compared to horizontal ones. The iteration of the angle and azimuth of the solar rays with slopes and aspects produce topographic effects resulting in variable illumination.

Several techniques for mitigating topographic effect have evolved in recent years (Eastman 1997). According to Mather (1999), Lambertian surfaces assume that Lambert's law applies or, the radiance measured at an observation angle θ is the same as that measured at an observation angle of 0° adjusted for the fact that the projection of the unit surface at a view angle of θ is proportional to $\cos\theta$. The author however, mentioned that it is usually the case that the distribution of reflected flux from a surface depends on the geometrical conditions of measurement and illumination, and although this assumption, target reflectance is generally non-Lambertian. Consequently, the component relief affects the reflectance curves obtained by TM imagery and its impact will depend on how high the relief is.

Generally, corrections for topographic effects require a digital elevation model (DEM). The image is registered to the DEM and estimates of slope angle and slope direction can be derived for each pixel position in the image (Mather 1999). Mispan (1997) mentioned four physically-based approaches, which are: a) the statistical-empirical correction method, the cosine correction method, the Minnaert method and the C-correction method. Itte *et al.* (1993) considered that the C-correction method, the statistical and the Minnaert approaches were effective to visually reduce relief effects.

However, it is important to emphasise (Eastman 1997) that many of these techniques tend to be only appropriate for the specific environment in which they were developed (site-specific), or they require high-detail ancillary data that is often unavailable. The simplest method is to use a band ratioing technique between red and near-IR channel (Holben *et al.* 1980). Another accessible technique frequently used is the application of Principal Component Analysis (PCA), and both (band rationing and PCA) are discussed in **Chapter 5**. In the study area, however, the topographic effect was also attenuated because differences in altitude were in the maximum around 200 m.

3.4 Processing methods

Images generated by remote sensors, aerial photographs, or satellite imagery, are subject to many types of spatial distortion. Hence, the objects' position, their areas or the phenomenon's that they represent are not cartographically accurate (Crosta 1992). Therefore, for mapping purposes or for project integration using GIS, it is essential that any form of remotely sensed imagery be accurately registered to the proposed map base (Eastman 1997) or another registered image set. Corrections are normally made to the raw satellite data in primary processing to eliminate the distortions (Howard 1991). This is related to the image processing system, a set of analytical routines designed for the restoration, enhancement and computer-assisted interpretation of digital images. According to USDA (1995), image processing refers to the manipulation of digital image data and its goal is to create an accurate and easily interpretable map of a particular application. There are several different ways in which remote sensing data may be presented, manipulated and inspected and, as Barrett and Curtis (1999) stressed, there are also different stages in the inspection and utilisation of those data. Novo (1989) considered that digital image processing stages can be

classified into three groups: pre-processing techniques (preparation of the images), enhancement techniques and classification techniques.

More specifically, and according to Eastman (1997), digital image processing is largely concerned with four basic operations: a) image restoration; b) image enhancement; c) image transformation; and d) image classification. In this section, the methodology used in the image restoration is presented. The operations were performed using SPRING - System for Processing of Georeferenced Information, developed by INPE and presented in **Appendix 3.1**.

3.4.1 Image restoration

According to Barrett and Curtis (1999), data from remote sensing missions are rarely optimal in quality, and moreover, each system has its specific characteristic capability, instrument design dependency, platform altitude and performance of recording instrument. Consequently, pre-processing functions are necessary to reduce or eliminate such inadequacies. Image restoration (also called image *pre-processing*) seeks to remove degradation effects, which are intrinsically connected to remotely sensed images and especially remote sensing of the environment. According to Howard (1991) and Eastman (1997), image restoration can be broken into the two sub-areas of *geometric restoration* and *radiometric restoration*. Image restoration is essential as part of image analysis, forming its first step when, for example, the geometric accuracy of the scene needs to be further improved for thematic mapping, or when the data are to be used multi-temporally in monitoring changes of the forest cover (Howard 1991). A scheme of the usual stages for geometric restoration and possible approaches for radiometric restorations is presented in **Figure 3.13**.

The first procedures taken up in the image processing were performed at INPE (São José dos Campos-SP, Brazil), using a PC Pentium II platform. The selected image processing software is called SPRING (Georeferenced Image Processing System), Windows version, developed by INPE.

Defining the area boundaries: as mentioned before, the IRATI and TEIXEIRA SOARES topographic maps were considered as the limits of the working area; these maps involve the Irati National Forest and surroundings. The area has approximately 675 km^2, or 67,500 ha.

Co-ordinates: Latitude ..... S 25° 17' 00.00" to S 25° 28' 30.077"

Longitude .....W 50° 16' 29.94" to W 50° 42' 0.014"

3.4.1.1 Geometric restoration

As mentioned before, a raw image is not a map. *Rectification* is the process of projecting the data (raw image) onto a plane and making it conform to a map projection system. According to USDA (1995) rectification (also called geometric restoration) reorients the image to compensate for the Earth's rotation and for variation in the position and attitude of the satellite. This procedure may also includes positioning or warping an image into a map projection system so that accurate measurements can be made (Smith *et al.*1994).

Image registration (also called *georeferencing*) is related to the process of assigning map co-ordinates to the image is; this refers to the location of a layer or coverage in space as defined by a known co-ordinate referencing system. A common form of georeferencing is to indicate the reference system (e.g. latitude/longitude), the reference units (e.g. degrees) and the co-ordinate positions of the left, right, top and bottom edges of the image. Consequently, georeferencing is not only a restoration method but a method of re-orienting the data to satisfy the specific desires and project requirements of the data user (Eastman 1997). By this means it is also possible to position two sequenced images of the same region to be coincident with respect to each other (Thomas 1987).

As mentioned by Barrett and Curtis (1999), rectification can be performed by optical and electronic means or by digital techniques but, basically, the problem involves bringing separate data into congruence, or ensuring that point-to-point correspondences are achieved. The authors also mentioned that the relative ease with which digital satellite data such as from Landsat can be brought into congruence is one of the principal advantages of them. This characteristic is particularly important when change detection studies are performed because they usually involve precise geometrical processing of the image data set before interpretation (Igbokwe 1999). All the set of images involved in the process must be superimposed to each other with acceptable accuracy. Considering the historical aspect of the information based on satellites, sometimes the available image sets are multi-sensoral (for example Landsat MSS and Landsat TM, which are very common as combined temporal sets). Igbokwe (1999) presented a method for geometrical rectification and image registration, which considered fine registration with two-dimensional image correlation. The method was used for the purposes of change detection studies and the procedure resulted in image registration accuracy of 0.28 pixel, considered good for the mentioned purposes.

According to Lillesand and Kiefer (1994) the geometric correction process is traditionally implemented as a two-step procedure. Firstly, systematic or the predicable distortions are considered and they are normally made by the distributors of the imagery who correct most elements associated with image capture, in different correction levels. The second step considers the unpredictable or random distortions. These distortions may be due to several factors, including: a) the perspective of the sensor optics; b) the motion of the scanning system; c) the motion of the platform; d) the platform altitude, attitude, and velocity; e) the terrain relief; and f) the curvature and

rotation of the Earth. Geometric corrections are intended to compensate for these distortions so that the geometric representation of the imagery will be as close as possible to the real world (CCRS 1999).

Campbell (1996) mentioned two approaches to (random) geometric correction. The author called the first one *analytical* approach which needs as inputs: a) the satellite altitude, b) satellite trajectory, c) the shape of the Earth's surface, d) motions of the scanner and the Earth's surface, in relation to the satellite. This process can correct only some of the geometric distortions even if the input information is known and precise. As an example of the analytical approach, the geodetic error determination method shortened to GED can be mentioned. It was proposed by Labovitz *et al.* (1986) and it requires information on orbital and platform altitude at the time of data acquisition. In the process to geometric correction, the parameters are adjusted and the benchmark array recomputed. Again, the practicality of the method is doubtful due to limited access to orbital and altitude information. Recently Huseby *et al.* (1999) developed a method for geometric correction of satellite images where several parallel, rigid lines are acquired at the same time. The methodology includes an algorithm for automatic correction of ground control points (GCPs). The authors concluded that the developed automatic geometric correction algorithm represents an accurate method for correction of NOAA/AVHRR and EOS MODIS data, but has to be improved for application of Landsat TM data.

3.4.1.1.1 Geometric restoration by resampling

Another approach to geometric correction, called *resampling*, does not involve information about the system geometry (Campbell 1996) but treats the images as arrays of values that are manipulated to create a second array with the desired geometry. Hence, resampling is the process of calculating file values for the rectified image by creating new files. The procedure is supposed to deal with the random (or unpredictable) type of distortions, which as mentioned by Barrett and Curtis (1999), are more difficult than corrections for systematic distortions to solve. The process is usually accomplished through reference to carefully selected (GCPs), as detailed below in the **section 3.4.1.1.2**. The location of these points establishes the geometry of the output image and its relationship to the input image (Campbell 1996).

According to Barrett and Curtis (1999), resampling involves: a) the definition of a geometrically uniform output matrix; b) the transformation by computer of the co-ordinates of each output cell into corresponding co-ordinates in the original data set; and c) the transferral of each pixel value from the image data set to its appropriate place in the output matrix. This is called *image-to-map registration*. Once several well-distributed GCP pairs have been identified, the co-ordinate information is processed by the computer. The procedure determines the proper transformation equations to apply to the original (row and column) image co-ordinates to map them into their new ground co-ordinates (CCRS 1999). Resampling is used to determine the digital values to place in the new pixel locations of the corrected output image. The process calculates the new pixel values

from the original digital pixel values in the uncorrected image (CCRS 1999). However, according to USDA (1995), besides the inherent resolution of the acquired data, certain image processing functions can influence the usefulness of the data. This means that the data correction method needs to be carefully chosen and based on the knowledge of the implications of its use. Different interpolators can assign the new grey levels to the registered image in the resampling process (Mather 1999). The majority of the digital image processing packages provides resampling algorithms and the most available processes are: a) nearest neighbour; b) bilinear interpolation; and c) cubic convolution (**Figure 3.6**).

The nearest neighbour interpolator is the simplest one and assigns to the corrected image pixel the same value as the pixel that is nearest to the position to be occupied. Advantages of the method: simplicity and ability to preserve the original values in the unaltered scene. Disadvantages: it may create noticeable positional errors that can be severe in linear features where the realignment of pixels is obvious (Campbell 1996).

The bilinear interpolator ensures that the grey level to be attributed to the pixel of the corrected image is determined by the "weighted" average value of the four nearest input pixels. Campbell (1996) mentioned that "weighted" means that nearer pixels values are given greater influence in calculating output values than are pixels that are more distant. The advantage of the method is that the output image has a more natural look. Its disadvantage is that the original brightness values are lost, or more clearly, the pixels values are not the same as in the input image. Campbell (1996) also mentioned that because the averaging process is conducted over areas, there is a decrease in the spatial resolution.

Cubic convolution uses the same philosophy as bilinear interpolation, but extends a weighted average of values within a neighbourhood of about two pixels in each direction, consequently encompassing 16 pixels. Advantages of the method: the output image is more attractive (Campbell 1996) than are those of other procedures. Disadvantages: performed alterations are naturally much more drastic than those that occur using the other methods.

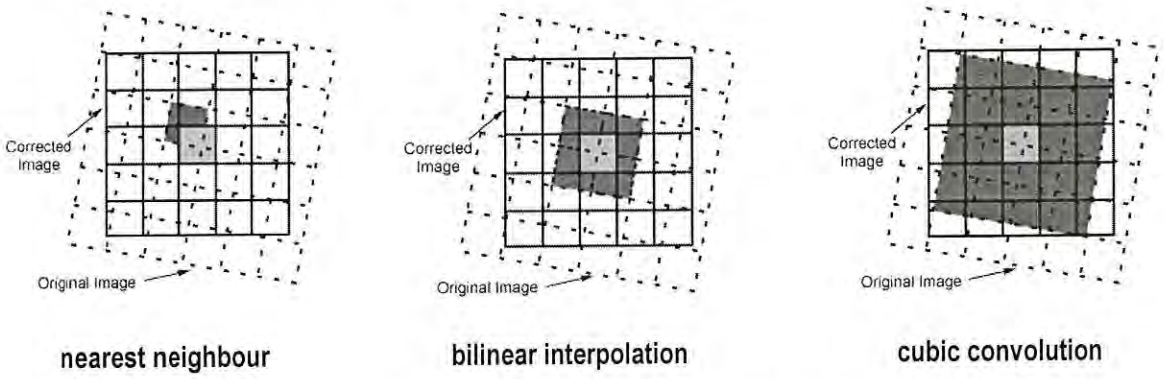


Figure 3.6 Different resampling methods to assign the appropriate DN to an output pixel.
Source: Adapted from CCRS, 1999

NETGIS (1998) mentioned a fourth method denominated hybrid interpolation, which is actually composed by the application of the nearest neighbour or the bilinear interpolator, depending on the local characteristics of the grey levels in the image. Nearest neighbour is applied in the areas of the image where there is not heterogeneity in the values of grey level. When this heterogeneity does occur, the bilinear interpolator is applied.

In this study, the selection of the area window and cutting of the images were performed using the IMPIMA module from SPRING. The entire quadrant A of each image was examined to define the region boundaries. Bands 3 and 5 were used in order to distinguish the different land covers. The module was also utilised to make the conversion from raw format (.i) to grib format (.GRB, used by SPRING). After this analysis, an area with approximately the same boundaries of the topographic maps (IRATI and TEIXEIRA SOARES) was performed using each image set. The initial number of path/row for each set was mentioned in **Table 3.2**.

Specificity of Landsat images: Compensations must be made before accurate analyses are performed using these data (Barrett and Curtis 1999). Both, *systematic* (predictable) *distortions* and *random* (unpredictable) *distortions* occur in the Landsat environment, although they are relatively free from panoramic distortions and relief displacements. On the ground, a Landsat pixel is basically trapezoidal in form, as a result of the general viewing angle of its sensor system. These effects are much more strongly marked with MSS than TM. Corrections for systematic distortions are, in the case of commercial satellite imagery, normally made by the distributors of the imagery who correct most elements associated with image capture, in different correction levels. As mentioned before, INPE provides 3 different levels of corrections and the imagery used in this study was previously correct by level 4. The software SPRING (module IMPIMA) recognises which level was used and if level 5 or 6 was utilised, it does not enable the option " SYSTEM CORRECTION" because it was actually already done in the systematic corrections stage. Barrett and Curtis (1999) mentioned that some systematic corrections could be made relatively easily by the distributor of imagery, for example the skewed-parallelogram effect of Earth rotation beneath the satellite during imaging.

In the present study, the image from 1987 was selected to be the reference image in the process, because it had the best visibility level (**Table 3.2**). The images were resampled by linear interpolation (grey level value to be attributed to the pixel in the new image will be defined according to the values of its four neighbour pixels in the original image). According to NOVO (1989), this procedure has as the advantage of better geometric precision, with the geometrical discontinuities disappearing; the disadvantage is that the implementation requires more calculation and changes the original grey level value.

3.4.1.1.2 Ground control points

As mentioned before, the types of distortions can be broken into systematic or random. Corrections for random distortions are usually more complicate (Barrett and Curtis 1999). However, registration can usually be achieved through the use of a rubber sheet transformation process that warps an image (through the use of polynomial equations) based on the known positions of a set of widely dispersed control points, usually called Ground Control Points (GCPs). GCP is a location within an image for which the map co-ordinates are known. To be useful in the registration process, the co-ordinates of GCPs must be known in the system to which the data are being rectified (Smith *et al.* 1994). Hence, as mentioned by Mather (1999), the method relies on the availability of an accurate map of the area at a suitable scale. When this tool is not available, the only alternative is to locate suitable control points on the ground using Global Positioning System (GPS). Ground Control Points can be recorded from a Viewer (computer display), from a GPS file or from a digitising table. Veronese *et al.* (1998) presented results of a comparative analysis of four methods for geometric corrections of Landsat TM imagery based on polynomial fit. For each method, a different source of reference GCPs was taken. The study was carried out on two areas of state of Bahia, Brazil. The four methods of reference GCPs acquisition were: a) hard-copy topographical maps; b) digital topographical maps; c) hand GPS; and d) differential GPS. The assessment of the correction accuracy was based on the residual errors and the root mean square errors (RMS) of the GCPs and the conclusions were: a) corrections based on both, hand and differential GPS gave the best results; b) the comparison between the two GPS-based corrections showed a slightly better accuracy for the differential GPS than the hand GPS, but not enough to justify the higher cost of the differential GPS co-ordinate acquisition; and c) hard-copy and the digital maps GCPs co-ordinate acquisition yielded similar results.

According to Smith *et al.* (1994) and NETGIS (1998), a minimum number of GCPs is necessary to calculate the transformation, depending on the order of the transformation. This number of points is:

$$N = (t^2 + 3t + 2)/2$$

Where:
N = minimum number of points
t = order of the transformation

To register the image it is necessary to define the number of GCPs and the degree of the polynomials. The degree of the polynomials to be used in the points mapping is proportional to the number of points to be utilised. As the degree of the polynomials rises, the degrees of freedom also increase and consequently the bigger will be the number of points necessary to use the equation. The danger of using higher orders of transformation is that the more complicated the equations become the less regular and predictable the results will be (Smith *et al.* 1994). Even if the value found for N is few (three), usually at least six points are chosen to fit first-order polynomials. Mather

(1999) mentioned that experience suggests that, for Landsat TM, 10-15 control points will give acceptable results for a first-order fit and a small image area (up to 1024^2 pixels).

In this study, the selection of the GCPs was carried out in SPRING (module "REGISTRO"). Twelve well-distributed GCPs were selected from the available topographic maps (described in the **section 3.2.2**). Those included road intersections, man-made forest stands boundaries and transmission line intersections (**Figure 3.7**). A digitising table was connected to a SUN workstation, in a pairwise mode (which allows the user to rectify or register the raster image in the viewer/display).

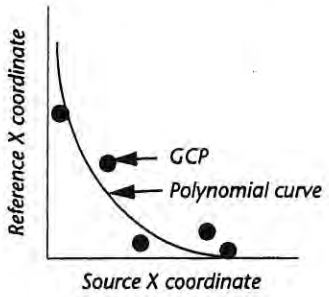


Figure 3.7 Polynomial Curve
Source: Smith *et al.* (1994)

Band 3 was selected and the transformation coefficients were then calculated using a first-order polynomial transformation. The module allows the user to drop and add points, and to evaluate the effects on the a) overall RMS (root-mean-square) error and the b) residuals of points as they are fitted to a new equation. The image sets available in the present study were rectified to a standard UTM-reference frame (Datum Córrego Alegre, Zone 22-South) using a first-order polynomial transformation. UTM was the pre-existing geographic reference system, used by the original maps (IRATI and TEIXEIRA SOARES) to whose the image set was registered. A summary of the rectification process as it was made in this study is presented in **Figure 3.8** and **Table 3.5.**

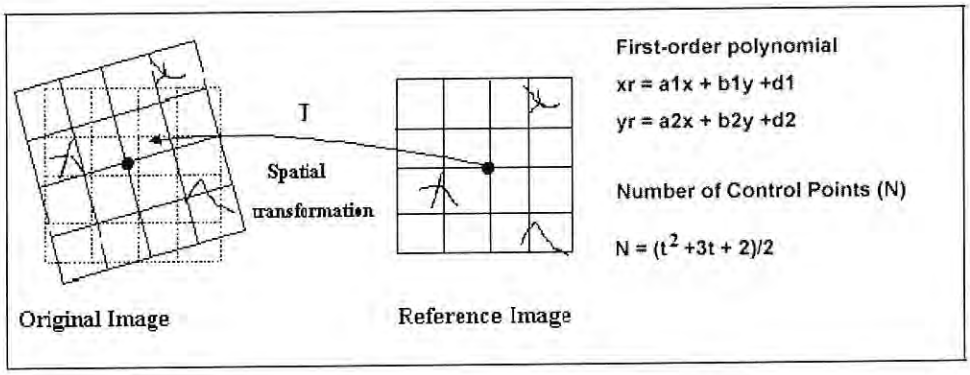


Figure 3.8 The rectification process, adopted polynomial and the number of control points definition.
Adapted from NETGIS (1998)

Table 3.5 Sequence followed by SPRING in the rectification process

Sequence	Box
Acquire the image: Read image from raw format	IMPIMA: Module Import
Load the Database and Project	Database and Project
Calibrate map – digitising table	Calibration
Load image in the Viewer 5 (software peculiarity)	Image selection
Acquire control points	Image registration
Select points to register and analyse RMS errors	Image registration
Save the image to be registered	Image registration
Import the image to the Project	Data import

3.4.1.1.3 Error assessment

According to Silva (1998) errors in the digitising process can be: a) inherent errors, those that were pre-existent (in the originals); and b) operational errors, which area result of the digitising process and of the original data manipulation. The mechanism normally used to control errors is to calculate the root-mean-square error (RMS). RMS is a measure of the variability of measurements. The RMS error is estimated by taking a sample of measurements and comparing them to their true values. The formula of RMS is:

$$RMS_{error} = \sqrt{\left[(x_r - x_i)^2 + (y_r - y_i)^2\right]}$$

Where:

x_i and y_i are the input source co-ordinates

x_r and y_r are the retransformed co-ordinates

Determining the acceptable RMS – According to Smith *et al.* (1994), RMS error is expressed as a distance in the source co-ordinate system. If the data file co-ordinates are the source co-ordinates, then the RMS error is a distance in pixel widths. As mentioned before, the final results will be a balance between the defined order of transformation and an acceptable error (**Figure 3.9**). Sometimes it is considered better to tolerate a certain amount of error than take a higher order of transformation. As a practical rule, NETGIS (1998) advocates that in urban areas, it is acceptable an error of 0.5 pixel, for the resolution of 30 m. Otherwise, due to the complexity of forested areas and the difficulty to find good GCPs arrangement, RMS errors of 3 pixels might be acceptable (NETGIS 1998). For multi-temporal analysis Lunetta (1999) mentioned that any misregistration greater than one pixel will provide an anomalous result for that pixel and suggested, as a precaution, that the RMS between two dates should not exceed 0.5 pixel. Another practical rule is to consider the scale in which it is being worked. For example, if the map is in the 1:50 000 scale, the acceptable error is the half of the scale integer value (divided by 1000), or 25 m. Logically this results in less than one pixel, for resolution equal to 30 m. This study presented a calculated RMS of 0.852 pixels (approximately 25.56 m) and considering the difficulties related to determination of the best GPS arrangement, the approach was considered acceptable.

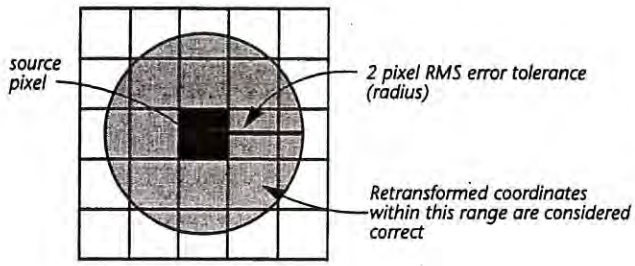


Figure 3.9 RMS Error Tolerance
Source: Smith *et al.* (1994)

3.4.1.2 Geometric co-registration (registration of the image to image)

Geometric registration may also be performed by registering one (or more) images to another image, instead of to geographic co-ordinates. This is called image co-registration or *image-to-image registration* and is often done prior to performing various image transformation procedures (CCRS 1999) and it is useful for registering multi-temporal sets. The schematic matching is showed in **Figure 3.10**.

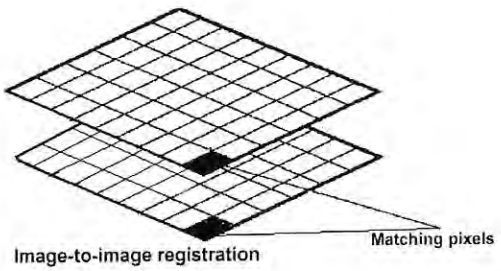


Figure 3.10 Schematic geometric co-registration

According to Novo (1989), terrestrial targets may have their spectral response changed in time, because of inherent (their own nature) or external alterations (illumination, anthropological modifications, and so on). The author also mentioned that the natural targets most liable to inherent modification are the ones related to vegetation. The procedures of detecting such differences can be eventually simple (image differencing, for example). However, as Campbell (1996) noted, practical problems can occur, as: a) the two sets of images must share a common base before they can be registered to one another; and b) minor differences in placement of boundaries can create differences that do not form evidence of land use change but are merely artefacts of the interpretation process.

Therefore, multi-temporal data set analysis requires accurate geometric co-registration. According to Mouat (1993) accuracy assessment procedures are critical in ensuring that the detected change

is real. Townshend *et al.* (1992) exhaustively discussed the impact of misregistration on change detection. Milne (1988) considered that registration errors could potentially be interpreted as a land cover change if two (or more) images are not registered within an accuracy of one pixel or less. Igbokwe (1999) summarised the subject asserting that the accuracy of change detection depends on the accuracy with which the images were registered to each other.

In the process of image to image co-registration, a GCP is selected in the reference image and precisely matched to its counterpart in the other image using spectral characteristics of the data and the transformation matrix. Highly accurate results can be obtained using a fine tune procedure. The mentioned transformation matrix is a set of coefficients used to convert the co-ordinates to the new projection. The process is based on the reflectance values, so histogram matching is eventually a necessary tool to guarantee that there is no offset between those images that should have similar spectral wavelengths (such as two visible bands).

In this study, band 3 from the 1987 image was selected to be the reference image and the 1997 image was co-registered to that first one, using a first-degree polynomial. The resultant RMS was smaller than one pixel (0.486 or 14.58 m approximately). The same criteria used in the **section 3.4.1.1.3** were considered to decide about the suitableness of the results.

3.4.1.3 Radiometric restoration

Radiometric restoration refers to the removal or reduction of distortions in the degree of electromagnetic energy registered by each detector. According to Eastman (1997) and as mentioned before, a variety of agents can cause distortion in the values recorded for image cells. It accounts for variations in the image resulting from sensor anomalies or environmental conditions, such as haze (USDA 1995). Campbell (1996) mentioned radiometric restoration as being the effort of removing undesirable influence of a) atmospheric interference; b) system noise; and c) sensor motion.

According to Mouat *et al.* (1993), remote-sensing detection is a process for determining and evaluating differences in a variety of surface phenomena over time. Guyot *et al.* (1994) stated that statistical classification, based on consistent differences within a scene, does not require radiometric corrections, but it is often necessary to normalise radiance values to a standard set of conditions or to retrieve the target reflectance factors, when multi-temporal information is used. As Lillesand and Kiefer (1994) mentioned, these standard sets would involve a sun elevation correction and an earth-sun distance correction, considering that these values are different from date to date. The normalisation of all included data to standard conditions is also the reason mentioned by Hill *et al.* (1991) to perform radiometric restoration, when the study is multi-temporal. According to Sousa (1997), the process is necessary because of calibration differences motivated by different sensor utilisation or wearing of the detectors over time.

Therefore, besides the process of the different sets of images co-registration, it is necessary to calibrate radiance values when the goal is analysing patterns of reflectance at different times. The objective is that the absolute reflectance levels of objects measured by different sensors correspond across separate images (Eastman 1997), since apparent temporal differences result from many sources may mask changes in surface material (Mouat *et al.* 1993).

The type of radiometric correction obviously varies according to type of the sensor but, summarising, two steps are normally followed to deal with the radiometric restoration. The first one, which can be called *radiometric correction,* is the conversion of digital numbers (DNs) to absolute radiance and reflectance values, including in this step the sun elevation correction and an earth-sun distance correction. The second approach, called *radiometric calibration* is related to the problems derived from atmosphere and illumination differences when multi-temporal data sets are being used. Both procedures are described in the next sections.

3.4.1.3.1 Radiometric correction

The term radiometric correction was adopted here to name the procedures related to the conversion of DNs to physical values, which will reduce errors in the analysis of those data (Robinove 1982). The author mentioned that Landsat digital images are commonly analysed by using the digital numbers DNs for each pixel. This procedure may be satisfactory when only a single image is used for analysis as in mosaic or temporal overlays. However, it may produce incorrect results if more than one image is used. According to the author, the digital numbers for each pixel should be converted to their dimensioned equivalents (number with physical meaning) such as radiance as measured at the satellite, in milliwatts per square centimetre per steradian (mW $cm^{-2}sr^{-1}\mu m^{-1}$) or reflectance of the ground (also called albedo of the ground). Therefore, reflectance and radiance are crucial elements at this stage. Reflectance, according to USDA (1995) is the total solar energy incident on a given feature minus the energy that is either absorbed or transmitted by the feature. Reflectance is dependent upon the material type and condition, and permits us to distinguish different features in a visual image. Radiance is the reflectance of an object modified by the atmosphere between the sensor and the ground, also known as brightness (USDA 1995). Rather than the true reflectance of ground objects, radiance is what the remote sensor measures.

3.4.1.3.1.1 Converting DN to radiance

As discussed before, to study patterns of reflectance at different times or across mosaics of imagery it is necessary to calibrate radiance values so that the absolute reflectance levels of objects measured by different sensors correspond across separated images. Robinove (1982) enumerated the particular analytical conditions for which conversion to reflectance should be done before comparisons as being: a) band ratioing; b) analysis of a digital mosaic of two or more scenes in which sun angles are different; c) analysis of a mosaic of two or more scenes taken by two or more Landsat image or before and after a change in scanner calibration of a single Landsat

image; and e) comparison or detection of changes between two or more images taken under conditions of different sun angles and different calibrations. The author suggested that the conversion require the computational algorithms to be available on the computer. Another important fact to be considered is that related to reflectance differences between solar elevation angles in multi-temporal images. Eastman (1997) mentioned that it might be necessary to normalise solar angles and earth-sun distances in the process. According to Mouat *et al.* (1993) the level of accuracy is influenced by several factors, including scene dependent solar zenith angle, the amplitude of the surface reflectance and the additive atmospheric scattering component. The transformation of radiance values to reflectance considers the differences between angles and corrects the images for this factor, by normalisation.

According to Eastman (1997) detectors are usually calibrated so that there is a linear relationship between DN and spectral radiance. Each spectral band of the TM has its own response function, and its characteristics are monitored using onboard calibration lamps (Lillesand and Kiefer 1994). The digital products produced by DGI/INPE at level 4 only are processed with a systematic correction, which means geometric correction at rows level to remove variations in the satellite's mirrors movement and to align pixels between adjacent sweeps. According to Lillesand and Kiefer (1994) the absolute spectral radiance output of calibration sources is known from pre-launch calibration and is assumed to be stable over the life of the sensor. Hence, the basis for constructing the radiometric response function is the *onboard calibration source*, which relates known radiance values incident on the detectors to the resulting DNs. One question relates to the reliability of the obtained radiance and/or reflectance, since the sensor the Landsat TM 5 has been in continuous operation for a period beyond its foreseen life cycle (Chen *et al.* 1996). The authors also mentioned their literature review of several radiometric studies with field measurements and confirmed the degradation of TM visible bands, especially band 1 and suggested that due to the sensors' degradation, the updated coefficients should be used. There are no orbit calibration studies using Brazilian field measures. Thus, calibration coefficients published in literature might be the best approach.

The linear function (linear relationship between DN and spectral radiance) is described by the following parameters: a) the range of DN values in the image; and b) the lowest (L_{min}) and highest (L_{max}) radiances (both dated constants) measured by a detector over the spectral bandwidth of the channel. Most commonly, the data are distributed in 8-bit format corresponding to 256 DN levels. L_{min} is the spectral radiance corresponding to the minimum DN values (usually, a value of 0). L_{max} is the radiance corresponding to the maximum DN (usually, the value of 255). Not only each sensor, but also each band within the same sensor, has its own L_{min} and L_{max}.

The equation relating DN to radiance in remotely sensed data is deeply discussed in Lillesand and Kiefer (1994). In this study, the procedures and coefficients for post-calibration that were proposed by Markham and Barker (1986) and Markham and Barker (1987) were utilised, with the follow equation, for each band:

$$L_i = \left(\frac{DNi}{255}\right) * \left(L_{maxi} - L_{min_i}\right) + L_{min_i}$$

Where:
L_i = Radiance
i = band
DN_i = digital value of the pixel, in the band i
L_{min} = spectral radiance corresponding to the minimum DN values (usually, a value of 0)
L_{max} = radiance corresponding to the maximum DN (usually, the value of 255)

In the present study, this stage was performed using the software SPRING. To implement the procedure a computer program using LEGAL language (the macro language incorporated to the software) was especially prepared (by the author of this work) in order to convert digital values to radiance values. The values of L_{max} and L_{min} from calibration of Landsat TM 5 calculated and suggested by Markham and Barker (1987) are presented in **Table 3.6**.

Table 3.6 Values of L_{max} and L_{min} used in the conversion to radiance

TM Band	L_{max} (mW $cm^{-2}sr^{-1}\mu m^{-1}$) *	L_{min} (mW $cm^{-2}sr^{-1}\mu m^{-1}$) *
1	15.210	-0.150
2	29.680	-0.280
3	20.430	-0.120
4	20.620	-0.150
5	2.719	-0.037
7	1.438	-0.015

Source: Markham and Barker (1987)
Post-calibration dynamic ranges of U.S. processed TM data; spectral radiances L_{max} and L_{min} (**mW $cm^{-2}sr^{-1}\mu m^{-1}$**). The procedure was carried out in TIPS (TM Image Processing System at Godard) Long Term Parameter File Number TL5027.PAR for in-band radiance from full-width half maximum band-width for Landsat-5.

3.4.1.3.1.2 Converting radiance to reflectance

The next step was to convert the radiance values obtained by the equation above mentioned in reflectance values. According to Eastman (1997), to study patterns of reflectance at different times is often necessary to apply both a sun elevation (and solar angle) correction and an earth-to-sun distance correction. Images that are acquired under different solar illumination angles can be adjusted to one another by dividing each pixel by the cosine of the Sun's elevation from the zenith (90°). The Earth-Sun distance correction is applied to correct for seasonal changes in the distance between the sun and the earth. The combined corrections are related in the following equation, which uses solar irradiance values (E), for each band of Landsat TM 5:

$$R_i = \left[\frac{Li * \pi}{(E_i * \cos\ zen)} \right] * 255^{\#}$$

Where:
R = Reflectance
L = Radiance
i = band
zen = solar zenith angle on the time where the image was taken, in radians. This information was provided by INPE (**Table 3.7**).
E_i = Solar irradiance values
(#) To present the same comprehensiveness of the original digital values, reflectance values were multiplied by 255

Table 3.7 Solar Elevation Angle (in degrees) at the time where the images were taken

Date	Solar Elevation Angle (°)
20/06/87	27°
18/08/97	35°

Source: DGI (INPE)

The values of Solar Irradiance (E) - TM Solar Exoatmospheric Spectral Irradiances (mW cm^{-2} μm^{-1}) used in the conversion to reflectance are presented in **Table 3.8**. The values were suggested by Markham *et al.* (1985); Markham *et al.* (1986). The computer program REFLECTANCE was specially prepared using LEGAL language to perform this step.

Table 3.8 TM Solar exoatmospheric spectral irradiances at mean earth-sun distance used in the conversion to reflectance

Band TM	Irradiance (mW/cm^2 sr μm)
1	195.700
2	182.900
3	155.700
4	104.700
5	21.930
7	7.452

Source: Modified from Markham *et al.* (1986)

3.4.1.3.2 Calibration

The first step in the pre-process for preparing temporal sets of images, the conversion from digital numbers to radiance (which has physical meaning), was already performed. However, a second stage still had to be achieved, to ensure the comparability of the image sets. This second step deals with the interference of atmospheric and illumination conditions. As mentioned before, many atmospheric radiative transfers' codes that may estimate the effects of atmospheric scattering, (like its characteristics and the way to attenuate them) have being suggested in the literature. According to Mispan (1997), those characteristics can be: a) obtained from simultaneous *in situ* atmospheric

measurement; b) obtained from radiometric measurements over calibrated ground areas (Slater 1987); c) obtained from modelling calculations; and d) estimated by using special features of the target and the atmosphere. The two main approaches for making satellite data from different acquisitions comparable are *absolute calibration*, based on physical principles and *relative calibration*, related to statistical principles (Olsson 1994). The first two characteristics mentioned above belongs to the absolute calibration approach and the last two belongs to the process of relative calibration.

3.4.1.3.2.1 Absolute calibration

Absolute calibration aims to convert pixel values to reflectance at canopy level. The process can be broken in: a) *empirical approach* - radiometer measurements are made simultaneously with the satellite data acquisition as the technique mentioned by Vogelman *et al.* (1988). In Brazil, Ponzoni *et al.* (1998), working with data from *Eucalyptus* sp. stands, calibrated (by linear regression) TM transformed reflectance radiometric data using an airbone (SPECTRON SE 590 radiometer, installed in a helicopter). This kind of approach can be complicated and expensive and consequently, several methods have been developed; and b) *analytical or physical approach* - to model the influence of the atmosphere using a radiative transfer code, such as LOWTRAN and MODTRAN or 5S (**Figure 3.11**), described by Tanré *et al.* (1990).

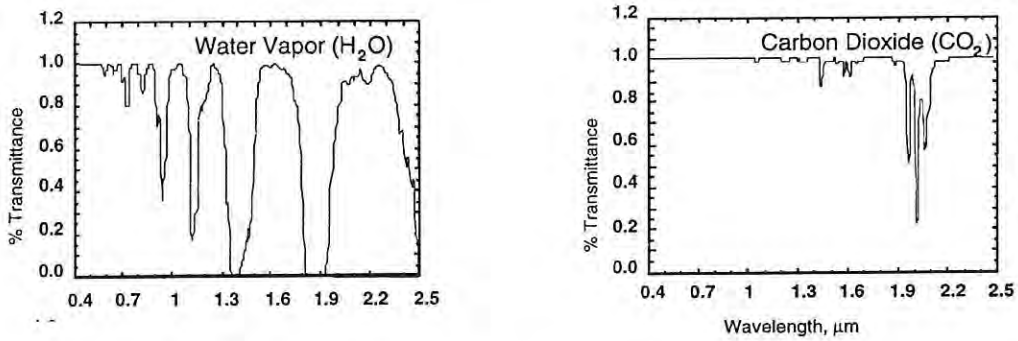


Figure 3.11 Atmosphere water vapour transmittance spectrum (left plot) and carbon dioxide transmitance spectrum (right plot) for specific conditions. The spectra was calculated using LOWTRAN7
Source: CSES (1996)

LOWTRAN7 (**Figure 3.11**) and MODTRAN were developed by U.S.Air Force and both calculate the atmosphere transmittance and atmosphere background radiance for a variety of atmospheric conditions (Campbell 1996). These programs estimate atmospheric absorption and emission by atmospheric gasses (H_2O, NO_2, O, O_3, CH_4, CO_2, N_2O, CO, O_2, NH_3, and SO_2).

The Simulation of Satellite Signal in the Solar Spectrum (5S) code and its successor 6S are semi-analytical atmospheric radiative transfer code used to model the signal received by a satellite sensor system due to solar radiation backscattered by the earth's surface features and atmospheric system. The method was analysed by Olsson (1994) in four different approaches and adopted by Mispan (1997). Yet classified as physical approach and aiming to avoid the need for

simultaneous measurements of the atmosphere, other methods which also use dark areas in the process where developed. One example is the method mentioned by Kaufman *et al.* (1988), whom developed an algorithm which requires some pixels in the image which correspond to dense dark vegetation as the surface cover. Once the presence of such pixels is detected, the algorithm automatically chooses those pixels and corrects the image.

3.4.1.3.2.2 Relative calibration

Relative calibration are based on statistical principles and are estimated from a reference populations of pixels (Olsson 1994) sampled from two images. Some studies consider the mentioned reference population as a sample of the whole area and other techniques use a few manually selected test areas. The most important principle in the majority of the procedures is that those test areas are bright and dark targets with an assumed stable reflectance.

Some methods are based on examination of data distribution and others on pairwise observations of pixel values (Campbell 1996; Olsson 1994) over the whole area. An example of the first ones is the histogram minimum method (HMM), developed by Chavez (1975). It looks for values in areas of known zero reflectance, like deep water. Any value above zero in these areas is likely to represent an overall increase in values across the image and can be subtracted easily from all values in the individual bands (Eastman 1997).

The second group can be represented by the regression method, mentioned by Mather (1999) and by its extension, CMM (covariance matrix method), that was described by Switzer *et al.* (1981). The method was adopted by Hadjimitsis *et al.* (1999a); the authors applied the method to a time series of Landsat TM images of large reservoirs. The methods are robust and often produce satisfactory results, as mentioned by Campbell (1996), who mentioned a modification in the strategy by using an automatic grouping procedure to ensure reflectances in fact uniform. Another effective method for making data from different sensors comparable involves the application of Principal Component Regression.

Campbell (1996) introduced an interesting question: how can the analyst decide whether atmospheric corrections are necessary? The author commented that this might be a difficult decision, as effects of atmospheric degradation are not always immediately obvious from casual inspection. He advised that good tools should be: a) summary statistics (each scene); b) means and variances; c) frequency histograms for suggestions of poor image quality; and d) the absence of dark values, specially if the image is known to show larger water bodies.

As mentioned before, some of the correction methods are based on pseudo-invariant targets. These procedures usually are based on the idea that atmospheric variations over multi-temporal scenes can be estimated in a relative way by taking one image as 'reference'. The approaches are robust but a question remains about how free from atmospheric effects are the reference image

and the selected pseudo-invariant targets. A possible solution to overcome these shortcomings was mentioned by Hadjimitsis *et al.* (1999b), where the authors recommend an approach based on the use of *in situ* measurements. Again, the practicability of the suggestion must be addressed because the collection of ground data necessary to investigate the spectral response of the targets sometimes is not possible.

Hall *et al.* (1991) also proposed a method of radiometric rectification. The method uses linear transformations between sets (dark and bright sets) in which are expected no change or little change over time. The technique, which compensates for sensor calibration, atmospheric and illumination differences between images, was used in the present study and is described in the next session. It assumed stable parts of the bi-temporal scatterogram of pixels values.

3.4.1.3.2.2.1 Radiometric rectification

According to Hall *et al.* (1991), a common radiometric response is required for quantitative analysis of multiple satellite images of a scene acquired on different dates with different sensors. Their technique involves the radiometric rectification of Landsat images of a scene to a reference image. The result is that all rectified images should appear as if they were acquired with the same sensor, while observing through the atmospheric and illumination conditions of the reference image. In their investigation, the authors found that radiometric rectification should be a useful alternative to atmospheric radiative transfer codes and sensor calibration approaches when reliable atmospheric optical depth data or calibration coefficients are not available. Perhaps the major difference between radiometric rectification and other traditional atmospheric corrections is that the first one does not require sensor calibration or atmospheric turgidity data, and thus corrects images from a common scene in a relative, rather than an absolute sense (Hall *et al.* 1991). The temporal set of images is rectified relatively to a selected reference image, determined by visual quality.

The basic principle of the technique is to identify features, which the authors called *radiometric control sets*, that are sets of scene landscape elements with a mean reflectance. After this first step, the average digital count values of these control sets are used to calculate linear transforms relating digital count values between images. According to Hall *et al.* (1991), in theory, some of these control sets may have nearly constant reflectance over time and other elements of the set may change reflectance between images (different dates). This can be due to precipitation, weathering (rock outcrops), growth of vegetation (bare soil) or, in the case of water bodies, be a consequence of sun glint, changing sediment or yet chlorophyll concentrations. The authors concluded that, consequently, the members of the radiometric control sets might not be the same pixels from image to image. The process contrasts with the process of geometric control points for spatial image rectification, which are composed of identical elements in each scene.

The components of the system are: a) the first component forms radiometric control sets that have little or no variation on their mean surface reflectance between images. As mentioned earlier, it is

not necessary that the members of these sets are the same pixel in both images, rather they reside at the non-vegetated extremes of the Kauth-Thomas (KT) greenness-brightness scattergram (Kauth and Thomas 1976, later modified by Crist and Cicone (1984)) and surely did have their spectral response not much altered in time; b) the second component radiometrically rectifies the images using a linear transformation with coefficients calculated to equate the individual band means (in raw digital counts) of the radiometric control sets in each image.

The key assumption for the first component of the system is that the extremes of the KT greenness-brightness histogram always contains at least some pixels that have the same average surface reflectance between images. The first control set is called *dark radiometric control set* and is composed by a group of pixels corresponding to water, low in KT brightness. Water pixels contaminated by sun glint or sediments must be excluded from the set. Within this water set, a KT greenness threshold is used to further eliminate lake-edge pixels and water containing floating vegetation or high sediment concentrations. The reverse order is applied to define the *bright radiometric control set*. Vegetated pixels are excluded and the set is normally composed by pixels representing bare soil. The calculation of the rectification transforms is made by computing the coefficients of linear transformations and the band-by-band average digital count values of the radiometric control sets are used to compute them.

In the present study, the image from 1987 was chosen to be the *reference image* because it was considered the scene that had the lowest contribution from atmospheric effects (clouds free and high contrasts) and the image from 1997 represented the untransformed subject image. To define the dark and bright radiometric control sets, initially two images were generated by transformation of each set of data. They were called respectively Greenness (G) and Brightness (B) images, after Crist and Cicone (1984) and following the reasoning developed by Kauth and Thomas (1976). The images were created for each date and were obtained by transformation of the original image using different weights (coefficients) by for each band. The model was implemented by running a LEGAL program (GREEN) especially written (by the author of this work) for this purpose. The next step was to choose the radiometric control sets, or non-vegetated areas, with little or no variation in their mean surface between images and, consequently, little or no difference in their values of reflectance. The procedure was operated by writing and running a LEGAL program (called GR_BR and wrote by the author of this work) to find the scene elements that resided at the unvegetated extremes of the two-dimensional Kauth-Thomas (KT) greenness-brightness scattergram. According to Hall *et al.* (1991) dark radiometric control set can be defined by selecting the pixels from the lower half of the water mode, in the KT brightness histogram. The upper half of the water mode can contain water pixels contaminated by sun glint or sediments. In this case, the histograms of each G and B image were examined and their extreme values were checked. The control sets were determined by the slicing procedure, performed on G and B images. There was one set of dark and one set of bright control set to G image, as well as to B image, for each date, since in the nonreference image, the reference-image threshold values do not apply, due to atmospheric and

sensor differences. Consequently, threshold values for nonreference image have to be recalculated (Hall *et al.* 1991). The lower values in the interval of reflectance were chosen as dark control set and the bright values were composed by the higher values. As mentioned before, water bodies without turbidity constitute usually the dark control set and the bright control set usually is bare soil.

The generated mask of the G image dark control set was crossed with the mask obtained with the B image dark control set, for each date. This procedure generated another mask, referring to the dark control set for each date. The same procedure was adopted with the bright control set, generating a mask for this set, for each date. Finally, these masks were overlapped to the reflectance images to extract their average. These values should be used as a reference to the rectification procedure. To operate this step in SPRING, other program in LEGAL language (CROSS) was prepared by the author of this work, as part of the training scheme. The sequence of operations is summarised in the **Figure 3.12**.

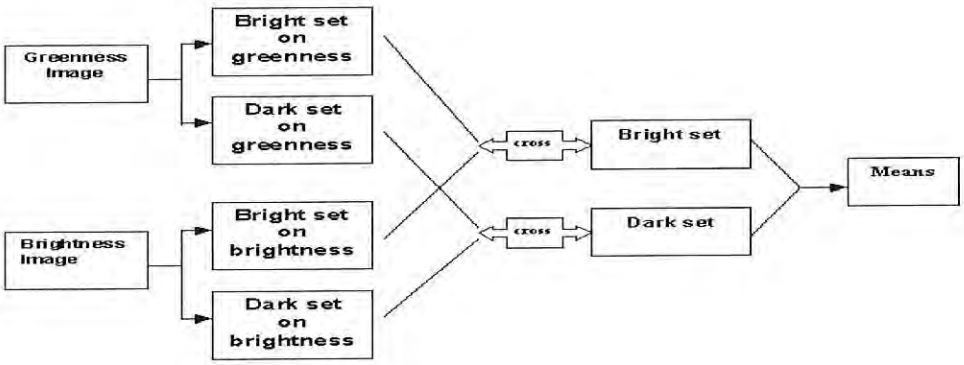


Figure 3.12 Scheme for determination of dark and bright radiometric sets for each date

As the second component mentioned before, the band-by-band average digital count values of the radiometric control sets were used to compute the coefficients of linear transformations relating all digital count values band by band between the two images being rectified, as is shown in **Figure 3.13**.

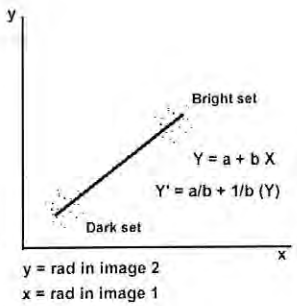


Figure 3.13 Calculation of radiometric transform – rectification transform approach

The coefficients (a and b, **Figure 3.13**) of the radiometric rectification transforms are the gain and offset of the line connecting the means of the dark and bright radiometric control sets in each band.

According to Hall *et al.* (1991) this empirical transformations accounting separately for sensor, illumination, and atmospheric effects, and consequently is linear itself.

As input data, the model requires the values obtained as means of the dark (*D*) and bright (*B*) radiometric control set in the *ith* band of the reference (*r*) image and the untransformed subject image. The results of the whole process are presented in **Table 3.9**.

Rectification transforms are a set of linear transforms, performed for a set of n bands, and expressed by

$$\{T_i(x_i) = m_i x_i + b_i\}_{i-1,n}$$

where:
n = number of bands
m_i = gain valued applied to the digital value in the *i*th band
b_i = offset value applied to the digital value in the *i*th band
x_i = corrected values

For the *i*th band, x_i are given by:

$$x_i = T_i(x_i)$$

The coefficients m_i and b_i are defined by equating the transformed means of the radiometric control sets in the subject image to those of the reference image, i.e.,

$$m_i = (Br_i - Dr_i)/(Bs_i - Ds_i)$$
$$b_i = (Dr_i * Bs_i - Ds_i * Br_i)/(Bs_i * Ds_i)$$
i = band

where:
Dr_i = means of the dark (D) radiometric control set in the *i*th band of the reference (r) image
Br_i = means of the bright (B) radiometric control set in the *i*th band of the reference (r) image
Ds_i = means of the dark (D) radiometric control set in the *i*th band of the untransformed subject (s) image
Bs_i = means of the bright (B) radiometric control set in the *i*th band of the untransformed subject (s) image

It was assumed that the bright and dark radiometric control set radiances were independent both within and between images. It was also considered that the value of the image element being transformed was independent of those quantities (Hall *et al.* 1991). The authors referred that the sampling error is reduced by choosing radiometric control sets at the extremes of the KT histogram, thus separating the mean values as much as possible.

Table 3.9 Reflectance means* of bright and dark control sets before rectification, and the difference of the means of the rectified image for the reference image (1987) and subject image (1997)

BRIGHT RADIOMETRIC CONTROL SETS					
BAND	(r) 1987 Reference image (1)	(s) 1997 Untransformed subject image (2)	1997 radiometricaly rectified image (3)	Differences (1-2)	(1-3)
1	23.1	20.4	21.9	-2.7	-1.2
2	22.0	20.6	21.0	-0.4	-1.0
3	26.1	22.5	24.6	-3.6	-1.5
4	51.0	54.7	49.9	3.7	-1.1
5	79.3	73.0	77.8	-6.3	-1.5
7	40.5	39.4	39.2	-1.1	-1.3
DARK RADIOMETRIC CONTROL SETS					
BAND	(r) 1987 Reference image (1)	(s) 1997 Untransformed subject image (2)	1997 radiometricaly rectified image (3)	Differences (1-2)	(1-3)
1	20.0	16.9	19.6	-3.1	-0.4
2	15.3	13.7	14.4	-1.6	-0.9
3	11.8	9.3	11.1	-2.5	-0.7
4	2.8	4.2	2.1	1.4	-0.7
5	0	0.1	0	0.1	0
7	0	0	0	0	0

Values of reflectance means* were multiplied by 255.

Evaluation of the quality of the rectified images was achieved by: a) dark and bright control sets reflectance means comparison from the date set before and after rectification, with the averages of the reference image; and b) visual analysis of the rectified images.

The first step of the analysis consisted of comparisons between the reflectance means of control sets in the rectified images and the reference image, band by band. Analysing **Table 3.9**, it should be mentioned that before rectification (column 1-2) substantial differences should be noticed in the means for bands 1, 3, 4 and 7 for the bright control set; there was also substantial differences in the means for bands 1, 3 and 4 for the dark control set. Bands 3, 4 and 5, which are very important for vegetation studies presented actually the most noticeable differences, even if the selected samples were supposed to be pseudo-invariants. The phenomenon was more characterised for the bright control set than it was for the dark control set. The possible explanation is related to the sampling procedure: in the second set (dark control set) the location of sample units was facilitated by the presence of deep-water bodies in the study area. The bright control set, on the contrary, was more difficult to be selected because it was searched looking at bare soil, what might bring more heterogeneous sample units. The column 1-3 (**Table 3.9)** presents the means of the control sets (bright and dark) after the rectification. According to Sousa (1997), from an operational point of view, the rectified image might have satisfactory results after rectification process if the differences (mean of reflectance before and after rectification, in relation to a reference image) are around one (1) in relation to the reference image, what was achieved. The visual analysis between the rectified image and the reference image band by band and by colour composition confirmed the overall similarity between the 1987 and rectified 1997 image.

3.5 Chapter summary

This chapter discussed basic correction procedures that any analysis or classification multi-temporal data attempts. The whole process of image restoration is summarised in **Figure 3.14**. The process of georeferencing and co-registration of all involved data sets provide key for trustworthy results, as stressed by Igbokwe (1999). In his summary he stressed the accuracy of change detection depends on the accuracy with which the images were registered against each other. Data for the study consisted of a temporal TM 5 data set; dates: 20/06/87 and 18/08/97; number of bands involved: 6 bands – 1 to 5, and 7 at a minimal visibility of 70%. Data were obtained in digital format and corrected by INPE's level 4 corrections, which means geometric correction at rows level to remove systematic distortions.

For the image restoration process, the image processing software used was SPRING (Georeferenced Image Processing System). Data were converted from raw format (.i) to grib format (.grb) used by Spring utilising module IMPIMA. Images were rectified and georeferenced to the UTM-projection system, zone 22-s and horizontal datum: Córrego Alegre. The image from 1987 was defined as the reference image and twelve well-distributed ground control points (GCPs) were selected from the available topographic maps (using the digitising table). Transformation coefficients were calculated using a first-order polynomial transformation. The resultant images were then resampled using a bilinear interpolator. The study presented a calculated RMS of 0.852 pixels (approximately 25.56 m), which was considered satisfactory for the purposes of vegetation analysis.

Geometric co-registration was performed using band 3 from each date. Image 1997 was co-registered with the 1987 reference image. The calculated RMS, for this procedure was 0.486 pixels, or approximately 14.58 m, which allowed performance of posterior analysis involving vegetation change detection.

In this study it is recognised that original digital numbers making up the imagery must be converted to physical values, considering the specificities of modelling or assessing the relationships between different remote sensing temporal sets. Multi-temporal sets should appear as if they were acquired within the same conditions, and consequently they can be analysed and compared using the same unit of measurement. Radiometric corrections were performed in two steps: first of all, digital numbers (DN) were converted to radiance using the approach and coefficients for post-calibration that were proposed by Markham and Barker (1986, 1987).

Multi-temporal analysis also requires that the measured radiance should be calibrated and converted to reflectance before further analysis (Robinove 1982). Therefore, radiometric correction is an essential and necessary procedure to ensure consistency and comparability over time

(Mather 1999). In the second step, radiance were converted to reflectance using the procedure and coefficients as proposed by Markham and Barker (1986).

Finally, radiometric restoration is another important step in the process, which deals with atmospheric and illumination differences between different image sets. Atmospheric correction is a complex process, which requires substantial modelling and computation. A major difficulty is obtaining appropriate input parameters for the models (Hadjimitsis *et al.* 1999b). The problem with the implementation of some of these procedures is that they need data about the sensor calibration parameters and/or atmospheric properties as inputs in the system. Frequently this kind of information is not available, which commonly happens if the study uses archived data. This study adopted a modified approach to the methodology proposed by Hall *et al.* (1991), dealing with atmospheric effects and illumination geometry in a relative way. The method does not require atmosphere radiative transfer codes or *in situ* measurements of atmospheric optical properties. However, the authors commented that in their tests, it performed better for the visible and near infrared bands than for the midi-infrared bands on TM, although results should be site specific. The proposed methodology has advantages, like: a) relative simplicity and easy of computation in a Windows environment; and b) it is image-based.

A review of suitable methods for correcting possible topographic effects was accomplished, during the pre-processing phase of this work. That literature review conducted to the decision of not performing any specific terrain corrections, mainly due to the assumption that the correction is necessary only if the study area has relief differences greater than 150 to 200 m (USDA 1995). The study area is characterised by gentle relief, showing some elevations of around 100 m. It was assumed that techniques such as image differencing and PCA, that were adopted in the study should also cope with the topographic effect. However, terrain effect was present in some areas, as mentioned and discussed in **Chapters 5, 6** and **7**.

This chapter also highlighted the procedures involved with the processing methods using the SPRING environment. SPRING is a Brazilian software with high potential for use in developing countries. It is an innovative database of geographical data of second generation, developed by INPE (National Institute of Space Researches) for UNIX and Windows environment. The basic motivation for the development of SPRING was integration of data and ease of use. In the first case, it was verified that complexity of the Brazilian environmental problems requires a high capacity for data integration of satellite images thematic maps, and numeric models. Also, many of the available systems in the national market are highly complex to use and require long-term learning. SPRING is freeware, eliminating economic concerns for institutional use. The programmes used in this study for the geometric and radiometric corrections were especially prepared by the author of this work.

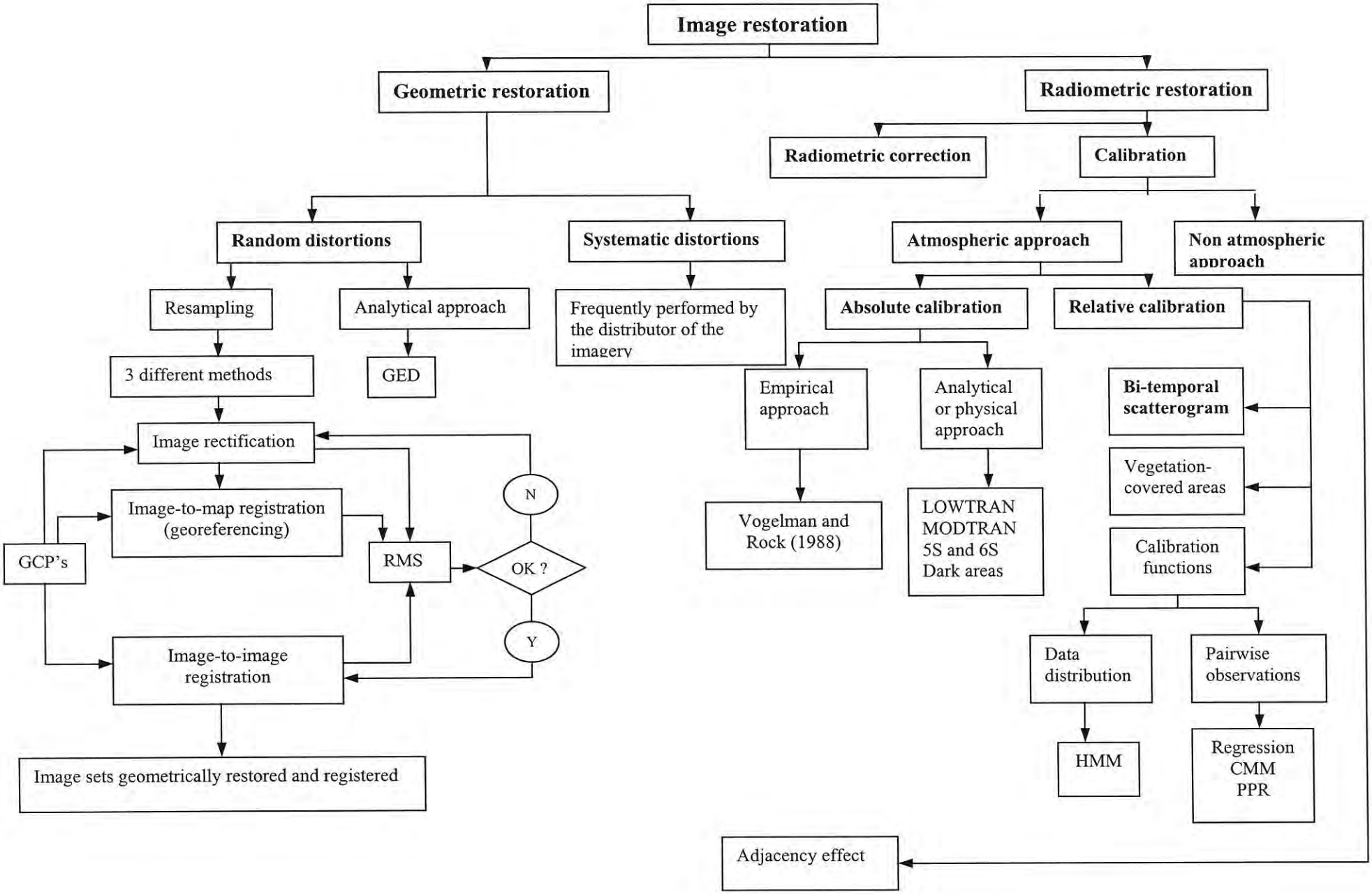


Figure 3.14 Scheme of stages for geometric restoration and possible approaches for radiometric restoration

Chapter 4

SPECTRAL CHARACTERISTICS OF THE REPRESENTATIVE FOREST TYPES

4.1 Introduction

Interpretation of remote sensing data obtained from space must take into account atmospheric conditions (discussed in **Chapter 3**) as well as the spectral characteristics of the surfaces to be investigated (Szekielda, 1986).

The radiation received by remote sensors is the result of the interaction of solar radiation with features on Earth. The main assumption of optical remote sensing (RS) of an object is that its radiation can be deciphered to determine some characteristics of that object. Vegetation is mainly recognised by its foliage, but the use of RS on vegetation patterns depends on the knowledge of spectral properties of plants and eventually depends on their interactions with soil.

In this study, criteria for discrimination of different forest patterns in *Araucaria* forest are investigated, based on the spectral response of each forest type. *Araucaria angustifolia* (paraná pine) is here studied as a key species in the process of classification refinement of the different forest types.

Facts relevant to the success of the discriminating process are: a) *Araucaria angustifolia,* the target species in this study, is one of the two **coniferous** species of its biome. It is a straight tree with almost cylindrical bole and commonly grows to heights of 30-50m, eventually with a 2m diameter at breast height. It belongs to the **upper-most canopy layer** and in the pristine areas *Araucaria* forest gives the impression that it is formed only by paraná pine. The second coniferous species is *Podocarpus lambertii* Klotz, which is not an overstory species but which normally occurs in-groups and it is not common in the Irati National Forest (INF) area; b) *Araucaria angustifolia* is a **dominant** species (representing usually between 15 - 40% of the forest trees) and the forests also support a complex, variable, and regional ecosystem composed of many **deciduous** species, some of which are endemic to this forest type; and c) The selected area belongs to a protected area (Irati National Forest - INF) with *Araucaria* Forest (natural forest), *Araucaria angustifolia* plantations and also plantations with exotic species (mainly *Pinus elliottii* Engelm and *Pinus taeda* L.). Pine plantations are very common in the *Araucaria* forest region, and are being used as a substitute to the *A.angustifolia* timber, after the remaining areas became poorer due to shifting cultivation, uncontrolled logging and unplanned land use. As coniferous species, **Pine species** can

present spectral responses similar to the *A.angustifolia* response. This can be true, especially when related to *A.angustifolia* plantation response, since trees were normally treated to have uniform structure and age classes. In this case, the architecture of the canopy can be the discrimination factor in the process. Pine species have cone-shaped crowns (pointed tops), and paraná pine presents rounded crowns. The area occupied by each crown is also much bigger in the *Araucaria* stands. As the stands are dense, the arrangement of canopy produces a **pattern** that is distinct for these species. The protected area of INF is considered an undisturbed area and it did not suffer any kind of forest management in the past 40 years.

In this Chapter, the reflective properties of the present forest types are discussed together with the strategy that was used to capture the different spectral signatures of each potential forest class and to analyse the differences between them. It is shown that *Araucaria* forest and *A.angustifolia* plantation are spectrally distinguishable from each other and from other conifer plantations like pine by means of the typical reflectance curves of each forest type built with data from the National Forest. Structural factors and spectral response relationship' were analysed using data Pine stands.

4.2 Physical and physiological properties of plants (Reflective properties of tree foliage)

4.2.1 Sources of energy

Remote sensing devices detect and record *radiant energy*, called electromagnetic radiation (EMR). This energy is emitted by or reflected from ground features and is transmitted to the sensing instrument in the form of waves (Robinson *et al.* 1995). Wavelength (λ) is the distance between successive wave peaks and frequency (ν) the number of wave peaks passing a given point in one second. Wavelength and frequency are inversely related, or $c = \nu\lambda$, where c is the constant velocity (3×10^8m/sec) of electromagnetic energy in a vacuum such as outer space (Lillesand and Kiefer 1994). **Figure 4.1** shows an electromagnetic wave. The range of possible EMR wavelengths and frequencies is termed the electromagnetic spectrum. For practical reasons, it is divided into energy regions with component wavelengths which are similar in some respect (Robinson *et al.*1995).

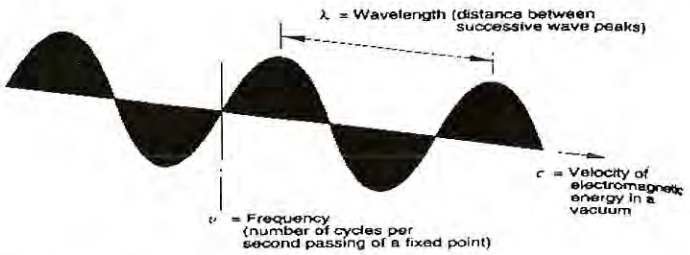


Figure 4.1 Electromagnetic wave

Source (Robinson *et al.* 1995)

Two objects may be indistinguishable in one portion of the electromagnetic spectrum (because their reflectance is similar in this portion), but highly separable in another (where their reflectance differs more markedly). To help capture the spectral uniqueness of ground objects, many sensors are designed to collect information in specific regions of the electromagnetic spectrum (USDA 1995). The electromagnetic spectrum is very broad (**Figure 4.2)** and not all wavelengths are equally effective for RS purposes (Eastman 1997).

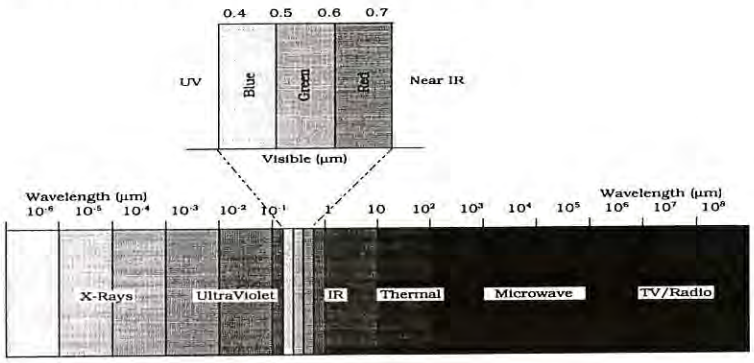


Figure 4.2 Regions of the electromagnetic spectrum
Source (USDA, 1995)

According to Howard (1991), the exact limits of the visible spectrum cannot be defined precisely, since it depends on the spectral sensitivity of the human eye. The boundary of the visible spectrum has been variously placed between 0.38-0.40 μm and 0.70-0.80 μm and even at 0.835 μm. The following scale (**Table 4.1**) is an adaptation of Howard (1991), Rosa (1992), Robinson (1995), USDA (1995) and Campbell (1996) scales:

Table 4.1 Limits of electromagnetic spectrum

Division	Limits
Gamma rays	values < 0.03 nm
X rays	values between 0.03 nm and 300 nm
Ultraviolet Spectrum	occurs at wavelengths below about 0.38 μm.
The visible Spectrum	constitutes a small portion of the full electromagnetic spectrum and it is defined approximately from 0.38 μm through the basic colours of blue (0.40-0.50 μm), green (0.50-0.60 μm) and red (0.60-0.70 μm).
The Infrared Spectrum	Wavelengths longer than the red portion of the visible spectrum are designed as the infrared region. It is large relative to the visible region, as it extends from 0.72 to 15 μm. Near infrared, mid infrared and far infrared can be recognised in this region • Near infrared (Near IR) starts approximately at 0.70 μm • Mid infrared (Mid-IR) is arbitrarily considered as commencing at 1.5 μm (although sometimes placed at 1.3 μm). • Far infrared (Far IR) commencing at 3.0 μm and extending to the range of microwaves commencing at 1 mm (3000 μm or 0.3kHz).
Microwave radiation	From 1 mm to 30 cm
Radio	Limits ≥ 30 cm (> 100 cm, as mentioned by Rosa (1992))

4.2.2 Spectral characteristics of natural systems

The study of the reflective properties of natural surfaces is not only important to those involved with remote sensing applied to forest resources, but also to many other scientists including plant physiologists and environmentalists concerned with the absorption of solar radiation by plant foliage (Howard 1991). The author also stressed that an understanding of the biophysical factors constraining the operation of passive Remote Sensing Systems and influencing image analysis is basic to the successful use of RS in forestry.

When the flow of energy (sunlight) is incident on an object, it interacts with its constituent material. Because of this interaction, the energy can be absorbed, reflected or transmitted, total or partially and, consequently, absorption, reflection and transmission are the three radiometric properties of the objects (**Figure 4.3**).



Figure 4.3 Radiometric properties of the objects: Absorption (A), Reflection (R) and Transmission (T)
Source (CCRS 1999)

The fraction of the incident flow that is reflected is the reflection factor or reflectance (ρ) of this object. It refers to the ratio of the total (hemispherical) radiation reflected by a body to the incident radiation of the body (Howard 1991). The absorbed fraction expresses the absorption (α) or absorptance factor and that transmitted part expresses the transmission factor or transmittance (τ) (Slater 1980). The total incident energy will interact with the surface in one or more of these three ways (CCRS 1999).

In SR, the reflected portion has high importance because it is the one registered by the sensors and, consequently, it is possible to measure the radiance reflected from targets. The relationship between reflection, absorption and transmission can be expressed as $\rho\%+\alpha\%+\tau\% = 100\%$, where 100% is the incident energy (Howard 1991).

According to Lillesand and Kiefer (1994) two points concerning this relationship should be noted. Firstly, the proportions of energy reflected, absorbed, and transmitted will vary from different earth features, depending on their material type and condition. These differences permit to distinguish different features on an image. Second, the wavelength dependency means that, even within a given feature type, the proportion of reflected, absorbed, and transmitted energy will vary at different wavelengths (**Figure 4.4**).

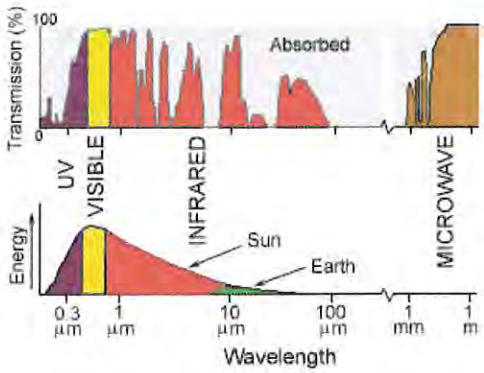


Figure 4.4 Interrelationships between energy sources and atmospheric absorption characteristics. Adapted from Lillesand and Kiefer (1994) and CCRS (1999)

4.3 Water, Soil and Vegetation

Spectral characteristics in the visible part of EMS differ significantly between targets. According to Szekielda (1986), for the identification and choice of proper wavelengths in monitoring natural surfaces with RS techniques, it is important to know the reflection or emission spectra of the target to be investigated. With this investigation, there is a chance of success in the discrimination between different formations. Ground objects selectively reflect and absorb electromagnetic energy due to differences in the molecular composition of their surfaces (Robinson 1995) and those characteristics make their *spectral pattern* predictable. There is a controversial term used for graphing the typical reflectance at each wavelength within one or more bands (Robinson 1995) and it is called *spectral signature*. It is controversial because some authors consider that for the beginning student the term can be misleading since it implies distinctiveness and a consistency that seldom can be observed in nature (Campbell 1996). There are no fixed spectral signatures for specific targets. They can change over time, over distance and local conditions but the term spectral signature conveys a more general concept that is useful (Campbell 1996) because each target has a general tendency of uniqueness that might be the key for its discrimination from other targets. According to the same author, this is the fundamental concept of spectral signature, or the notion that features display unique spectral responses that would permit clear identification, only by spectral information. Obviously the subject is not that simple and only few, if any, objects in nature can be uniquely identified at the desired levels of detail using only the spectral signature concept. The study of spectral properties of different forest types in the *Araucaria* forest region is an important part of this research. Distinct spectral signatures are expected for each forest type even if more than one species of conifers are presented in the area.

4.3.1 Water

The first fact to be mentioned is that water absorbs energy in two different manners, as a feature *per se* (such as lakes and streams) and as an element contained in vegetation or soil. Lillesand and Kiefer (1994) considered that probably the most distinctive characteristic of the spectral

reflectance of water is the energy absorption at near-infrared wavelengths. It is however, a complex matter because the reflectance properties of a water body are a function of not only the water. It is affected also by the material present in the water (or presence and concentration of dissolved and suspended organic and inorganic materials) and is affected by the water depth. Another aspect of the subject is that water in nature appears under different physical states (liquid, clouds and snows) that influence in a fundamental way its spectral reflectance (Novo 1988).

Although water spectral response involves complex energy/matter iteractions, its reflectance curve is less complex than the vegetation curve. The increase in the moisture contents results in a general decrease of the reflectance. Consequently, in the mid-infrared region the reflectance is controlled by water absorption bands (Rondeaux *et al.* 1996). The characteristic spectral reflectance curve for water shows a general reduction in reflectance with increasing wavelength, so that in the near-infrared the reflectance of deep, clear water is virtually zero (**Figure 4.5**) (Mather 1999). This fact increases its discrimination from soil and vegetation, since both have high reflectance values in that region.

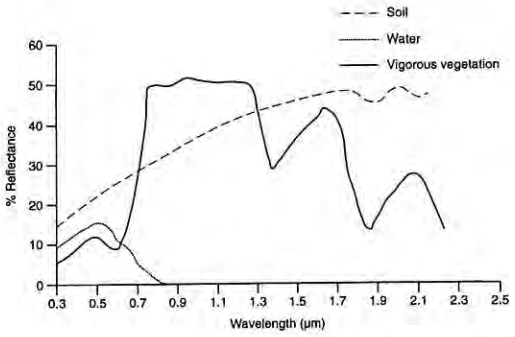


Figure 4.5 Typical spectral reflectance curve of vegetation, soil and water
Source (Mather 1999)

4.3.2 Soils

According to Mather (1999), the spectral reflectance curves of soils are generally characterised by a rise in reflectivity as wavelength increases. Its shape is the opposite of the spectral reflectance curve for water (**Figure 4.5**). In general, soil reflectance is relatively low (less or equal to 10%) in the blue channel, but increases monotonically with wavelength through the visible and near-infrared regions (Rondeaux *et al.*1996). This is a generalisation, however, because these optical properties vary with each soil type. Soil is, in reality, a complex mixture of many materials, with different physical/chemical characteristics. This affects its spectral response, sometimes drastically. For the moment, it is important to mention that searching the interactions of solar radiation with soil has led to the development of vegetation indices (VIs). VIs can be divided into two broad categories: slope based VIs and distance-based VIs. The former are arithmetic combinations that focus on the contrast between the spectral response patterns of the red and near-infrared regions (Thiam 1997). The most used example of slope-based VIs is the Normalised Difference Vegetation Index (NDVI). NDVI was calculated and it is analysed in **Chapter 5**. The distance-based VIs

measure the degree of vegetation cover present by gauging the difference on any pixel's reflectance from that of the bare soil. According to Thiam (1997), a plot of bare soil pixels of varying moisture levels in bi-spectral plot will tend to form a line known as the soil line. When vegetation is being analysed, it is important to remember that soil background is progressively obscured as vegetation canopy cover increases. Then, vegetated pixels tend towards increasing perpendicular distance from the soil line. The general purpose of the distance-based VIs is to cancel the effect of soil brightness when vegetation is sparse. The Perpendicular Vegetation Index (PVI) is one example of the most used indices in this category. Thiam (1997) reinforced the importance of this type of approach when sparse vegetation, like deserts and arid regions are analysed. However, the use of PVI was not considered strategic for the study area, because the natural vegetation is still considered dense and the man-made forests are mature.

4.3.3 Vegetation

Leaf structure contributes significantly to the appearance of vegetation. The leaf (or needle) is a complex structure with its internal tissues and external morphology varying between genera and species (Howard 1991). The author commented that even for the same species there might be differences in leaf structure. More precisely, on the same tree differences will exist between leaves fully exposed to sunlight (i.e. sun leaves) and shade-bearing leaves (i.e. shade leaves). The spectral reflectance of vegetation depends on many factors. Atmospheric conditions, species, soil, Leaf Area Index (LAI), phenological aspects, biomass, leaf structure, sensor type and canopy cover among others can contribute decisively to the aspect of the curve (Rosa 1992).

According to Lillesand and Kiefer (1994), spectral reflectance curves for healthy green vegetation usually manifest the "peak-and-valley" configuration (**Figure 4.5**). Chlorophyll strongly absorbs radiation in the red and blue wavelengths but reflects green wavelengths (CCRS 1999). This is why in the visible region the vegetation reflectance is relatively low (valley) and why human eyes perceive vegetation as green in colour. In the near-infrared portion of the spectrum, the reflectance of healthy vegetation increases dramatically. In that region, a plant leaf reflects 40 to 50% of the energy incident on it (peak). It is important to mention that absorption in this spectral region is less than 5%. In this study we are interested in the discrimination between *Araucaria angustifolia,* a coniferous species (in plantations and in natural stands) and deciduous species and/or other conifers (pine plantations). The spectral response for coniferous and deciduous species can overlap in most of the visible region of the spectrum. Sensors that can record near-infrared energy improve the discriminating capability (Lillesand and Kiefer 1994), since deciduous trees have higher infrared reflectance than conifers and appear much lighter in tone.

The reflectance characteristics of earth surface features may be quantified by measuring the portion of incident energy that is reflected. This is measured as a function of wavelength and is called spectral reflectance ($\rho\lambda$) (Lillesand and Kiefer 1994). A graph of the spectral reflectance of an object as a function of the wavelength is termed spectral reflectance curve (**Figure 4.5**).

4.3.4 Spectral characteristics of vegetation

Finding distinctive spectral response patterns is the key to most procedures for computer-assisted interpretation of remotely sensed imagery (Eastman 1997), but it is extremely difficult because of the factors that influence and determine each response. Surface materials do not have necessarily distinctive spectral response patterns that can allow them to be reliably detected by visual or digital means. Species identification is a complex problem of interpretation even for aerial photographs, where some interpretation keys have been produced over the years (Barrett and Curtis 1999). However, using multispectral data it is possible to consider the fact that chlorophyll is a very strong absorber of red visible wavelengths and, consequently, the structure of the plant leaves is correlated to the near-infrared portion of the spectrum. The first step should be the detection of differences between conifers and deciduous species, when both are present in the same area (**Figure 4.6**).

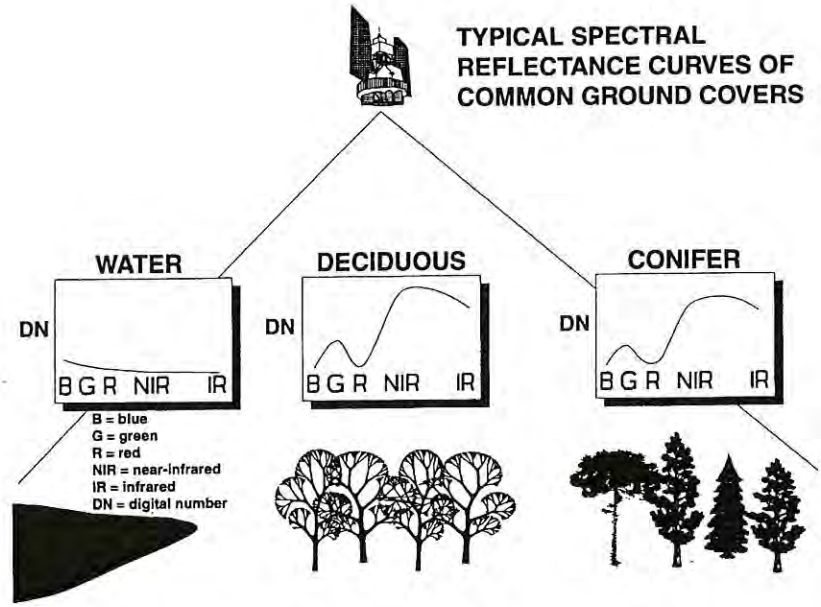


Figure 4.6 Differences between conifers and deciduous species spectral response
(Adapted from USDA 1995)

4.4 Spectral responses of the forest types in INF

Analysing vegetation by selected characteristics, such as crown closure, species, structural variables or age is an important step in the process of using digital satellite imagery. In the study area, the main group of species or biome (features) that can be studied separately by their patterns is summarised in **Table 4.2**. According to Mather (1999), *pattern* is a set of grey values for a single pixel measured in a number of spectral bands; *features* are the properties of the pixels derived from the spectral bands (such as texture and context), or not (such as land-surface elevation and soil type).

Table 4.2 Main forest features in the study area

Type of vegetation	Characteristics
Pinus elliottii	Coniferous
Pinus taeda	Coniferous
Araucaria angustifolia plantation	Coniferous
Araucaria forest	Mixed forest with coniferous and deciduous trees

As mentioned in **Chapter 2**, other species of the genus *Eucalyptus* sp. and *Cunninghamia* sp. are also present in the area, but only in a few small stands. **Figure 4.7** shows areas and distribution of the main group of features.

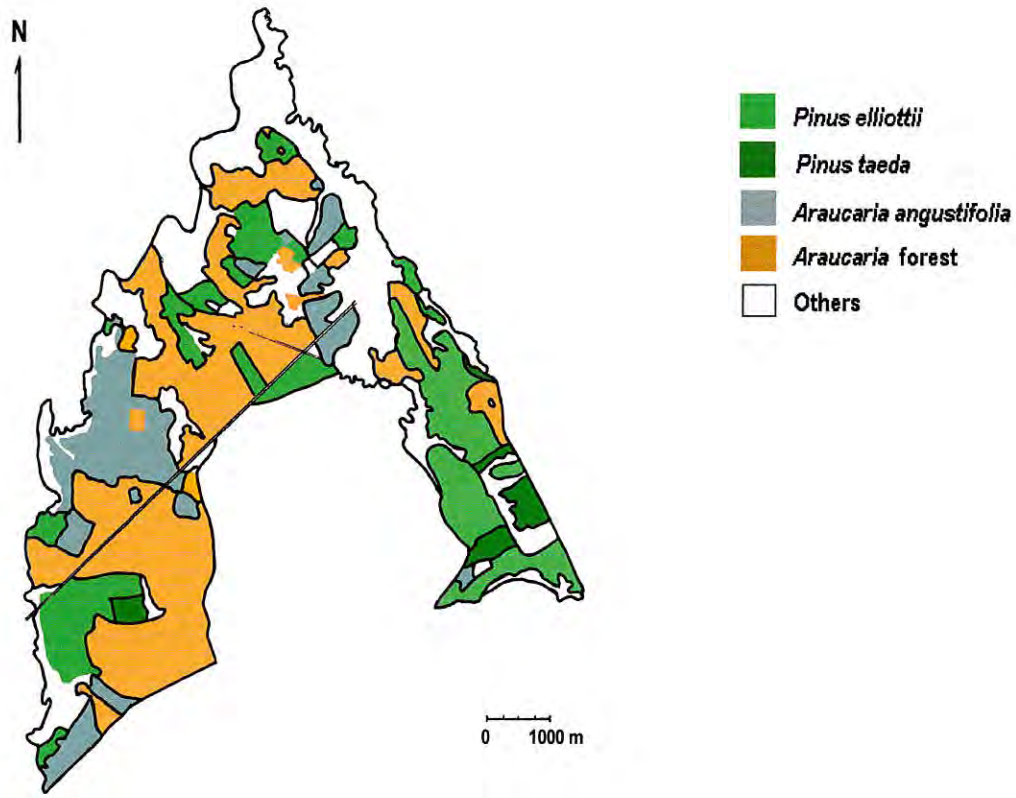


Figure 4.7 Main features of the INF
Source: Map of INF elaborated by FUPEF (1986)

Three main aspects were studied in the analysis of the spectral response of forest types in the study area: a) spectral responses and structural variables relationships; b) spectral responses and species composition and c) spectral characteristics of each feature and their separateness.

4.4.1 Pine species - spectral responses and structural variables relationships - Experiment 1

An experiment was set with the objective of analysing the spectral response of the two species of the genus *Pinus* (*Pinus elliotti* Engelm. – slash pine and *Pinus taeda* L. – loblolly pine) present in

the study area. Differences between species, ages and structural variables were verified. Some of these aspects were very important to the comprehensiveness of the differences between those coniferous species and also fundamental in the comparison of pattern differentiation between *Pinus* sp. and *Araucaria angustifolia* plantations or *Araucaria* forest. The experiment also sought to test the usefulness of quantitative remote sensing for forest management purposes in that region. It examined the relationship between structural variables such as diameter at breast height (DBH), height, basal area and volume and the spectral responses. Field-sampled information was correlated and regressed against Landsat TM reflectance values and transformed data.

4.4.1.1 Generalisation about the genus in Brazil

Different cycles of wood exploitation occurred in Southern Brazil over the past centuries. The *Araucaria* forest was so imposing and extensive that it was considered inexhaustible. However, the inevitable happened and both area and abundance declined sharply. Therefore, in the 1960s, were adopted the policies of fiscal incentives for the Brazilian Forest Sector, to increase the forest base activity, considering that areas with forest plantations were not very significant in relation to the needs of forest replacement. Pine and Eucalyptus species were introduced, because of their characteristics related to adaptability and fast growth and as a pulp alternative to *Araucaria angustifolia.* The initial goal of area expansion for raw material supply and energy was reached and around 1.8 million hectares of Pine species were planted. Subtropical species, such as *Pinus elliottii* and *P.taeda* are now the most common species and are strongly concentrated in the Southern Region.

4.4.1.2 Fieldwork

At the INF, *P.elliottii* (45 stands) and *P. taeda* (8 stands) were installed as commercial forests since early 1960s, usually up 2,000 to 2,500 trees per hectare. Programmed thinning were not always done at the planned times. The stands are distributed all over the INF area, with different sizes and ages. The available field data were collected in 1988. The methodology adopted was the standard approach to forest mensuration described in IBAMA (1989). The adopted sampling strategy was the stratified random sampling when the strata were the stands and the sample units (usually three plots of 20-by-30 m per stand) were designed to be permanent plots. All individuals with GBH (girth at breast height) equal to or greater than 15 cm (5 cm of diameter at breast height - DBH) were measured and the results subsequently converted into DBH. Total heights of the first 25 trees in the plot were also determined using a clinometer. The bole volume was estimated by the Smalian method, using trees specially felled for this purpose. The exact location of each plot was determined and its central point recorded in 1998 by using the "Global Positional System"(GPS) from GARMIN which allows precision around 1.40 m. Geodesic co-ordinates of the point were collected in the central part of the plot whenever it was possible (details about the procedure can be found in **Appendix 2.2**). Constraints such as accessibility of the stand, difficulties to locate the plot and weak GPS signal under canopy were frequent and were responsible for the final number of considered observations. Therefore, 49 plots of *P.elliottii* and 7 plots of *P.taeda* were considered

in the preliminary analysis. These plots belonged to 28 stands of *P.elliottii* and 4 stands of *P.taeda*, ranging in area from 1 to 35 hectares. Further analysis revealed that stratification by age should be considered. The reason was that some of the younger stands did not suffer any thinning operation, whereas the older should be in the second or third thinning, which sometimes reduce dramatically the number of trees to improve growth. Therefore, young stands have had in average, approximately a double of the trees of the older stands and the trees had also much smaller mean DBH, due to competition and age. The stratification by age however, was not performed in this study due to the small number of young plots. Therefore, they were eliminated of the data set. The final set comprehended 44 plots of *P.elliottii* and 6 plots of *P.taeda*, which were joined in one unique data set composed by 50 observations. Descriptive statistics of the field data are presented in the **Appendix 4.1-Table 4.1.1**. The decision was made after the analysis that considered, as a precaution, the presence and absence of *P.taeda* plots. It was assumed that average differences among the two species of the same genus were small and as shown in **Appendix 4.1-Table 4.1.2**. Therefore, *P.taeda* data were included into *P.elliottii* limits. As it can be seen in **Appendix 4.1-Table 4.1.3**, correlation results were not greatly affected by the inclusion of *P.taeda* in the analysis.

4.4.1.3 Imagery information

The 1987 set, bands 1 to 5 and 7, was used (details about imagery acquisition and pre-processing methodology are mentioned in Chapter 3). A window of 1428 columns by 713 rows that includes the INF was considered in the study.

4.4.1.4 Location of the sample units

As mentioned before, the exact location of each plot was recorded by GPS. The spectral response of each of these plots was also recorded and consisted in the first data set of this analysis. However, the error (RMS) of 0.852 pixels that occurred in the process of rectification had to be considered, which means that the exact location of the plot should be in one of the neighbour pixels. To analyse the impact on the results and also as an extra precaution, the defined procedure was to recover the spectral values in the 3-by-3 window around the plot location, totalling nine observations per plot. On TM imagery, each pixel covers 30 x 30 m, or an area of approximately 0.09 ha. With the adopted procedure, the recorded area totalled almost 1 hectare (3-by-3 pixels = 0.81 ha) and in such area the variation among pixels should be high. Therefore, an alternative solution to the problem of the windows size when compared with the filed plot size (600 m^2) was set up. It consisted in analysing all possible combinations of 2-by 2 windows (0.36 ha) and to compare them with the first data set. Band 5 was used in this analysis. **Figure 4.8** presents the adopted strategy. The highlighted pixels were measured using the different arrangements showed in the picture and recorded as: mean and v5 ("a" and "b" groups), v1245 ("c" group) etc.

The procedure should be considered excessively detailed for a homogeneous area. However, even for even-aged stands, variations will occur, due to genetic differences (the material was not highly selected) and punctual fast growing due to the thinning regime adopted.

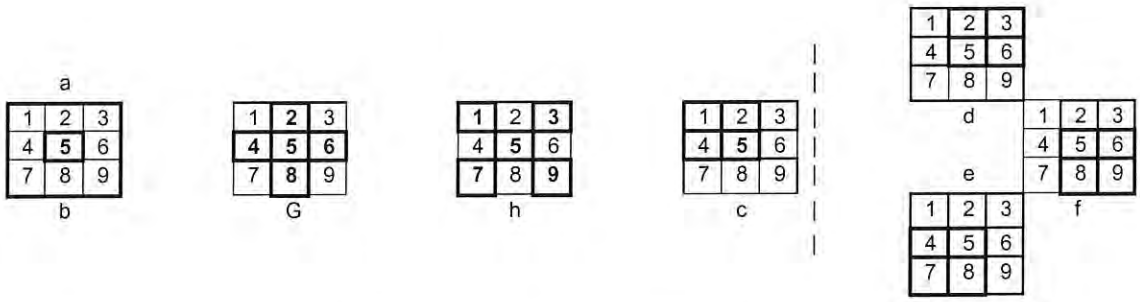


Figure 4.8 Central point determined by GPS (a), 3-by 3 window (b), combinations of 3-by 3 window (g and h) and combinations of 2-by 2 window (c,d,e and f).

4.4.1.5 Analysis of ground data - relationship between forest plot parameters

The statistical approach was designed to explore and analyse the relative impact of each ground variable on each of the six TM bands or on the calculated synthetic imagery based on them. As a first step in the analysis, the interdependence of ground data was studied (**Appendix 4.1-Table 4.1.3**). Analysis of these data showed that several of the variables were inter-correlated. Tree density (TREE) was negatively correlated (at the 95 per cent confidence level) with tree age (AGE, -0.437 and -0.345). This is certainly due to the management practice of systematic or selective thinning, with the consequent periodical removal of trees, although those practices were not conducted as planned during long periods at the INF. TREE was also negatively correlated to DBH (-0.681 and -0.721), HEIGHT (-0.057 and -0.083), as a natural consequence of initial density and thinning. DBH (0.525 and 0.379) and HEIGHT (0.520 and 0.544) were positively correlated with age, as expected, as they are related to tree growth. Both Basal Area and Volume were obtained from individual trees rather than from the average.

4.4.1.6 Spectral and structural relationships

4.4.1.6.1 Sample units

Using band 5 and the strategy mentioned in **Figure 4.8**, eight data sets were generated by collecting the spectral response in the different groups of pixels. To test the equality of means (reflectance) and to access the differences on those means ANOVA procedure with multiple comparisons was adopted, as showed in **Table 4.3**.

Table 4.3 Location of the sample units - ANOVA applied to the different data sets

One-way Analysis of Variance					
Source	DF	SS	MS	F	P
Factor	7	11.25	1.61	0.25	0.972
Error	392	2526.55	6.45		
Total	399	2537.80			

The F-test p-value of 0.972 indicates that there is no significant evidence at α=0.05 of differences among the means. These results indicate that, in the present study, v5 (the central point of the window and the pixel located by GPS) can be used as a reasonable approximation of the exact location of the plot (**Figure 4.9**).

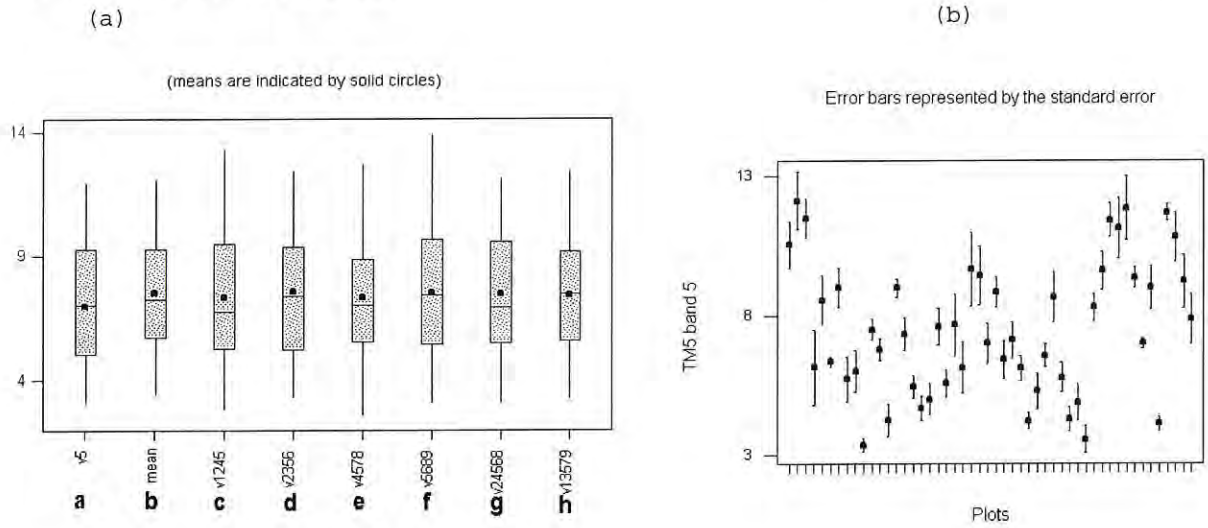


Figure 4.9 (a) Boxplots representing the magnitude of the differences between the selected data sets and (b)Interval plots representing the spread of data around the mean of the variable "Reflectance of band 5" on each plot

Where: v5 = Central point of the plot recorded by GPS ("a" group); Mean = Average of the 9 values on the 3-by 3 window ("b" group) allocated around v5; and v1245 ("c" group), v2356 ("d" group), v4578 ("e" group), v5689 ("f" group), v24568 ("g" group) and v13579 ("h" group) = Values summed following the graph on **Figure 4.8**

A boxplot with a graphic representation of the magnitude of the differences was produced to provide information about the shape and frequency of the distribution with additional information about tails of the distribution and the possible presence of outliers. According to Kitchens (1998) the box portion of the boxplot contains the middle 50% of the data, with the median dividing it in a half. In the case of the groups presented in **Figure 4.9**, the interpretation is that no outliers were presented. Groups "c" and "g" presented some skewing in the middle 50% of the distribution, but generally, it was symmetric in all eight groups. The rest of the data presented also symmetric distribution, but groups "c" and "f' were a little skewed, concentrating values greater than the others groups. The conclusion is that the group "a" should be considered as a non-tendency approach for catching up "reflectance of band 5".

An Interval plots graph were also produced to analyse the spread of data around the variable means of "reflectance of band 5" over each plot (**Figure 4.9**). The produced error bar extends one standard error ($s/\sqrt{n}$) above and below the symbol (solid circles) at the mean of each plot. There was no evidence that the variability within data is greater than the variability between plots (**Appendix 4.1-Table 4.1.4**).

4.4.1.6.2 Spectral responses

Further analyses were made using the values collected in the central point of the plot (v5 – group "a"). Again (**Table 4.4**), slash pine data seems to overlap loblolly pine limits. This confirms that they are not significantly discriminated from each other when the variable "reflectance of band 5" is being analysed. Log transformations for band 5 were also generated following Ripple *et al.* (1991) and Ardö (1992) which reported strong relationships between the logarithms of spectral data and volume. In the present study, generally, they did not increase consistently the correlation coefficients except for band 3. Because the results were not significant, they were not considered in the further steps of the analysis.

As expected, visible bands of TM (bands 2 and 3) were significantly inter-correlated (**Appendix 4.1-Table 4.1.6**). The coefficients for TM1 might suggest the possibility of interference due to atmospheric conditions. However, it must be considered the flat nature of the study area, the excellent visibility at the time of image acquisition (Minimum Visibility Level=7) and the fact that the imagery was processed in order to reduce these effects. Infrared bands (4 and 5) were also inter-correlated but the magnitude of this correlation was smaller than expected, suggesting relatively heterogeneity in the studied forest area, as mentioned by Ripple *et al.* (1991). Some other aspects will be discussed looking at each specific spectral band:

Table 4.4 Analysis of Variance for TM5

```
Source      DF        SS        MS         F        P
Spp          1     16.40     16.40      2.43    0.125
Error       48    323.60      6.74
Total       49    340.00

                                      Individual 95% CIs For Mean
                                      Based on Pooled StDev
Level        N      Mean     StDev  -+---------+---------+---------+-----
1           41     6.732     2.674  (------*------)
2            9     8.222     2.167        (--------------*-------------)
                                    -+---------+---------+---------+-----
Pooled StDev =     2.596           6.0       7.2       8.4       9.6
```

As mentioned in **section 4.4.1.2** slash pine data seems to overlap loblolly pine limits. This confirms that they are not significantly different from each other when the variable TM5 is being analysed. The significance (t) of the correlation [r] between TM5 and Basal Area (and Volume) was checked using the following test, where n is the number of observations:

$$t = \frac{r\sqrt{(n-2)}}{\sqrt{(1-r^2)}}$$

The obtained values (at the 0.01 significance level and n-2 degrees of freedom) were compared with the critical values of t, provided by Watts *et al.* (1996). Using this criterion, TM5 was considered to have high significant correlation with both selected ground variables (Basal Area and Volume).

4.4.1.6.2.1 Reflectance values of TM7

Middle-infrared region (Wavelength = 2.08-2.35 μm) and structural factors: band 7 presented reflectance values equal to zero in almost all the pine plots examined. Actually, the reflectance values resided in the interval zero-one in the majority of the analysed plots, as expected for conifers. Some authors have reported the sensitivity of TM7 in the process of identification of pine species, but it was noticed that some of those results referred to natural coniferous stands. Plantations are a special case of vegetation where floristic diversity is minimised and architecture of canopy is less complex. Jakubauskas *et al.* (1997) found promising results for TM7, in one area with lodgepole pine (*Pinus contorta var.latifolia*). Also working with pine plantation data, Souza *et al.* (1998) found significant sensitivity of TM7 on detecting differences in volume. Differences in sampling and local characteristics made comparisons difficult; for example, both authors analysed data at stand level, averaging the values by compartment, what probably means more heterogeneity on data than collection of just one observation, like in the present study, where TM7 was not able to detect differences among stands.

4.4.1.6.2.2 Reflectance values of TM5

Middle-infrared region (Wavelength = 1.55-1.75 μm) and structural factors: according to Ripple *et al.* (1985), there is an inverse relationship between vegetation amount and both visible reflectance and middle-infrared reflectance, what is explained by the highly absorbing plant pigments and strong absorption from water in the leaves respectively. LAI (Leaf Area Index) and TM3 and/or TM5 had a strong inverse curvilinear relationship in a study conducted by Spanner *et al.* (1990). A high correlation between reflectance in TM bands 5 and 7 has been noted in many studies, as reported by Ahern *et al.* (1991). This aspect was also stressed by Horler *et al.* (1986) who found TM5 and TM7 highly mutually correlated. In the case of the present thesis, TM5 and TM7 were significantly (at $p<0.05$) but not highly correlated (0.374 for *P.elliottii* and 0.393 for both, *P.elliottii* and *P.taeda)*, what might explain the negative results of TM7 in this experiment. Trotter *et al.* (1997) mentioned that the reflectance of Landsat TM shows a general negative correlation with wood-volume and volume-related parameters, with the exception that band 4 varies in different areas.

In the studied area, there was a significant negative correlation (**Appendix 4.1-Table 4.1.6**) between TM5 and VOLUME (from -0.706 only *P.elliotti* to -0.716 for both, *P.elliotti* and *P.taeda*) and TM5 and BASAL AREA (-0.642 and -0.639 respectively). It was the single band that showed the largest correlation against these ground variables. The significance of correlation between TM5 and ground variables was tested following a test proposed by Watts *et al.* (1996). The obtained values (at the 0.01 significance level and with n-2 degrees of freedom) were compared with the critical values of t. Using this criterion, TM5 was considered to be a highly significant correlation with both selected ground variables (Basal Area and Volume).

These results are coherent with the literature, as mentioned by Sousa *et al.* (1998). Kazmierczak (1992) actually found that TM5 seemed the isolated band to constitute the best independent

variable. De Wulf *et al.* (1990) using information from SPOT-1 and Corsican pine (*P.nigra* Arn.) found that only the infrared band displayed a consistent relationship with ground data. Ardö (1992) also mentioned the sensitivity of TM5 to volume changing for conifers (the relationship was stronger for compartments with small volume than for compartments with large volume). Jakubauskas *et al.* (1997), working with lodgepole pine (*Pinus contorta*), mentioned that several lodgepole factors are predictable from remotely sensed data, and TM5 and TM7 or a combination of visible bands and middle-infrared bands presented the best results. Horler *et al.* (1986) also mentioned the capability of TM5 for detecting forest vegetation density partly because it is less subject to atmospheric interference.

4.4.1.6.2.3 Reflectance values of bands 1,2,3 and 4 and structural factors

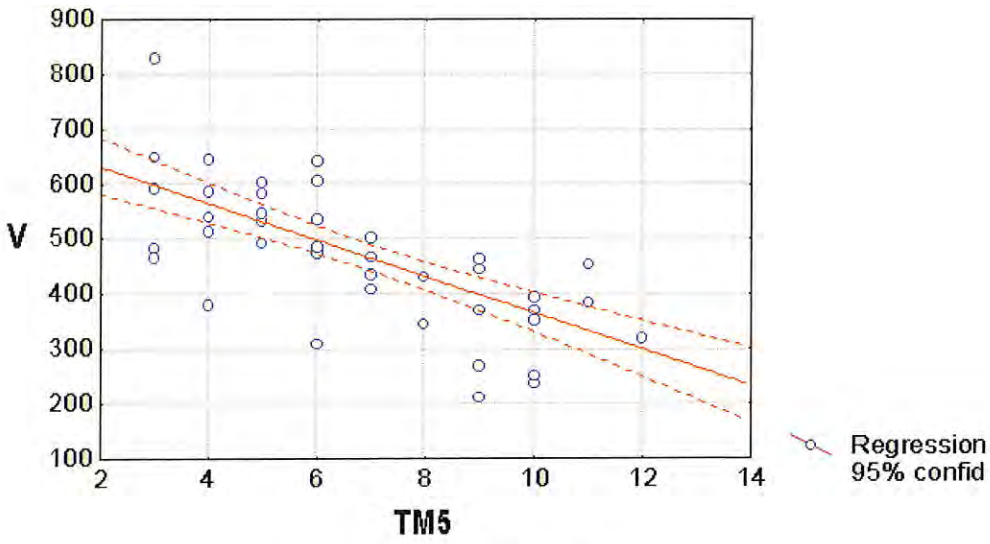
The visible bands TM1 or blue band (Wavelength = 0.45-0.52 μm), TM2 or green band (Wavelength = 0.52-0.60 μm) and TM3 or red band (Wavelength = 0.63-0.69 μm) were only weakly correlated with all the field variables. TM4, located on near infrared region (Wavelength = 0.76-0.90 μm) was expected to be useful for determining vegetation types, vigour and for biomass content (Lillesand and Kiefer 1994). It seems that band 4 presents better performance in heterogeneous stands due to diversity in species and diversity in the architecture of canopy, as found by Ahern *et al.* (1991). It was their most important independent variable to predict spruce (*Picea sp*) and balsam fir volume in natural stands. However, Trotter *et al.* (1997), using TM data from a *Pinus radiata* plantation, found that the two bands most highly correlated with wood volume were bands 3 and 4, although the significant correlations were weak (at the pixel scale). Spanner *et al.* (1990), using mixed coniferous stands data, mentioned that the radiance of band 4 was elevated in open stands with highly reflective understory vegetation and was low in old growth stands with a less reflective forest litter background and shadow. In the INF, as expected for pine plantations, the differences in reflectance were much smaller for bands 1,2 and 3 (visible region) than for band 4 and 5 (infrared region).

4.4.1.6.3 Regression analysis

Based on the fact that some transforms (operations involving two different bands) should improve the results of correlation between structural and spectral variables, some analysis were made and are presented in **Appendix 4.1-Transforms**. The results, however, were less promising than expected, when compared with those obtained by the untransformed band 5. Forward stepwise linear regression procedure, including the TM original bands (bands 1 to 5) and the statistically significant transformed images (**Appendix 4.1-Table 4.1.7**) was therefore performed but the only significant variable was still band 5 (**Table 4.5** and **Figure 4.10**).

Table 4.5 Linear regression between TM5 and Volume

Dep. Var.: Volume, with N=50				
Multiple R: 0.7161		F= 50.51		
R^2: 0.5128		df= 1,48		
Adjusted R^2: 0.5026		p=0		
Standard Error of estimate: 86.16				
Regression Summary				
Predictor	Coef.	Std.dev.	T	p-level
Constant	698.11	34.91	20.0	0
TM5	-33.213	4.673	0	0



Volume = 698.11 - 33.2 * TM5
Correlation: r = -0.7161

Fig.4.10 Bivariate scatterplot - Linear regression between Volume in m^3 (V) as the dependent variable and reflectance of band 5 (TM5) as the independent variable

4.4.1.6.4 Summary of results and conclusions for Experiment 1

Several ground variables were inter-correlated. Volume was chosen as being the dependent variable because it is the variable most related to biomass. However, basal area should be also considered, as it presented correlation almost as good as volume. Actually, volume is a variation of basal area, which includes in its formula one additional dimension, the commercial or total height. For the purposes of the experiment, it was assumed that the average differences (at ground level and at spectral level) between slash and loblolly pine were small, but the species presents spectral signatures different enough to be discriminated and classified separately, as it is discussed in **Chapter 6**. The alternative stratification by species was not possible due to the few observations for *Pinus taeda* at the INF.

Preliminary analysis using *Pinus elliottii* data suggested that stratification by age should be necessary. Some of the younger stands did not suffer any thinning operation, whereas some of the mature areas were in the second or third thinning, which sometimes reduce dramatically the number of trees with the objective of improving growth. Those young stands clearly formed a new

stratum. However, the number of observations was too few to be included in the analysis and for this reason, data from the youngest stands were eliminated from the database.

Band 7 (middle-infrared region) and its possible transforms presented reflectance values close to nil in almost all of the plots, being consequently not able to detect volume differences among stands. Band 5 (infrared region) was the single band that showed the largest correlation against basal area and volume. This result is in agreement with many other researches involving conifer species what stresses the importance of the middle-infrared region for biomass detection using remote sensing. Some authors reported that, after certain age (or volume) differences in biomass were not detectable. However, it seems that, for the analysed data set, that "saturation effect" did not affect the results, since differentiation in volume was detected in all stages of crown closure.

To analyse the impact of the geometric restoration on the exact location of each plot, a strategy was designed, involving 3-by-3 and 2-by-2 windows around the central pixel. It was concluded that the area was homogeneous enough to be sampled only by the central point of the window. Stepwise regression was performed including all the original bands and respective transformations, but the only significant variable was band 5. It was concluded that there is a potential for quantification of Pine plantation at INF using Landsat TM but the standard error of estimate was 86.16 m^3 ha^{-1}. In the fieldwork, the standard error of estimate for the same age classes was 57.33 m^3 ha^{-1}.

4.4.2 *Araucaria angustifolia* plantation - spectral responses and structural variables relationships - Experiment 2

An experiment was set up in order to analyse patterns of spectral responses of *Araucaria angustifolia* in plantations. This is the appropriate environment also to analyse possible approaches to capture the pure pixel of *Araucaria angustifolia*, important information for the success of the linear mixture modelling (LMM), as mentioned in **Chapter 6**.

4.4.2.1 Generalisation about the genus in plantations

The legal requirements to replace logged trees in natural stands, fiscal incentives and other initiatives have resulted in around 200,000 ha being planted with *Araucaria angustifolia.* Growth of individual trees has been much faster in plantations than in natural areas but, in recent years, the expansion of plantations has declined sharply to about 300 ha each year. Instead, *Pinus elliottii* , *Pinus taeda* and *Eucalyptus* spp. have been introduced since they are less demanding and have higher growth rates (Shimizu *et al.* 1980). For the same reason, tree improvement planning (progeny and provenance trials) with *Araucaria angustifolia* have remained limited, although preservation of the valuable genetic stock was established *ex situ*.

The objective of the experiment was to characterise the spectral response of paraná-pine plantations, based on stands ages and different stocks. The number of information pairs (stands)

was considered too small to adopt the same strategy as used for pine, where it was possible to predict volume based on satellite information. Apart from this, the environment was more complex, because of the age of stands and the presence of eventual understory vegetation.

4.4.2.2 Field work and imagery information

There are 26 stands of *Araucaria angustifolia* planted in the INF. They were installed since the early 1940's, as an alternative to the use of natural forests and a long time before the introduction of the exotic species of pine. In the first decades, the majority of the stands were managed as commercial plantations. More recently, because of the success of pine plantations, they received no more silvicultural interference and the understory was left intact, being then eventually invaded by pioneer forest and non-forest species. Nowadays the stands still have around 400 trees of *Araucaria angustifolia* per hectare, in contrast with natural stands, where the number of trees per hectare varies very much. In the fieldwork, around 44 trees, in average (10% of the number mentioned for plantations) of *Araucaria angustifolia* per hectare were found, and 187 trees per hectare of the other deciduous species (DBH larger than 20 cm). The available field data were collected at the same time as the pine species data set, as mentioned in **section 4.4.1.2**. The methodology adopted was reported by IBAMA (1989) and it can be summarised in: a) sample units (usually two or three plots of 20-by-30 m per stand, depending of its size) were installed; b) all individuals with GBH (girth at breast height) equal to or greater than 5 cm of diameter were measured and the results subsequently converted into DBH (diameter at breast height); c) total heights of the first 25 trees in the plot were determined using a clinometer; and d) bole volume was estimated by the Smalian method, using trees specially felled for this purpose.

The imagery data comprised the 1987 set, bands 1 to 5 and 7 (imagery acquisition and pre-processing methodology are detailed in Chapter 3). The same window of 1428 columns by 713 rows used to analyse pine data set was also used in the study.

4.4.2.3 Location of the sample units

As mentioned before, the location of pine sampling units were made using GPS. The same procedure was not followed for the *A.angustifolia* plantation data. In this case, 21 of the total 26 stands were located in the imagery. The five stands not included in the study were too small or had suffered severe interference in the past. For the location of stands, aerial photographs, GPS points and the map of the INF were used as ancillary data. The adopted strategy was: a) in the imagery, to draw a mask around each stand, making it smaller than its actual size and so being sure that the information really belonged to that specific stand; b) to write a program in LEGAL language to capture the mean spectral response inside the mask, for each band; and c) to correlate this information with the averages of field information for each stand.

4.4.2.4 Ground data – relationships between forest plot parameters

Descriptive information about the available data is summarised in **Table 4.6**, presenting a snapshot of the data set variability. The stands were originally planted as a monoculture, and still have 10 times more trees of *A.angustifolia* than *Araucaria* forest. However, It is a complex data set because of, among other reasons, the tree reduction effect created by thinning during the commercial forest management phase and because of the differences in age of the stands. It presented also too few observations to analyse, in all aspects, the impact of tree number and age interaction over variables like Basal Area and Volume.

Table 4.6 Descriptive statistics of field data (*Araucaria angustifolia*) collected from stands within forest management units

Species: *Araucaria angustifolia* : Number of stands = 21				
Variable	Mean	St.Dev.	Maximum	Minimum
Age (years) - AGE	34.48	6.47	45.00	14.00
Density (trees ha^{-1}) - TREE	407.70	169.10	825.00	216.00
DBH (cm) - DBH	24.51	6.07	33.80	10.00
Height (m) - HEIGHT	17.55	3.13	21.40	6.18
Dominant Height (m) - DOMH	19.54	2.99	22.70	8.09
Basal Area (m^2 ha^{-1}) - BA	18.49	6.31	30.91	6.50
Volume (m^3 ha^{-1}) - VOL	200.90	79.30	361.90	30.80

4.4.2.5 Spectral and structural relationships

The results of the correlation analysis are presented in **Table 4.7**. *A.angustifolia* presented a significant correlation between field and spectral variables in all bands, but this is probably a site-specific indication. For *A.angustifolia,* TM5 was still the band with the highest correlation, but the other bands from the visible or infrared region also performed reasonable correlations. TM7 presented reflectance values equal to or higher than 1 more frequently in the paraná-pine plantation data set than in the pine data set. The interpretation should be that paraná-pine in plantations might present reflectance values that reflect its condition as a conifer but also some interference of the understory presence, captured by TM7. There were a few younger stands in the data set, but the effect of age was not analysed (too few observations) although their mean reflectance values were higher than in the older stands. Mean DBH is a variable potentially useful for the next steps of this study and presented significant correlation and highest values with TM5 and TM7. The same tendency was found with its highly correlated basal area, which presented a significant inverse relationship with all bands (but TM1) and high values in TM5 (nearly the same as obtained for pine). Basal area incorporates the concept of spatial distribution, instead of variables such as mean diameter. Mean HEIGHT presented also even higher correlations with bands 3,5 and 7, which certainly contributed to the reflectance values found when volume (VOL) was analysed.

Table 4.7 Correlations between forest stand parameters and spectral responses (Pearson)

Species: Araucaria angustifolia (plantation)							
	AGE	TM1	TM2	TM3	TM4	TM5	TM7
TM1	-0.534*						
TM2	-0.337*	0.821*					
TM3	-0.542*	0.818*	0.845*				
TM4	-0.008	0.501*	0.549*	0.265			
TM5	-0.432*	0.922*	0.861*	0.868*	0.486*		
TM7	-0.725*	0.801*	0.658*	0.845*	0.263	0.822*	
VOL	0.138	-0.592*	-0.511*	-0.425*	-0.522*	-0.676*	-0.492*
DBH	0.459*	-0.504*	-0.259	-0.343*	-0.199	-0.524*	-0.548*
HEIGHT	0.608*	-0.834*	-0.604*	-0.688*	-0.304*	-0.838*	-0.831*
DOMH	0.602*	-0.871*	-0.681*	-0.772*	-0.281	-0.893*	-0.878*
BA	0.073	-0.517*	-0.464*	-0.365*	-0.528*	-0.613*	-0.425*
TREE	-0.725*	0.416*	0.087	0.376*	-0.064	0.348*	0.565*

(*) Marked correlations are significant at $p < 0.05$

It was not possible to analyse the effect of age or number of trees separately over the general spectral responses because of the small number of observations. Apparently the "saturation effect" was responsible by the less expressive correlation between spectral responses and AGE but the analysis is complex because of the effect of the commercial management (thinnings) implemented in the first years of the plantation.

4.4.3 *Araucaria* forest - spectral responses and species composition - Experiment 3

The greater complexity of mixed forests compared to pure forests is reflected by variation among species in crown form, phenology, growth rate, longevity, and size (Wulder 1999). Complex, multi-age and multi-species forests such as *Araucaria* forest are especially difficult to assess with traditional remote sensing tools and techniques. However, an investigation about spectral behaviour under different species concentration might help in the discrimination of one particular species.

4.4.3.1 General comments on the biome

An analysis about the environment of the natural forest in INF is presented in **Chapter 2**, but in this Chapter some aspects were picked up to analyse their contribution to the reflectance values over the different bands.

Plants within communities do not occur in equal proportions; certain species tend to dominate and often are used to name the communities. The *Araucaria* forest is not an exception. *Araucaria angustifolia* (commonly called paraná-pine), a conifer, is the dominant species, frequently representing between 15 and 40% of the forest trees. In the study area, the number of trees per hectare varies very much. As mentioned before, in the fieldwork, around 44 trees (10% of the number mentioned for plantations) of *Araucaria angustifolia* per hectare were found on average,

and 187 trees per hectare of the other deciduous species (DBH larger then 20 cm). The plots with more concentration of the species *A.angustifolia* presented almost 100 trees (DBH larger than 20 cm) per hectare, which means around 40% of the total number of trees. The plots with lower concentrations of the species presented between 6 and 15 trees per hectare (3% to 8% of the total number of trees).

In the field inventory, 80 different woody species (DBH > 20 cm) were found (**Appendix 2.3**). As a snapshot of the species variability and their abundance, some of the other species present (all deciduous species) in the INF, ordered by abundance were: *Nectandra lanceolata* Nees et Martius ex Nees (16.07 trees per hectare), *Capsicodendron dinisii* (Schwacke) Occhioni (11.33 trees per hectare), *Nectandra grandiflora* Nees (11.27 trees per hectare), *Prunus brasiliensis* (Chamisso & Schlechdt) D.Dietr. (8.47 trees per hectare), *Ocotea puberula* Nees et Martius Nees (7.80 trees per hectare), *Matayba elaeagnoides* Radl. (6.33 trees per hectare), *Rapanea* (Myrsine) *umbellata* (Mart.ex A.DC.) (5.47 trees per hectare), *Lithraea brasiliensis* L.March (7.73 trees per hectare), *Nectandra megapotamica* Mez (5.20 trees per hectare), *Ocotea odorifera* (Vellozo) Rohwer (4.40 trees per hectare), *Ocotea porosa* (Nees ex Martius) L. Barroso (4.33 trees per hectare), *Cedrela fissilis* Vellozo (3.53 trees per hectare), *Ilex theezans* Martius (3.40 trees per hectare) and *Cinnamomum sellowianum* (Nees) Kosterm.(2.80 trees per hectare).

According to Campbell (1996), communities tend to be stratified vertically, with some species forming an upper canopy and others a lower stratum. Vegetation of smaller-scale like shrubs and mosses forms other layers nearer to the ground. In the *Araucaria* forest, the upper-most canopy layer is normally composed only by the species *Araucaria angustifolia.* It is a straight tree with almost cylindrical bole and commonly grows to heights of 30-50m (sometimes 10 to 25% higher than the second layer of species). In the pristine areas, it was very common to find trees with 2m of diameter at breast height (DBH).

4.4.3.2 Field work and imagery information

As already mentioned, the area with native forests in INF has 1508.09 hectares distributed in patches of different shapes and sizes. At the same time as the other data sets (pine and paraná-pine plantations) a third data set composed by information from natural stands was collected. The methodology adopted is described in IBAMA (1990). The systematic approach was selected as the sampling process, with sample units distributed proportionally to the size of the different patches that compose the area. The plots were distributed in strips (east-west orientation) approximately 500 m apart. The dimension of the sample units was 15 m-by-100 m and they have located at intervals of 100 m. Using this procedure, some baselines received several units and others just one or two. The sample units were installed as permanent plots and 93 plots were installed. In each plot, the following information was obtained: a) identification of the species; b) all individuals with GBH (girth at breast height) equal to or greater than 20 cm of diameter; c) timber height (bole length); d) total height of the paraná-pine trees; e) evaluation of the stem quality; f)

phytosociological position; g) characteristics of the understory, access and slope. Volume was calculated by specific equations developed for the paraná-pine (conifer) and for all the other species (deciduous). In this experiment, the 1987 Landsat TM imagery was used, using bands 1 to 5 and 7 (imagery acquisition and pre-processing methodology are detailed in Chapter 3).

4.4.3.3 Location of the sample units

The position of each strip and plot was determined and recorded in 1998 by using the "Global Positional System"(GPS) from GARMIN (precision around 1.40 m). The geodesic co-ordinates were collected in the central part of the plot whenever it was possible. Constraints such as accessibility of the stand and weak GPS signal under canopy were frequent. However, locality the plots were facilitated by the systematic approach adopted in the sampling method, by the presence of field boundaries demarcation and by the maps, aerial photographs and the sampling layout. The spectral response of each of those recorded points was recorded, for bands 1 to 5 and 7.

4.4.3.4 The K/d ratio

There is a special interest in the analysis of the possible correlation between DBH (and, consequently basal area) and spectral responses. This is because of its straight correlation with crown parameters like crown diameter and crown area (sometimes called crown closure and abbreviated to K). In field inventories, the volume is usually obtained by mensuration of appropriate variables in trees (individuals). The most important variables in the volume calculation are DBH (and, consequently, basal area) and the tree height. A similar procedure can be adopted when using remote sensing, mainly when high spatial resolution digital imagery is available. However, in such situations the direct mensuration of DBH is impossible (Aldred *et al.*1972). Consequently, determining volume depends on the correlation among bole diameter (usually DBH) and the dimensions of the crown (K), a correlation that can be established for one or for a group of species (Aldred 1976).

According to Philip (1994) many researchers have investigated the ratio between crown width (K) and DBH (d). Dawkins (1963) studied 18 tropical forest species and deduced that for any given species the ratio was remarkably stable regardless of site, age or stocking. Actually, there is a tendency of agreement that the crown diameter and DBH are closely correlated for a given species and that one may be predicted from the other relatively precisely (Philip 1994). According to Dawkins (1963), a practical interpretation of the relationship between the crown diameter and bole diameter is the straight line, with the exception of the extremes of the distribution (big or small trees). In many cases, the relationship is slightly curvilinear, approaching, a straight line. The form of the regression varies with the species and spatial location in the sampled population. Freely growing trees and the edge effect also must be considered in the analysis. However, as a preliminary rule valid for North America, the DBH in inches (original units) should be considered as the same as 3/4 of the crown diameter in feet, for average diameter classes of conifers in the USA. Specifically for pines of the south on States of the USA, Minor (1951) suggested that the DBH in

inches should be approximately equal to the half of the crown diameter in feet (original units) , plus 5 (as a correction factor). Avery (1958) built a volume table for conifers and deciduous species of the south of the USA, based on the premise that the crowns of the studied trees tended to occupy areas closed to the square of their average diameters. Nowadays, with the increased level of details, it seems that the spatial resolutions of image available from multispectral sensors, scanned aerial photographs and upcoming earth observation satellites imply a change in the digital analysis paradigm for forestry. Special attention is given to tree crown delineation (Gougeon 1999; Pollock 1999) and its relation with structural data (Wulder 1999; Magnussen *et al.* 1999).

In Brazil, Volkart (1969) and Silva (1978) investigated the relationship among the two variables, in natural stands of *A.angustifolia*. The comparison of the results of the two studies can lead to the conclusion that such a relationship, *a priori*, would be close to a straight line, agreeing with Dawkins (1963). Oliveira (1980), working with data from natural stands of *Araucaria angustifolia* also found that that relationship could be expressed by a first-degree polynomial. **Table 4.8** presents statistical information about the analysed database.

Table 4.8 Descriptive statistics of field data collected from patches by sampling

Biome: Araucaria forest - number of plots = 93				
Variable	Mean	St.Dev.	Max	Min
MD_Aa	39.03	14.16	74.15	0.0000
Nt/pl	35.22	11.09	60.00	4.0000
Nt_Aa	6.60	5.35	28.00	0.0000
V/pl	19.57	8.59	55.16	2.2400
V_Aa	10.68	9.23	51.92	0.0000
BA/pl	2.98	0.92	5.11	0.4500
BA_Aa	0.93	0.73	3.88	0.0000
CD_p	50.95	39.95	200.39	0.0000
CD_a	54.40	42.90	231.28	0.0000
CA_p	4.09	3.61	21.74	0.0000
CA_a	3.67	2.91	15.09	0.0000

Where: MD_Aa=diameter corresponding to the average tree of *Araucaria angustifolia*; Nt/pl=total number of trees in the plot; Nt_Aa=number of trees of *Araucaria angustifolia*; V/pl=total volume per plot; V_Aa=volume of *Araucaria angustifolia*; BA/pl=total basal area in the plot; BA_Aa=basal area of *Araucaria angustifolia*; CD_p=crown diameter of *Araucaria angustifolia* obtained by regression (sum of CD's of all trees of *Araucaria angustifolia* in the plot); CD_a=crown diameter corresponding to the tree of *Araucaria angustifolia* with mean DBH; CA_p=Crown area (percentage) obtained by regression; CA_a=Crown area (percentage) diameter corresponding to the tree of *Araucaria angustifolia* with mean DBH.

4.4.3.5 Ground data – relationships between forest plot parameters

Inventory procedures can provide detailed data on the location and extent of forest resources, such as species composition, age, height, tolerance level, density, and crown closure (Wulder, 1999). In the study area, however, information about crown closure was not directly collected, but as mentioned in the **section 4.4.3.4**, there is a strong relationship between DBH and crown dimension in *Araucaria angustifolia* and it was used in the process. In this study, although information about individual variables like DBH and height was available, they were only used to calculate variables related to the area units, like Basal Area and Volume. These variables are being considered as a

surrogate for biomass (total of absolute amount of vegetation present in one area), a biophysical parameter, usually used to simplify the measurement of forest structure into a single measure.

4.4.3.6 Crown competition - a density indicator

Crown density is defined as the compactness of tree crowns per unit of area and, generally, it is expressed as a percentage of ground area covered. It reflects the relationship between numbers of trees (or volumetric values) and the area unit, usually the hectare. The density is, therefore, used to aid the determination of the maximum and minimum space that the medium tree of the stand could compose (Krajicek *et al.*1961). According to Silva (1978), the crown surfaces and lengths, associated to the environmental conditions might help explaining production and wood increment. If the average crown diameter is known and assuming that tree crowns tend to be to circular, the area covered by the crown projections can be estimated. The concept of this variable is related to the idea of basal area, which is the cross-sectional area of all stems of a species or all stems in a stand measured at breast height and expressed per unit of land area. CA is usually presented in percentage.

$$CA = \left[\frac{\pi (CD)^2}{4} \right] / 100 \qquad \therefore \qquad CA = 0.0079 (CD)^2$$

where:
CA = crown area
CD = crown diameter

Initially, the relationship among CD/DBH is established. Oliveira (1980) described a study carried out in order to: a) verify the possibility of measuring *Araucaria angustifolia* trees in natural stands by using aerial photographs; and, b) study and select regression models conceived to correlate dendrometric variables easily observed in aerial photographs and field parameters. Crown dimensions were measured in the field and in the aerial photographs. The association level (correlation) between DBH and CD was 0.8530 and it was found that:

$$CD = -0.089347838 + (0.204941414 \times DBH)$$

This equation was obtained with determination coefficient (R^2) of 0.73 and standard error of estimate of approximately 18%. It was also found that:

$$CA = 0.0000632 - (0.00028914 x DBH) + (0.0003318 \times DBH^2)$$

In the present study these equations were used to estimate CD (unknown) as a function of DBH (known). They were used also to create a variable (CA) that could express the percentage of area occupied by crown projections using DBH and the number of trees in the defined area. As the DBH of *Araucaria angustifolia* trees were known, CA was calculated for each plot.

The investigation concentrated on the association between the following variables: a) number of trees (total and the specific for the species *Araucaria angustifolia*); b) volume (using both approaches); c) basal area, all representing field variables; d) TM1 to TM5 and TM7 represented the spectral responses (reflectance) and their spectral values were collected in 3-by 3 window design (TM1_a to TM7_a). The strategy was designed because of the size of the plots in the fieldwork. Because of the sensitivity of TM7 to the differences in the plots, a second approach was designed to capture reflectance in this band. It consisted in considering only the spectral value collected in the central point of the window, ignoring in this way the natural variability between pixels (TM7_p). This approach assumes that the co-ordinates captured by GPS were exact, which it is almost impossible. As mentioned before (**section 4.4.1.4**), an error (RMS) of 0.852 pixels occurred in the process of rectification, which means that the exact location of the centre should be in one of the neighbour pixels. However, in this case, differently from the pine approach, the plots were three pixels larger in size and, so the location of the centre might be inside the plot; and e) the last set was composed by the transformed variables CD and CA. Two different entries of DBH were tested to calculate CD's and CA's. The first one, CD_a (and CA_a) were obtained by regression using the diameter corresponding to the tree of *Araucaria angustifolia* with the mean diameter as dependent variable. The second entries, CD_p (or CMA_p) were obtained by the sum of the CD's or CA's of all trees of *Araucaria angustifolia* and calculated by the equation mentioned above (**Table 4.9**).

Table 4.9 Matrix of correlation.

	Nt_pl	Nt_Aa	V/pl	V_Aa	BA/pl	BA_Aa	TM1_a	TM2_a	TM3_a	TM4_a	TM5_a	TM7_p	TM7_a
Nt_Aa	0.285												
V/pl	0.464*	0.705*											
V_Aa	0.248	0.861*	0.865*										
BA/pl	0.737*	0.348	0.799*	0.459*									
BA_Aa	0.271	0.862*	0.828*	0.967*	0.473*								
TM1_a	-0.137	-0.129	-0.231	-0.186	-0.215	-0.167							
TM2_a	-0.139	-0.088	-0.181	-0.150	-0.208	-0.148	0.361*						
TM3_a	-0.197	-0.009	-0.155	-0.093	-0.212	-0.092	0.483*	0.617*					
TM4_a	-0.210	-0.210	-0.239	-0.281	-0.182	-0.291	0.438*	0.472*	0.530*				
TM5_a	-0.303	-0.463*	-0.477*	-0.529*	-0.332	-0.533*	0.304	0.362*	0.428*	0.689*			
TM7_p	-0.308	-0.498*	-0.382*	-0.483*	-0.269	-0.520*	0.318	0.272	0.328	0.587*	0.669*		
TM7_a	-0.315	-0.492*	-0.457*	-0.486*	-0.321	-0.495*	0.332	0.367*	0.424*	0.599*	0.809*	0.762*	
mD-Aa	0.044	0.142	0.291	0.404*	0.170	0.412*	-0.181	-0.157	-0.152	-0.354	-0.355	-0.299	-0.250
CD_p	0.280	0.960*	0.782*	0.929*	0.411*	0.923*	-0.139	-0.118	-0.022	-0.240	-0.505*	-0.516*	-0.514*
CD_a	0.278	0.970*	0.797*	0.949*	0.418*	0.947*	-0.152	-0.109	-0.038	-0.255	-0.513*	-0.519*	-0.515*
CA_p	0.249	0.874*	0.840*	0.980*	0.463*	0.975*	-0.174	-0.117	-0.069	-0.286	-0.537*	-0.511*	-0.508*
CA_a	0.163	0.840*	0.723*	0.866*	0.342*	0.847*	-0.153	-0.048	-0.018	-0.235	-0.483*	-0.490*	-0.493*

(*) Marked correlations are significant at $p < 0.05$

Where: Nt_Aa = *Araucaria angustifolia* number of trees; V/pl = Total volume per plot; V_Aa = Volume of *Araucaria angustifolia*; BA/pl = Total basal area in the plot; BA_Aa = Basal area of *Araucaria angustifolia*; TM1_a to TM7_a = spectral values collected in 3-by 3 window and TM7_p = spectral value collected in the central point of the window; CD_p=crown diameter of *Araucaria angustifolia* obtained by regression (sum of CD's of all trees of *Araucaria angustifolia* in the plot); CD_a=crown diameter corresponding to the tree of *Araucaria angustifolia* with mean DBH; CA_p=Crown area (percentage) obtained by regression; CA_a=Crown area (percentage) diameter corresponding to the tree of *Araucaria angustifolia* with mean DBH.

Analysing the association level for band 5, the same tendency of association with volume (total or the species *Araucaria angustifolia*) and basal area (only for *Araucaria angustifolia*) was found as in

the experiments 1 and 2. Correlations of the same magnitude (*r* approximately 0.5) were obtained with band 5 and the transformed variables CD and CA.

For TM7, there was no evidence of differences among the reflectance captured by the 3-by-3 window (TM7_a) or the central point of located plot (TM7_p). It was noted that band 7 presented a different signature tendency in *Araucaria* forest when compared to pine or *A.angustifolia* plantation. In both of man-made forests the reflectance of TM7 was always nil (or close to nil, in the stands of *Araucaria angustifolia* with less density or trees with smaller dimension) and in the natural environment it was close to 0 only when the *A.angustifolia* presence in the plot was considerable. Because of this tendency, the precaution of checking the pixel value only in the central point of the plot was adopted, but the results of the two approaches were very similar. As a conclusion, it seems that band 7 is sensitive to the presence of coniferous *Araucaria angustifolia* in the analysed area.

The transformed variables CD and CA performed better in terms of association level with TM5 and TM7 than the associated field variable DBH, but apparently did not improve substantially the results obtained by basal area of the species *Araucaria angustifolia*. CA, in the original formula derived by Oliveira (1980) incorporated the concept of percentage but did not differ substantially of the results obtained by the transformed variable CD.

4.4.3.7 Testing the equality of population means

The natural forest in the INF is not concentrated in one extensive area. Neighbouring patches of natural forest, plantations were established since the 1940's in stands, with area varying from 1 to 50 hectares, composed of different species in monocultures of different ages. Nowadays there are some small areas with *Araucaria* forest among the infrastructure and the plantations and there are also three clearly identifiable major patches (**Figure 4.11**), forming a patchwork of different forest cover types. In the inventory procedure, the sample units were installed proportionally to the area of these patches. Because they are not continuous, but separated by other types of land cover, a question remained about the eventual statistical differences between the patches. Although systematic sampling does not accomplish the independence normally required for statistical analysis, if the designer is convinced that the systematic selection has resulted in an essentially random order of sample selection, then the procedures for random sample may be applied (Philip, 1994). Four groups of patches were considered in the investigation. The three mentioned patches and one forth group formed by small and close patches. The question was: are the differences in the sample means different enough to conclude that the population of the species *Araucaria angustifolia* means are different? One-way analysis of variance (ANOVA) was performed with the objective of testing the equality of population means and the variables basal area (BA_Aa) and crown diameter (CD_p) were chosen to be investigated.



Figure 4.11 *Araucaria* forest patches in the INF

4.4.3.7.1 Analysis of variance for basal area, crown diameter and spectral responses

As mentioned before, oneway analysis of variance was used to analyse the population basal area (BA/plot) with all trees including *Araucaria angustifolia*. The objective was verifying if there were statistically significant differences not only in basal area, but also crown diameter and the reflectance of band 5 in the four analysed patches. For basal area, the null hypothesis could not be rejected or, more specifically, there was insufficient evidence to say that the patches differ when that variable was analysed (accepting p-value > 0.10). The details of the analysis are presented in the **Appendix 4.2**. When analysing only the *Araucaria angustifolia* population (BA_a), instead, the results were different. There was highly significant evidence that the basal area means for that species performed differently in the four patches. This may be explained by the fact that in some plots, *A.angustifolia* was dominant in frequency and the species is definitely dominant in size. There was evidences that the four patches have different concentrations of *A.angustifolia* (represented by basal area). The next step was to perform a statistical approach to detect which patch presented differences in basal area.

This analysis was carried of using Tukey's multiple comparison procedure. It was adopted to check the evidence of differences in basal area of the four patches; the resultant tables are also presented in **Appendix 4.2**. By analysing those tables, it was verified that there was a significant difference between $\mu 1$ and $\mu 2$, implying that the basal area (BA_a) contents of the species

Araucaria angustifolia is statistically different in patches 1 and 2 (the summary of the analysis is in **Table 4.10**).

The same procedure was adopted to analyse the variable crown diameter obtained by regression (CD_p) (**Appendix 4.2**) performance. Patches 1 and 2 and patches1 and 4 were found to present significant (error rate was 5%) differences from each other (summary of the analysis is presented in **Table 4.10**). In fact, patch 1 seems to present fewer trees of paraná-pine (and with smaller dimension) than patches 2 and 4 specifically. Patch 3 seems to present an intermediate situation. Finally, the same procedure was adopted to investigate if the spectral responses (specifically band 5-TM5_a and band 7-TM7_a) could also be recognised as different in the four patches. Significant differences (summary of the analysis is in **Table 4.10**) between patch 1 and patch 4 were found when band 5 was the analysed variable (complete analysis is showed in **Appendix 4.2**). Summarising the results for ground variables (basal area and crown diameter) and spectral variables (reflectance of bands 5 and 7), it was concluded that there is a difference in basal area of paraná-pine between the sampled areas and that this difference can be captured by TM5. To clarify the adopted procedure, the summary of the Tukey's pairwise comparisons for basal area, crown diameter and reflectance of band 5 is showed in **Table 4.10**.

Table 4.10 Summary of Tukey's pairwise comparisons for BA_A, CD_P and TM5

Tukey's pairwise comparisons for BA_a			
	P A T C H E S		
	1	2	3
P 2	-0.8770		
A	-0.0204		
T			
C 3	-0.9570	-0.5273	
H	0.5433	1.0111	
E			
S 4	-1.2646	-0.8380	-1.3205
	0.0270	0.4977	0.4965
Tukey's pairwise comparisons for CD_P			
	P A T C H E S		
	1	2	3
P 2	-46.6		
A	-1.2		
T			
C 3	-51.6	-28.7	
H	27.7	52.6	
E			
S 4	-81.9	-59.2	-83.9
	-13.6	11.5	12.3
Tukey's pairwise comparisons for TM5			
	P A T C H E S		
	1	2	3
P 2	-0.391		
A	3.201		
T			
C 3	-1.435	-2.920	
H	4.856	3.531	
E			
S 4	0.235	-1.263	-2.577
	5.651	4.338	5.042

As mentioned before, the approach was considered to refine the analysis of variance. This further multiple comparison test was built to verify which patches actually presented differences when those variables were analysed. Numbered columns and rows identify the analysed patches (1,2,3 or 4). The fact that the origin (zero) is inside the interval of confidence indicates that there is no significant difference between the compared patches (1 and 2, 1 and 3 or 2 and 3, for example). The evidence that there are differences between patches, for a specific variable is highlighted in bold. The complete analysis of variance and Tukey's pairwise comparisons for BA_plot and BA_Aa, CD_p, TM5 and TM7 is presented in **Appendix 4.2**.

After analysis of the individuals plots, it become clear how heterogeneous the natural stands are, even within the patches. It is very common to have more than 10 trees of paraná-pine in one plot and just one or two, or even none in the next one, at a minimal distance of only 100 m. Patch 1 has a smaller *Araucaria angustifolia* basal area, on average. The plots with smaller basal area or crown percentages of *Araucaria angustifolia* are located in patch 1. However, patch 1 contains some areas when the presence of the species of interest is massive. Even with such a difficult scenario, in a few cases it was possible to detect areas with more than three neighbouring plots that have a high concentration of the species, creating a different niche. This spatial analysis is presented in **Chapter 6**, when the results of the Linear Mixture Modelling are superimposed on the plot distribution.

4.5 Typical spectral reflectance curves in the study area

In spite of all these investigations, it is still hard to present a typical curve for each species or biome due to the natural heterogeneity, even in planted areas. In the case of the INF, age and stand density or patch species composition varies very much. A typical curve should not contain all this variability. However, it is possible to characterise the differences between the different features by the mean curves for each component. The results are summarised in **Table 4.11**.

The response patterns of the pine species and paraná-pine plantation were obtained by averaging the reflectance values presented in the experiments 1 and 2. The curves for *Araucaria* forest were obtained by subdividing the data set in two parts: a) plots with less than 3 trees of *Araucaria angustifolia* (20 trees per hectare) were considered as secondary forests (A.for_s); b) plots with at least 10 trees of *Araucaria angustifolia* (more than 66 trees per hectare) were considered primary (A.for_p). **Figures 4.12** and **4.13** present the results graphically. A mask was drawn around the identified non-vegetated areas (soil and water). They were selected by using ancillary data such as the map of the area and aerial photographs. Twenty points in those different areas of the images were taken randomly and considered enough to capture the main variance between different backgrounds; the adopted value was based in the mean of those samples.

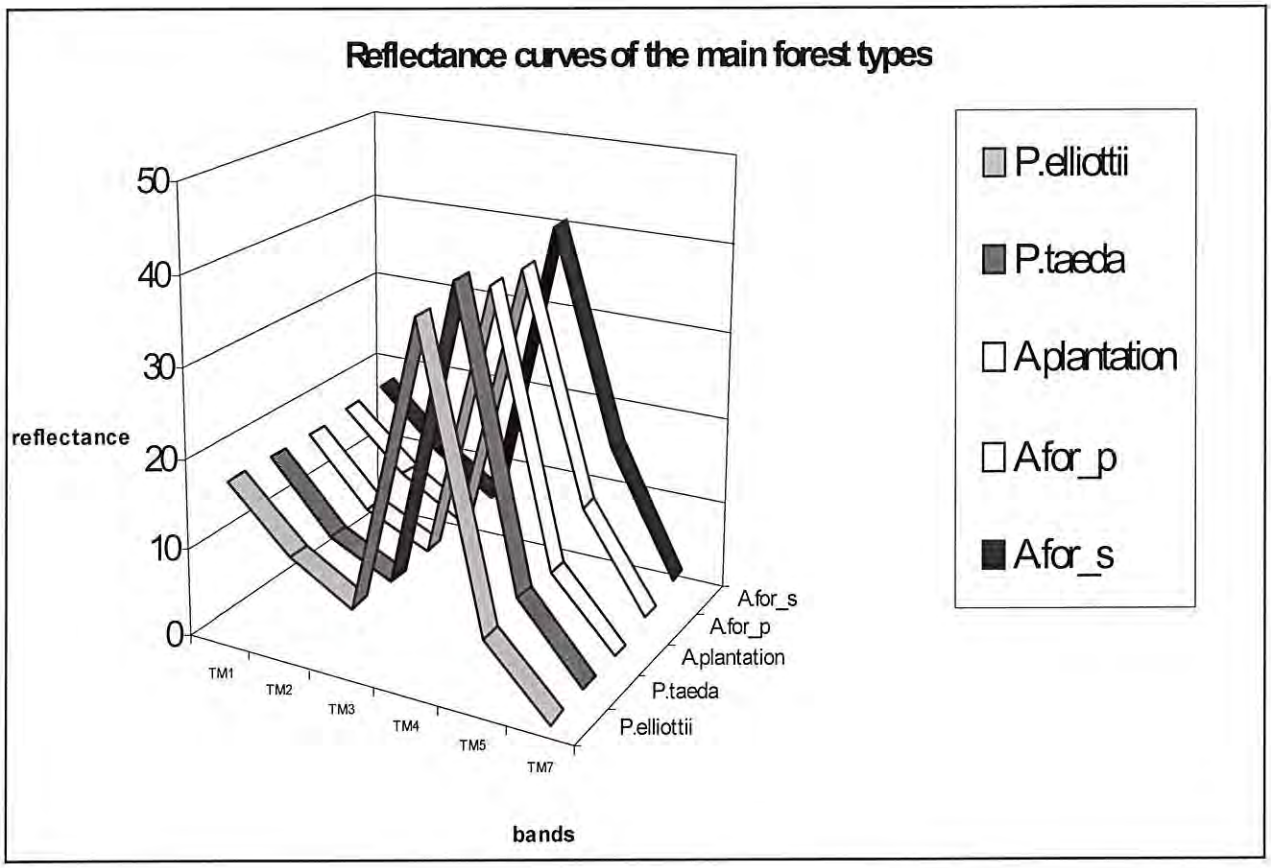


Figure 4.12 Typical reflectance curves of the main forest types in the study area

Where: P.elliottii = Pinus elliottii; P.taeda = Pinus taeda; A.plantation = plantations of *Araucaria angustifolia*; A.for_p = *Araucaria* forest plots with at least 10 trees of *Araucaria angustifolia*; A.for_s = *Araucaria* forest plots with less than 3 trees of *Araucaria angustifolia.*

Table 4.11 Typical spectral values (reflectance * 255) for the analysed forest types

Features	TM1	TM2	TM3	TM4	TM5	TM7
P.elliottii	17.6	10.7	6.7	39.2	7.1	0
P.taeda	17.5	9.9	6.5	40.8	8.2	0
A.plantation	17.5	10.2	6.9	38.1	7.9	0
A.for_p	18.1	11.9	6.9	40.0	12.9	1.1
A.for_s	18.0	11.4	7.2	43.0	16.2	2.0
Cunninghamia/Cupressus	18.9	11.2	7.8	42.1	8.5	0
Eucalyptus spp.	17.9	10.9	7.7	55.0	24.0	4.0
Sub-tropical forest	18.3	11.3	7.2	44.4	19.0	8.9
Riparian areas	18.9	12.4	9.1	36.9	15.6	10.6
Non-forest	19.8	15.2	14.3	39.3	35.3	16.1
Swamps	19.5	14.0	13.3	39.4	36.6	18.7
Flood-plain areas	19.7	14.0	10.9	40.4	23.2	15.3
Water	19.9	15.0	13.0	1.0	0	0
Soil	26.6	28.0	38.2	50.0	72.2	55.0

Where: A.plantation= *Araucaria angustifolia* plantation; A.for_p = plots with at least 10 trees of *Araucaria angustifolia* (more than 66 trees per hectare); A.for_s = plots with less than 3 trees of *Araucaria angustifolia* (20 trees per hectare); Cunninghamia/Cupressus = stands of *Cunninghamia lanceolata* or *Cupressus lusitanica*;

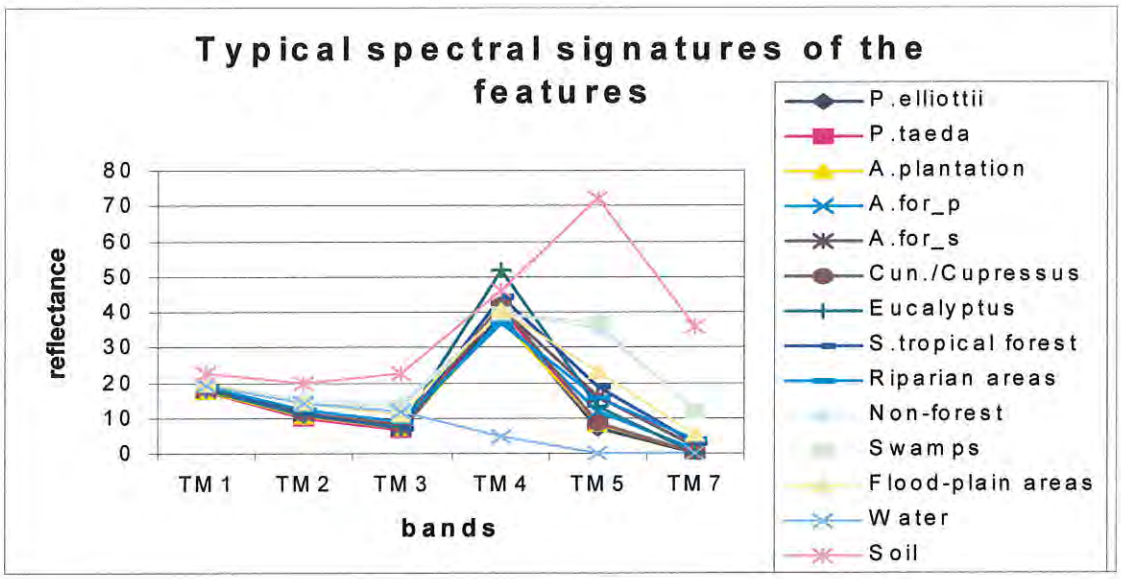


Figure 4.13 Typical reflectance curves of vegetated and non-vegetated areas (14 classes)

Where: Reflectance = reflectance * 255; Cun./Cupressus = stands of *Cunninghamia lanceolata* or *Cupressus lusitanica*; E.viminalis = stands of *Eucalyptus viminalis*; S.tropical forest = areas with sub-tropical forest where the species *Araucaria angustifolia* is not present; Riparian areas: vegetation located on the bank of the Antas River and Jacu Stream; Non-forest = agriculture or pasture areas; Swamps = non vegetated (or low vegetated) areas always covered by a water table; flood-plain areas = areas that are subject to overflow flooding during periods; water = water bodies like lakes and reservoirs; soil = non-vegetated areas captured usually in the urban zone.

4.6 Conclusions

All the analysis were performed with the objective of characterising remotely sensed differences between different forest land covers (pine and paraná-pine plantations versus *Araucaria* forest) and verifying possible differences in the natural areas that should be captured by Landsat TM. The conclusions can be summarised as: a) spectral responses of *P.elliottii* were similar to that reported by Sousa (1997) in plantations located in São Paulo State, Brazil. The species presents brighter values than *P.taeda*; b) other coniferous plantations (*Cunninghamia lanceolata* and *Cupressus lusitanica*) unfortunately are small stands and their areas probably are not representative, but it was found that their curves have a different aspect, with higher values in the region of band 5.

Experiment 1 - Based upon a study of pine species stands (*P.elliottii,* slash pine, and *P.taeda,* loblolly pine) of the Irati National Forest, Brazil, an evaluation of the contribution of remote sensing to forestry monitoring and management using Landsat TM was made. This was achieved by correlating canopy structural variables with the individual TM wavebands. The results illustrated that, at that spatial resolution, scenes contain enough information to discriminate effectively between areas with high and low volume and basal area. TM5 was found to be highly significantly correlated with both selected ground variables (Basal Area and Volume). TM4 was found to be correlated with DBH and height variables, but the relationships were not strong. The other visible

bands, TM1 (Wavelength = 0.45-0.52 μm), TM2 (Wavelength = 0.52-0.60 μm) and TM3 (Wavelength = 0.63-0.69 μm) were only weakly correlated with all the field variables.

Experiment 2 – In the plantations, *A.angustifolia* presented a significant correlation between field and spectral variables in all bands. This differs from the results obtained with pine plantations data set, when only band 5 had a significant correlation with field factors. TM5 was again the band with the highest correlation against mean diameter and basal area, but the other bands from the visible or infrared region also presented high association levels. Some variation among *A.angustifolia* and pine plantation was observed for TM7. Reflectance equal to or higher than 1 was more frequently found in the *A.angustifolia* plantation data set than in the pine data set. Basal area presented a significant inverse relationship with all bands (but TM1), with higher values in TM5, as found for pine.

Experiment 3 - An experiment was set up in order to analyse patterns of spectral responses from *Araucaria angustifolia* in a complex environment as the natural stands of *Araucaria* forest. TM5 and TM7 were the TM bands more correlated with field variables like volume and basal area. The transformed variables crown diameter and crown area presented almost the same level of association as basal area. Statistical analysis detected differences in basal area, crown diameter and spectral response (band 5) between the four patches presented in the area. By pairwise comparison, it was concluded that there is significant differences in basal area for *A.angustifolia* between the sampled areas (patches 1 and 2) and that this difference was captured by TM5 (patches 1 and 4 were significant different).

Chapter 5

THE CONTRIBUTION OF IMAGE TRANSFORMATIONS: DIMENSIONALITY REDUCTION OF THE FEATURE SPACE

5.1 Introduction

The analysis of vegetation patterns is a key to natural resource assessment and monitoring. This chapter analyses how different Image transformations can improve the process of forest use discrimination in the study area, with an emphasis on distinguishing *A.angustifolia*. Image transformations (usually called image transforms) involve the manipulation of multiple bands of data, whether from a single multispectral image or from multitemporal image sets (CCRS 1999). The term image transform is classic in remote sensing (RS). Image transforms are pixel-to-pixel operations which use different images through a defined rule, therefore generating a new band that combines the information of the original bands (Eastman 1997; Mather 1999). The new image is supposed to highlight particular features or properties of interest in a better format than the original input images.

This chapter deals with image transforms in a multispectral environment (Landsat TM5). **Chapter 7** deals with multitemporal sets (1987 and 1997), where vegetation indices (NDVI) and basic image transformations (as image differencing and ratio) were calculated, due to their applicability in the change detection process.

5.2. Vegetation Indices (VIs) approach

According to Huete *et al.* (1994), vegetation indices (VIs) are dimensionless radiometric measures that serve as indicators of relative growth and/or vigour of green vegetation. According to Campbell (1996), usually a vegetation index is formed from combinations of several spectral values that are added, divided, or multiplied in a manner to catch information related to the amount or vigour of vegetation within a pixel. VIs form part of the tools that are used to organise and synthesise information from multispectral bands into information that can be associated with the physical characteristics of ground classes (Bosdogianni 1996). Any spectral band can be involved in the process and this is why there are so many VIs available in the literature. Traditionally, however, the relation between red:infrared bands have been hugely explored mainly because of the inverse relationship between vegetation brightness in the red and infrared region (Campbell 1996) . Therefore, VIs could be classified broadly in two basic groups: the *slope-based* and the *distance-*

based approaches. According to Eastman (1997) the slope-based VIs are arithmetic combinations that focus on the contrast between the spectral response patterns of vegetation in the red and near-infrared portions of the electromagnetic spectrum. The main objective of distance-based VIs is to cancel the effect of soil brightness in cases where vegetation is sparse and pixels contain a mixture of green vegetation and soil background (Eastman 1997). Apparently, distance-based VIs have more specific applications in arid and semi-arid environments.

The number of different VIs available presently is so considerable that it is beyond the scope to deal with in this study. The aim of this work was not a comprehensive comparison of the many different VIs. Instead, the strategy adopted in the present study was to perform preliminary investigations, by selecting and calculating a number of slope-based and distance-based vegetation indices, as quoted in the literature or their slightly modified versions. Most of these VIs involved bands 3,4,5 and 7; they were compared visually on their capability of contributing to the general cover-type discrimination or, more specifically, discrimination between the different forest types present at the INF. However, the resultant images were just comparable to NDVI or RATIO, which were the most mentioned in the literature. Those indices were used in the training process and were useful as indicators in the process of determining which bands and relations between bands tended to be more useful for the purposes of this work. However, they are not discussed in this work, because of their poor contribution to the objectives of the thesis. The distance-based VIs in particular did not improve the representation of the spectral knowledge. The patches of *Araucaria* forest, although threatened by human action resulting in a reduction in size and vegetation density, are still considered densely vegetated and distance – based VIs are recognised to be successful in areas with sparse vegetation (Thiam 1977). The list of the previously calculated VIs is presented in the **Appendix 5.1**.

5.2.1 The RATIO

The index was proposed by J.Rouse *et al.* in 1974 (Eastman 1997) to separate green vegetation from soil background and is produced by dividing the reflectance values contained in the near-infrared band (TM4, wavelength 0.76-0.90 μm, for Landsat TM) by those contained in the red band (TM3, wavelength 0.63-0.69 μm), i.e.:

$$\text{Ratio} = \frac{\text{Near Infrared band}}{\text{Red band}}$$

According to Thiam (1997), the choice of these wavebands is based upon their inverse relationship in measuring photosynthetically active vegetation. Non-vegetated areas (e.g. soil, water) have high values in the red region and modest values in the infrared regions. On the contrary, healthy vegetation displays low values in the red and high values in the infrared region. Consequently, the index is expected to serve well for biomass estimation. It is also expected to experience a reduction

of undesirable effects like the ones provoked by illumination and caused by variations in topography (Mather 1999).

In the study area, as mentioned in **Chapter 4**, TM3 was not particularly useful in discrimination between the different forest types. Its standard deviation in all the analysed areas was, as expected, smaller than for TM4, showing that the variability of pixel values is greater in the infrared region. One reason for this failure to aid discrimination between coniferous and deciduous species is that TM3 is in the visible region of the spectrum, where this separateness is uncommon. **Figure 5.1** presents the result of the transformation.

Areas with forest vegetation were discriminated from the non-vegetated areas (contrast between the red and infrared bands for vegetated pixels). The water content was also enhanced (dark colour) but the expected minimisation of the effects of variable illumination as a result of topography, did not occur. Although the index was constructed as a ratio, topography was still a factor confusing in the definition of different vegetation classes. Some young stands of *Pinus* spp. presented a brighter tone, but no significant discrimination was noticed between coniferous and deciduous stands or between natural and planted areas. The index was kept in the analysis mainly because the further discussion used a variation of it, the ratio between bands 4 and 7 (labelled as ratio47). As mentioned in **Chapter 4**, band 7 plays an important contribution when the native stands of *Araucaria* forest are being analysed. The transformation (TM4/TM7) is presented in **section 5.2.1.2.2**, when the low correlation between bands is analysed and the results are compared with the results of selective PCA.

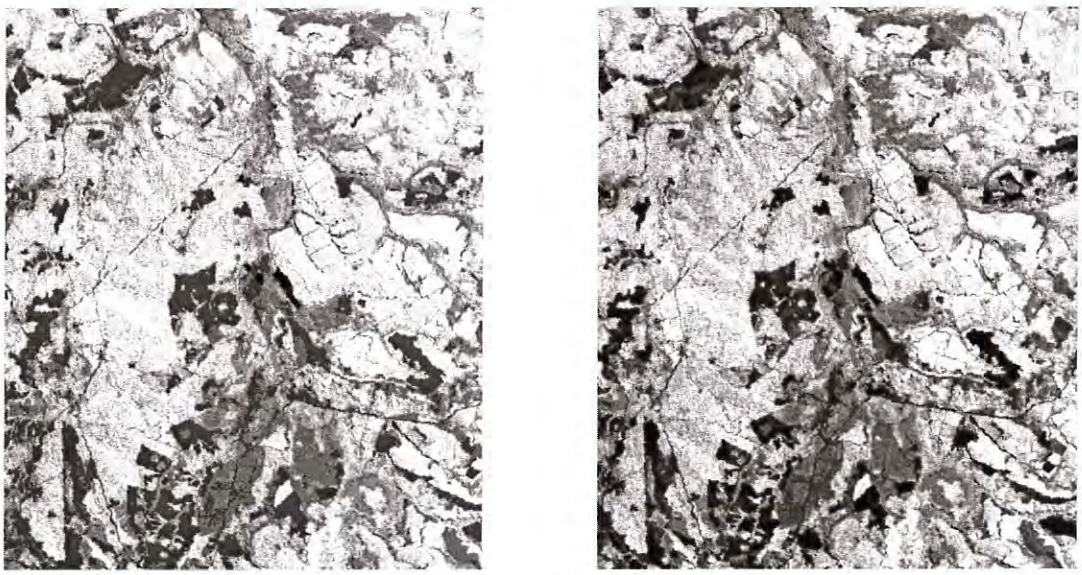


Figure 5.1 (a) RATIO VI image and (b) NDVI image focusing on the area of the INF. Both images have been subjected to a histogram equalisation contrast stretch.

5.2.2 The NDVI

The most used but not necessarily successful VI is still the q research since the 1970's (Huete *et al.* 1994). It was also introduced by J.Rouse *et al.* in 1974 (Eastman 1997), with the objective of separating green vegetation from its background soil brightness. It is expressed as the difference between the near infrared and red bands normalised by the sum of those bands, i.e.:

$$\text{NDVI} = \frac{\text{Near Infrared band} - \text{Red band}}{\text{Near infrared band} + \text{Red band}}$$

It is preferred to the RATIO index (mentioned above) because of the ratio value is not affected by the absolute pixel values in the near infrared and red bands (Mather 1999). The approach is based on the observation that high values reveal pixels dominated by high proportions of green biomass. Thiam (1997) mentioned studies calculating that the normalisation reduces the effect of sensor calibration degradation by approximately 6% of the overall index value. The author also mentioned its capability of eliminate topographic effects and variations in the Sun illumination angle and atmospheric contaminations that result from aerosol scattering and absorption.

The results of the present study (**Figure 5.1**) indicate that it was possible actually to separate the vegetated areas from the non-vegetated areas. However, the expected minimisation of the topographic effects was not achieved. Consequently, considering the purposes of the present project, the results are no more useful than those provided by the RATIO image, even if by using NDVI the proportions of vegetation are more discernible. As mentioned by Mather (1999), apparently the difference between RATIO and NDVI is not great. However, in the case of NDVI, sums and differences of bands are used rather than absolute values, making it more appropriate for use in comparisons over time, as is discussed in **Chapter 7**. Actually, according to Perry *et al.* (1984) RATIO and NDVI are functionally equivalent. The authors found that they have a relationship that allows going from one to the other. They indicate that dividing both the numerator and denominator of RATIO produces:

$$NDVI = \frac{RATIO - 1}{RATIO + 1}$$

The study area, as mentioned before is composed by four main forest types: *P.elliottii*, *Pinus taeda*, *Araucaria angustifolia* plantation, and *Araucaria* natural growth forest. Patches of *Araucaria* forest can be considered highly abundant in terms of number of trees (at least within the National Forest), when compared with patches in the adjacent areas and the NDVI discriminated only slightly the amount of forest vegetation in the whole area. Certainly, the position of the cover line is varying with the brightness of the vegetation cover, but the same amount of vegetation can appear differently accordingly to the soil type. According to Bosdogianni (1996) the cover line is best used

to detect change over time in an area with a particular soil background rather than to determine absolute cover classes across a range of soil types.

Hall *et al.* (1995) considered that NDVI is not a reliable predictor of biophysical parameters for the dominant coniferous species of the boreal forest. The authors concluded that NDVI is shown to be highly correlated with the fraction of sunlit canopy, which itself is relatively insensitive to both biomass and Leaf Area Index (LAI). Both transformations (RATIO and NDVI) are analysed with regard to their contribution to the change detection process, as discussed in **Chapter 7**.

5.3. The orthogonal transformations

This chapter considers derivations of vegetation indices determined through orthogonal transformation such the tasselled cap from Kauth-Thomas transformation and principal component analysis (PCA). The objective is to verify the usefulness of these indices in the process of enhancing and mapping the spectral differences or contrasts between the four main different forest types present in the INF: slash pine, loblolly pine, *Araucaria angustifolia* plantation, and *Araucaria* forest.

5.3.1 Principal Components Analysis (PCA)

In a multispectral remotely-sensed image, adjacent bands can be and generally are correlated (Mather 1999): if vegetated areas are analysed, because of spectral characteristics of vegetation, negative correlations between the near-infrared and visible red bands and positive correlations among the visible bands will occur. As the greenness (vigour) increases, the red reflectance diminishes and the near infrared reflectance increases. High correlation implies redundancy in the data and, as Mather (1999) mentioned, some information is being repeated. According to Lillesand and Kiefer (1994), images generated by digital data from various wavelength bands often appear similar and convey essentially the same information.

Principal Components Analysis (PCA) is also referred to as a Hotelling transform (Carr *et al.* 1999) or Karhunen-Loève transform (Crosta 1992; Mather 1999). According to Loughlin (1991), it is robust, fast and reliable and requires no atmospheric or radiometric corrections. The PCA is a statistical technique that transforms a multivariate data set consisting of inter-correlated variables into a data set consisting of variables that are uncorrelated linear combinations of the original variables (Ingebritsen *et al.* 1985). Hence, PCA can be used in remote sensing as a means of data compression (Eastman 1997), a transformation technique that reduces or removes the spectral redundancy (NETGIS 1998) or even as an image enhancement procedure (Picchiotti *et al.* 1997). However, it is important to reinforce that some information might be excluded by the procedure, mainly when the low-order principal components are discharged or omitted. The use of lower-order components depends on the aim of the project (Matter 1999) and, for some subjects like mineral

detection, it is likely that, due to its specific characteristics they are concentrated in the two or three low-order PCs (Crosta 1992).

According to Eastman (1997), given a set of image bands, PCA produces a new set of images, known as *components*. These principal components (PCs) or axes are orthogonal to each other and the new images have no mathematical correlation with one another (Chavez *et al.* 1989). In addition to being uncorrelated, they are ordered in terms of the amount of variance they explain from the original data set. Consequently, the procedure generates a new group of images and the largest amount of the total variance is mapped to the first component, with decreasing variance going to each of the following components. The sum of the variance in all components is equal to the total variance present in the scene (Chavez *et al.* 1989).

For a typical multispectral image band set, it is common to find that the first two or three components are able to explain most of the original variability in reflectance values. Later components tend to be dominated by noise effects. In theory, by rejecting these later components, the volume of data is reduced with eventually no appreciable loss of information. Chavez *et al.* (1989) mentioned operational advantages, because reduction in dimensionality means smaller volume of data and less computational demands. Ingebritsen *et al.* (1985) also considered that it is often difficult to appreciate visually all of the information available in the multispectral data. PCA may allow the available information to be displayed in two or three dimensions. Mather (1999) pondered that as the Landsat TM provides seven bands of data, the data compression property should be useful to create a conventional RGB false-colour composite without losing too much information. The procedure creates a convenient three image set but, as considered by Chavez *et al.* (1989) and Mather (1999), two problems can be encountered when using this approach: a) information of interest might possibly be mathematically mapped within one of the unused (discharged) components; and b) a colour composite can be difficult to interpret.

According to Crosta (1992), the number of spectral bands generated by the latest technology in RS tends to increase, and consequently, the technique of PCA is becoming increasingly relevant as a procedure which objectives dimensionality reduction. He also considered PCA as one of the most powerful image processing functions available for multispectral analysis. However it is, at the same time, one of the least understood techniques and, consequently, perhaps one of the most misused procedures. Some reasons for this, according to him might be: a) limitation of its potential through using repetitive procedures, derived from its historical connection to electric of engineering techniques; b) complexity of the mathematical conception of PCA; and c) difficulties in the analysis and extrapolation of results, which are site-specific and dependent on scene spectral characteristics.

The core of the procedure resides in the high correlation between bands. Therefore, the main function of PCA is to determine the extent of the correlation (or define the dimensionality of the

data set) and to remove redundancy, using mathematical procedures. According to Crosta (1992) the causes of this correlation are many, but the most important are: a) topographic shadowing; and b) relationship between superficial materials spectral reflectance and spectral bands.

Two bands are highly correlated when it is possible to deduce a pixel value in the second image (y) by knowing the value in the first one (x). The graph of this kind of relationship will be a straight line (AB) sloping upwards to the right (Mather 1999). When the correlation is perfect, r=1.0. However, even if the correlation is not perfect maybe a dominant direction of variability (AB) could be present yet. After this major axis is defined, a second minor axis (CD) might be drawn (where some variability might resides) at right angles to it; a plot using axes AB and CD should therefore represent more clearly the data structures (Mather 1999). **Figure 5.2** shows this idea and characterises the distinction between the number of variables (spectral bands) and intrinsic dimensionality of the data set. Note that the correlation remains high and positive in the second graph (b) in comparison to the theoretical perfect correlation depicted in **Figure 5.2**(a).

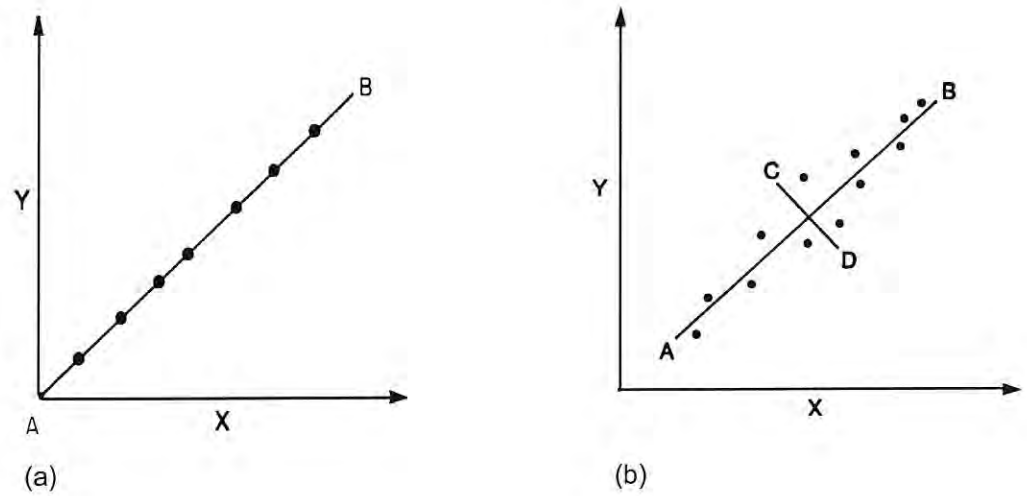


Figure 5.2 (a) Two perfectly correlated variables (x and y) and (b) two highly correlated variables (x and y). Source: Mather (1999)

For the example above, the number of variables is two (x and y) but the dimensionality is one (as shown in the scatter plot). The direction of axis AB is defined by the correlation between variables x and y. High positive correlation results in the scatter of points being restricted to an elliptical region in two-dimensional space (Mather 1999). The simplest parameter to be specified is the position of the centroid in the distribution. **Figure 5.3** deals with this idea, representing the location point of a hypothetical pixel of mean intensity; in this case, what is represented is the distribution gravity centre, not the geometric centre (Crosta 1992).

The size of distribution or the stretching measure of each band is given by the variance (which is proportional to the scatter of points in the direction parallel to the axis representing that variable). The correlation level between any two bands is specified by the covariance, which defines the shape of the ellipse enclosing the scatter of points. As summarised by Mather (1999), the mean vector and the variance-covariance matrix define the location and shape of the scatter of points in a *p*-dimensional space.

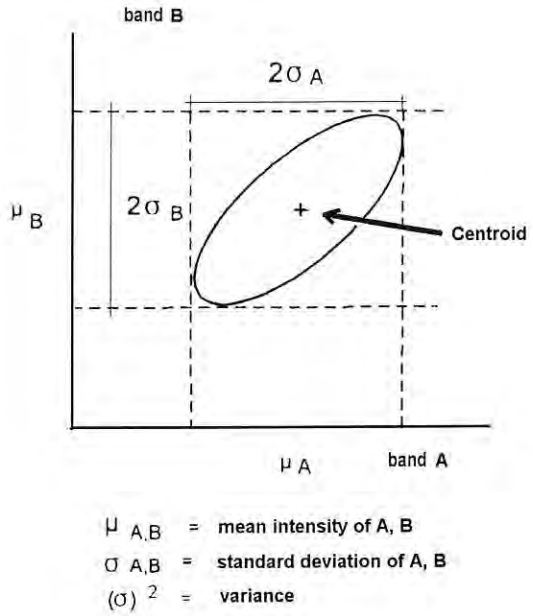


Figure 5.3 Basic statistical parameters of the space attribute
Source: Crosta (1992)

A variation of the traditional PCA is called *standardised* PCA and uses within- and between-band correlation coefficients to assemble a matrix that is eigen decomposed. According to Carr *et al.* (1999) the method is often known as factor analysis and aims to minimise the numerical influence of bands having a larger variance. They stated that, realistically, principal components and factor analysis are similar because (co)variance is directly proportional to the correlation coefficient. Mather (1999), however, described factor analysis as being somehow similar to the one proposed by Roger (1996) which is a noise-reduction method. Despite complication with definitions, some authors reported that for multitemporal change detection, standardised principal components are more accurate that non-standardised ones (Picchiotti 1997). However, Mather (1999) enlightened the theoretical assumption of each procedure by explaining the reasons for use of the correlation matrix (for standardised PCA) or the variance-covariance procedure. He clearly demonstrated that if the *p* variables making up the data set are measured on different and incompatible scales (raw images, for example), then the variances of these variables are not comparable. For him, in this case, the choice should be the use of a correlation matrix to standardise the measurement scale of each band or conversion of the image data to radiances followed by principal component analysis

of the variance-covariance matrix. As Fung *et al.* (1987) stated, standardised PCA is especially useful in multitemporal analysis because standardisation can minimise the differences due to atmospheric conditions or sun angles. However, as stressed by Mather (1999), if all the variables are measured in terms of the same units (such as milliwatts per square centimetre per steradian) then standardisation should not be applied because it removes the effects of changes in variability between the bands.

As mentioned above, the transformation derived from the covariance (or correlation) matrix among the bands generates a new group of images where each pixel value is a linear combination of the original values. The number of consequent principal components (in bands) and the original spectral bands is the same and they are organised in decreasing order of the grey level variance. The first PC has the largest variance (larger contrast) and the last, the smallest variance.

The covariances (or coefficient of correlation) are used to determine the eigenvalues, which represent the length of the principal axes of the ellipsoid whose shape is defined by the variance-covariance (or correlation) matrix (Crosta 1992) and are measured in units of variance (Mather 1999). Associated with each eingenvalue are eingenvectors. They are a set of co-ordinates (loadings) responsible for the direction of the associated principal axes. Both eigenvalues (lengths) and eigenvectors (direction) are important indicators in the interpretation of the results of PCA.

The technique has been successfully used in the detection of minerals by Loughlin (1991) in USA and Osako *et al.* (1996) in Brazil. Recently, PCA has been shown to have special application to environmental monitoring such as: a) monitoring forest defoliation (Muchoney *et al.* 1994); b) morphologic features of the Venice Lagoon (Picchiotti *et al.* 1997); c) detection of vegetation changes associated with extensive flooding in a forested ecosystem (Michener *et al.* 1997). Several of these studies have been reported using the technique in order to identify areas of change. This particular aspect of PCA is detailed in **Chapter 7**.

5.3.1.1 Standard PCA - Experiment 1

This study adopted the terminology suggested by Chaves *et al.* (1989) about different approaches involving PCA. According to the author, working with *standard* PCA means that all the available bands with same resolution (normally 1 to 5 and 7, for Landsat TM) are used as inputs to the analysis. *Selective* PCA, instead, means that only a subset of those bands are used as inputs. The word standard in this case has nothing in common with the term *standardised*, as mentioned above and which is related to the use of a correlation matrix instead of covariance matrix to perform PCA. The objective of the experiment was: a) to verify the usefulness of standard PCA to enhance and map the spectral differences or contrasts between the four main different forest types present in the INF (*Pinus elliottii* – slash pine, *Pinus taeda* – loblolly pine, *Araucaria angustifolia* plantation, and *Araucaria* forest). The importance of each feature for the region's economy and details about characteristics of each forest type were mentioned in **Chapter 4**.

Table 5.1 shows the correlation matrix for the six Landsat TM5 bands. As the second experiment is dealing with selective PCA, the correlation matrix can also be useful to identify the combination of bands to be used as input.

Table 5.1 Correlation Matrix for the six reflective bands (1987 set)

Bands	1	2	3	4	5	7	Mean	Std dev.
1	1.00	0.90	0.89	0.09	0.82	0.83	19.55	2.23
2	0.90	1.00	0.95	0.13	0.86	0.87	14.39	3.84
3	0.89	0.95	1.00	0.00	0.88	0.92	12.93	6.66
4	0.09	0.13	0.00	1.00	0.15	0.00	41.41	10.54
5	0.82	0.86	0.88	0.15	1.00	0.94	30.11	15.61
7	0.83	0.87	0.92	0.00	0.94	1.00	10.61	10.34

Where:
Std dev. = standard deviation

High positive correlations among almost all reflective bands were found. The exception was band 4 and, excluding it, the lowest correlation between any pair of bands was 0.82. The higher standard deviations were concentrated in the infrared region (bands 4,5 and 7).

Table 5.2a and **5.2b** present a list of the image statistics, eigenvalues and eigenvector loadings for a principal component transform, using the covariance matrix, on the 1987 Landsat TM imagery. The set consisted of the six reflective (visible plus near- and mid-infrared) bands (numbered 1 to 5 and 7) of a Landsat TM5 from the INF and adjacencies. Covariances were used in this case because the imagery was radiometrically corrected and rectified (calibrated) previously and, therefore, it was assumed that the digital counts in each band were related to the same physical units, as recommended by Mather (1999). The imagery acquisition and pre-processing methodology is detailed in **Chapter 3**.

Table 5.2a Variance-covariance matrix for the INF Landsat TM, for the 1987 set.

Bands	1	2	3	4	5	7	V_{min}	V_{max}	Mean	Std dv.
1	5.05	7.74	13.30	2.29	28.61	19.09	0	136	19.55	2.25
2	7.74	15.03	24.12	5.49	51.81	34.61	0	200	14.39	3.88
3	13.30	24.12	44.21	0.41	91.86	62.98	0	220	12.93	6.65
4	2.29	5.49	0.41	112.03	25.61	0.29	0	247	41.41	10.58
5	28.61	51.81	91.86	25.61	244.09	151.77	0	218	30.11	15.62
7	19.09	34.61	62.98	0.29	151.77	107.07	0	186	10.60	10.35

TM5 has a variance of 244.09, which corresponds to 46.27% of the total variance. In reality, the infrared bands accumulate approximately 88% of total variance and consequently they dominate the highest-order principal components (**Table 5.2b**). Loughlin (1991) stressed the importance of the magnitude of standard deviation (Std dv.) over the actual weighting of original bands mapped

into successive components. The eigenvalues give an indication of decreasing variance in successive PCs.

Table 5.2b Principal components loadings (eigenvector matrix) for the six principal components of INF TM imagery, based on the covariance matrix

Bands	PC 1	PC 2	PC 3	PC 4	PC 5	PC 7
1	0.10	-0.01	0.20	-0.17	0.30	0.91
2	0.17	0.00	0.40	-0.29	0.75	-0.41
3	0.31	-0.08	0.63	-0.39	-0.59	-0.05
4	0.08	0.99	0.11	0.09	-0.04	0.00
5	0.78	0.03	-0.56	-0.29	0.00	-0.01
7	0.50	-0.14	0.28	0.80	0.05	0.01
Eigenvalue	396.20	112.69	10.76	5.75	1.25	0.82
%var	75.11	21.36	2.04	1.09	0.24	0.16
Cum.% var	75.11	96.47	98.51	99.60	99.84	100.00

Where:
PC1 to PC7 = Order of principal components
%var = percentage variance
cum.% var = cumulative percentage variance

The first result to be noticed is as expected, the largest portion of the variability is concentrated in the PC1 (75.11%). It is possible to verify by Column 1 of **Table 5.2b** (which gives the relationship between the first PC and the six bands) that the largest loadings came from TM5 and 7. This points to a tendency for concentration of information in this region, and possibly indicating a contrast between the infrared and the visible region. PC1 represents the overall scene brightness and it is sometimes called the "albedo" image, in which the soil background is also represented. According to Loughlin (1991), the remaining components should therefore account for decreasing variance caused by differences between spectral regions and between individual bands.

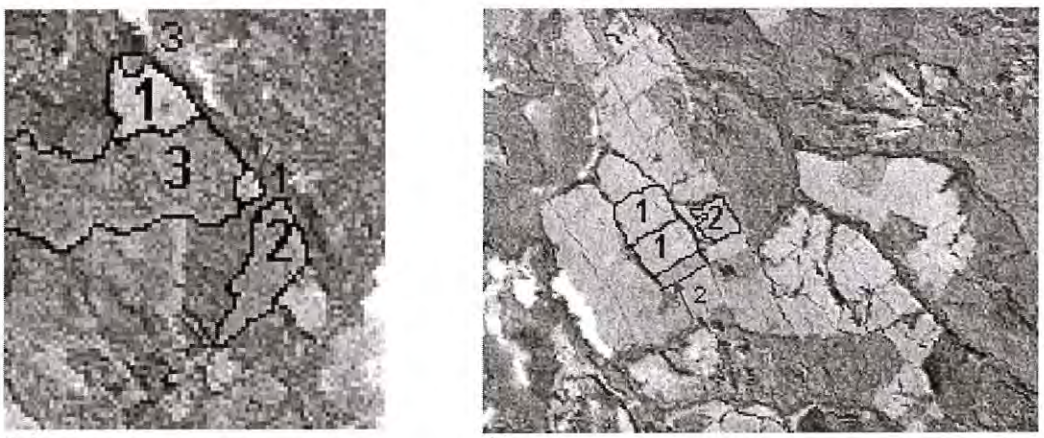
Loughlin (1991) mentioned that examination of the magnitude and sign of eigenvector loadings gives an indication of which spectral properties of vegetation, rocks, and soil are responsible for the statistical variance mapped into each PC. PC2 presented positive loading on the near-infrared and negative loading on the visible band TM3, resulting that the green pattern was highlighted in this component. It was massively dominated by TM4 but also was disturbed by the presence of topographic effect. The image presented no tendency of differentiation among the forest types despite a 21.26% concentration in the variability. Chavez (1992), found that there is a lack of contrast between vegetated and non-vegetated scene elements in the near-infrared bands, with respect to the visible and mid-infrared bands. The fact was reaffirmed by Ricotta *et al.* (1999), when the analysis of the mean local variance as a function of spatial resolution in the Olbian region, Italy. In the present study, TM4 did not present a great contribution for forest discrimination type. Even presenting a darker tone in some *A.angustifolia* plantation stands and in some pine plantations areas, forest types were not easily distinguishable as they were in PC1 and PC3. However, it seems that PC2 is more sensitive to total biomass than to general cover-type

discrimination. **Figure 5.4** shows the difference in tone for very young pine stands outside the INF (1) and the older pine stands inside the INF (2) in a sub-scene of PC2.



Figure 5.4 Sub-scene of the non-standardised principal component 2 (PC2) of TM bands 1-5 and 7. A young pine stand outside the INF (1) and an older pine stand inside the INF (2) are marked.

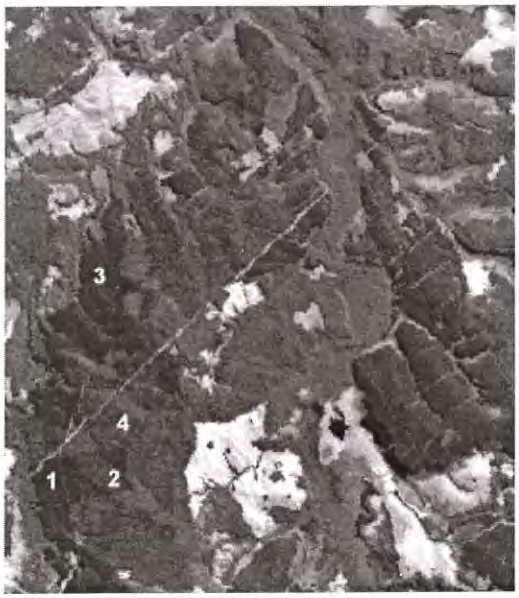
Analysing PC3, TM4 presents very low entries; the other bands were relatively well represented with prominence for TM3, TM5 and TM7. These ones actually seem to be the bands that most contribute to the discrimination of forest types in the study area. PC3 represents actually only around 2% of the variability but the image is also strongly dominated by vegetation. Traditionally, the aim of PCA is to find principal components in accordance with maximum variance of data matrix. However, as mentioned by Ricotta *et al.* (1999), the major consequence of ranking PCs in terms of the amount of variance explained is that the influence of noise becomes increasingly important in the higher-order PCs which are dominated by residual variance. In a recently published paper, Chang and Du (1999) mentioned that a band with small variance does not necessarily mean poor quality; they developed some approaches that considers noise ratio-based to rank transformed principal component. In the case of the present work, clear discrimination between the two species of Pine (slash and loblolly pine), and between *Araucaria angustifolia* (plantation) and *Araucaria* forest were achieved (**Figure 5.5**).



(a) (b)

Figure 5.5 Sub-scenes of the non-standardised principal component 3 (PC3) of TM bands 1-5 and 7. (a) shows (1) pine plantation, (2) *A.angustifolia* plantation and (3) *Araucaria* forest. (b) shows (1) *P.elliottii* and (2) *P.taeda* stands.

A less influence of topographical conditions was noted. PC3 appeared much more flat then PC1 and PC2 and this effect is desired because it can contribute more to the process of segmentation and classification. **Figure 5.6** shows the three first PCs obtained by non-standardised PCA, in a sub-scene, which includes the INF. The other three PCs, as expected, were dominated by noise, providing much less information that is related to this study.



(a) PC1 Albedo image



(b) PC2 Vegetation image



(c) PC3 Vegetation image

(a) PC1 (albedo image); (b) PC2 and (c) PC3 (vegetation images). Labelled are the main features of the INF. Stands of slash pine are labelled with (1), loblolly pine with (2), *Araucaria angustifolia* plantation with (3) and *Araucaria* forest with (4)

Figure 5.6 Sub-scene of the study area, focusing the INF. (a) to (c) respectively show the non-standardised principal components 1-3 (PC1, PC2 and PC3) of TM bands 1-5 and 7.

5.3.1.2 Selective PCA - Experiment 2

Experiment 2 is based on selective PCA. The basic theory underlying this is that selective PCA can be used to enhance and map the spectral differences or contrast between different spectral regions. According to Chavez *et al.* (1989) it can be used when the user is interested in information that is unique to some spectral band rather than in the spectral characteristics of all bands.

A variation of this technique was presented by Loughlin (1991). It relies on the ability of the principal component transform to map increasing subtleties of data variance into successive components. The technique has been informally named "Crosta technique" after Crosta *et al.* (1989). Basically, it advocates the idea of reducing the number of bands input for PCA. Crosta *et al.* (1989) commented that this ensures that certain materials will not be mapped and increases the likelihood that others will be unequivocally mapped into only one of the principal component images. Actually, the technique was developed to detect minerals and was intended to reduce the effect of vegetation, which frequently dominates landscape scenes. In the present study, one of the most important objectives was to test different techniques which could help to discriminate the species *Araucaria angustifolia* even in the natural stands and standard or selective PCA were tested to verify their applicability for this subject.

In this study, two approaches were adopted to test the suitability of selective PCA. Both approaches used the correlation matrix (**Table 5.1**) to identify pairs (or groups) of bands. The first seeks (**section 5.3.1.2.1**) for high correlation among bands and dimensionality reduction; the second approach (**section 5.3.1.2.2**) looks for medium to low correlation among bands for spectral contrast mapping.

5.3.1.2.1 Approach 1 when high correlation exists between bands

The analysis in **Table 5.1** shows that as good candidates for the dimensionality reduction should be grouping the bands of the visible region (TM1, TM2 and TM3) together and the bands of middle-infrared (TM5 and TM7) in a second group. Chavez *et al.* (1989) suggested that a colour composite made by using the first component of each of these two groups plus TM4 should look very similar to a colour composite made from three TM bands, one from each of the spectral regions. The results for grouping the visible bands separately from the two mid-infrared bands are shown in the **Table 5.3**.

The objective of using highly correlated images was dimensionality reduction and it was achieved. Basically, TM3 (with a variance of 44.21, which corresponds to 68.77% of the total variance) dominated the PC1 in the first PCA set (a), as expected.

Table 5.3 Principal components loadings (eigenvector matrix) for selective PCA (based on the covariance matrix) of INF TM imagery – highly correlated bands

	Bands	PC 1	PC 2	PC 3
(a)	1	0.26	0.31	0.91
	2	0.47	0.78	-0.40
	3	0.84	-0.54	-0.06
	Eigenvalue	61.97	1.50	0.82
	%var	96.39	2.33	1.28
	Cum.% var	96.39	98.72	100.00
	Bands	PC 1	PC 2	
(b)	5	0.84	-0.54	
	7	0.54	0.84	
	Eigenvalue	342.10	9.06	
	%var	97.42	2.58	
	Cum.% var	97.42	100.00	

The largest portion of the variability is actually concentrated in the PC1 (over 96%). Vegetation and non-vegetation features are distinguishable but as stated by Lillesand and Kiefer (1994) it is virtually impossible to distinguish between coniferous and deciduous vegetation in the visible spectrum, as confirmed in this study. PC2 and PC3 are dominated by TM2 and TM1 respectively and they presented no useful additional information. In fact, PC3 is very noisy, in part due to the high degree of correlation between the input images (Chavez *et al.* 1989) and maybe in part because of the atmospheric interference (frequent in TM1), despite the radiometric corrections performed in the data set. If space in disk and computation time is a concern, then it might be suggested to use only TM3 (representing the visible region) and the infrared bands, and it seems that not much information will be lost.

TM5 and TM7 were combined because of their high correlation. By visual analysis, it was noted that PC1 shows similarities with the first PC from the standard PCA and PC2 is quite similar to PC3 from that standard PCA. It is not difficult to understand why this happened because of the dominance of these two bands over the highest-order principal component images in this study. PC3 is now much more influenced by topography than in the first analysis and the pine species were not clearly distinguished as they were when the six bands were involved. It seems that more balanced entries were accomplished using TM5 and TM7 but still 97.42% of the variability is concentrated in PC1. Chavez *et al.* (1989) recommended visual analysis before discarding the second PC of TM5 and TM7 because they are not adjacent to one another as the visible bands are and perhaps some contrast of interest will be present. In this case, coniferous and deciduous vegetation presented different tones but not different texture as in PC3 from standard PCA. **Figure 5.7** presents a colour composition where PC1(R), PC5-7 (G) and TM4 (B).

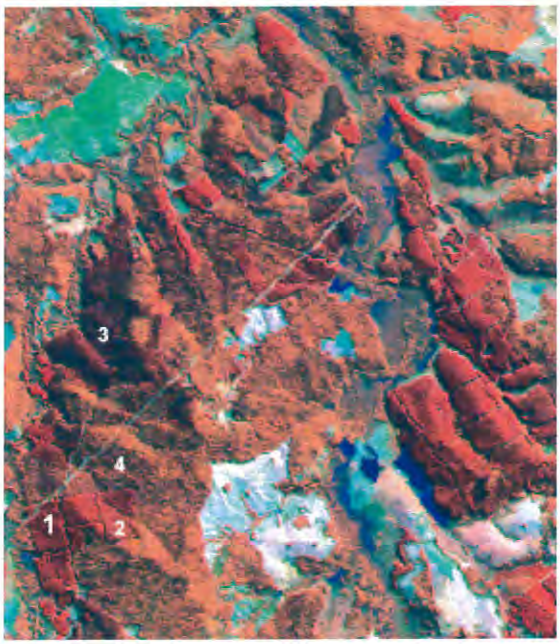


Figure 5.7 Colour composite of PC1 (from selective principal components of TM1 to TM3), PC1 (from selective principal components of TM5 and TM7) and TM4. Numbered are the main features of the INF. Stands of slash pine is labelled with (1), loblolly pine with (2), *Araucaria angustifolia* plantation with (3) and *Araucaria* forest with (4)

The image is thus affected by topography from the influence of TM4. A window of about 512 by 512 pixels (covering the whole INF) was selected and stands of slash pine are labelled with (1), loblolly pine with (2), *Araucaria angustifolia* plantation with (3) and *Araucaria* forest with (4). The Imbituva river, Antas river and their tributaries and flood-plain area (alluvial soils) around them are represented in tones of blue.

5.3.1.2.2 Approach 2 when low correlation exists between bands

The objective of this approach was to look at the spectral contrasts in the study area instead of its similarities. The analysis of **Table 5.1** showed that some bands presented very low degrees of correlation, sometimes being null. It was also noted that, actually band 4 presented very low degrees of correlation with all other bands. Therefore, the analysis was centred in this band and its combination with all other bands was tested. The following pairs of TM bands were considered, after the mentioned selection by analysing **Table 5.4**.

Table 5.4 Correlation coefficients, percent of variance mapped to PC1 and PC2, loadings (eigenvectors) for each pair of TM bands – spectral contrast for INF

TM Pair	Correlation Coeficient	PC1% Variance	PC2% Variance	PC1 Loadings	PC2 Loadings
TM1 and 4	0.09	95.73	4.27	(0.02, 1.00)	(1.00, -0.02)
TM2 and 4	0.13	88.42	11.58	(0.06, 1.00)	(1.00, -0.06)
TM3 and 4	0.00	71.71	28.29	(0.01, 1.00)	(1.00, -0.01)
TM4 and 5	0.15	69.89	30.11	(0.18, 0.98)	(0.98, -0.18)
TM4 and 7	0.00	51.14	48.86	(1.00, 0.06)	(-0.06, 1.00)

In the study area, low correlation refers to any band and particularly TM4. Explaining the reasoning of the method, Chavez (1989) mentioned that, instead of using highly correlated images (0.90 or greater) as input to PCA and using the first component, pairs of only two images with medium to low correlation are used as input. In the present study, some of the combinations like TM1/TM4, TM2/TM4 and even TM3/TM4 did not present any advantage in terms of contrast, when visually compared with others images obtained by standard PCA. **Table 5.4** shows the correlation coefficients, percent of variance mapped to principal components 1 and 2 (PC1 and PC2) and the values of the eigenvectors that are used as multiplication coefficients or loadings in PCA for each pair of TM bands. For example, for the pair TM1/TM4, band TM1 dominated the resultant PC1 (95.73% of the variance in the analysis, which considers only that pair of bands). As considered before, the contribution of the bands of the visible region, especially bands 1 and 2 were not noticeable, in terms of forest types discrimination. The PC1 resultant of the pair TM4/TM5 should be considered as a candidate for further analysis, but it presented less useful information than PC1 obtained from standard PCA.

According to Chavez (1989), by using only two bands as input, the information that is common to both will be mapped to the first component and information that is unique to either one of the two images will be mapped to the second component. **Table 5.4** shows that high percentages of variance in PC2 are related to the low correlation between the two input bands. The most remarkable results were found when TM4 and TM7 were combined. Chavez (1989) commented that the lower the correlation between the two input bands, the higher the percent of variance that gets mapped to component number two. In fact, PC4/PC7 presented balanced variance between bands (51.14 and 48.86%, respectively) and large variance mapped to PC2, which means a larger amount of spectral contrast. As mentioned in **Chapter 4**, band 7 was sensitive to the presence of the species *Araucaria angustifolia* in the natural stands of *Araucaria* forest. In the first PC, despite the image being dominated by TM4, *Araucaria angustifolia* plantation appeared with its typical and distinguishable tone and texture. Chavez (1989) mentioned that, for selective PCA using the approach were contrasts are searched, the second component is the one of interest. For this study, the second PC (PC2) did not present variable illumination caused by variations in topography and coniferous and deciduous species presented differences in tone. For pine and paraná-pine plantations, the tone was quite uniform, very dark. In the areas of *Araucaria* forest, a mixed grid of two tones is present (dark and bright), and it seems that the image is capable of detecting concentration of coniferous in the natural stands. As the only conifer that can vegetate in *Araucaria* forest is *A.angustifolia*, it is possible to say that the species can be identified in the natural areas, when the density (number of trees) is reasonable. There is a marked coincidence between the dark pixel areas and the regions (patches) where higher concentration of the species was detected during the fieldwork. There are no dark dots in the regions of sub-tropical forest, riparian areas, or in alluvial soils region (floodplain), where the species *Araucaria angustifolia* certainly does not grow **(Figure 5.8**).

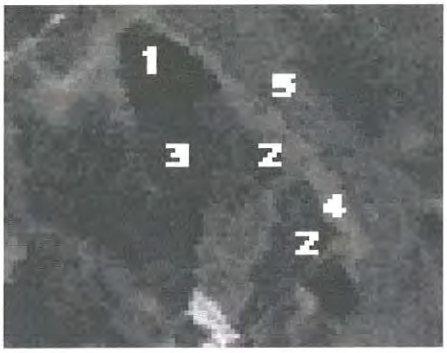


Figure 5.8 Sub-scene showing the presence/absence of conifers in the INF: (1) Pine plantation; (2) *A.angustifolia* plantation; (3) *Araucaria* forest; (4) alluvial soil (floodplain areas – absence of conifers); and (5) riparian areas (area with absence of conifers)

The fact was confirmed by analysis of the available aerial photographs, scale 1:25.000. Besides the analysis of the net of plots allocated in the field, personal communication from INF researcher confirmed the concentration of primary *Araucaria* forest remnants in the area labelled as (4) in **Figure 5.9** (**a** and **b**).



(a)



(b)



(c)



(d)



(e)

Figure 5.9 Series of images illustrating presence/absence of conifers.

In the **Figure 5.9**, image (a) is a magnified picture of a sub-scene of INF, showing one of areas with concentration of the species *Araucaria angustifolia* in the natural areas of *Araucaria* forest (under label 4). Label (1) identifies *P.elliottii*, label (2) identifies *P.taeda*, label (3) identifies *A.angustifolia* plantation. Image (b) is a scanned picture of the aerial photograph available from the area, in scale 1:25000 (more information in **section 3.2.2**). Picture (c) and (d) shows the first (PC1) and second (PC2) components of the selective PCA (derived from the matrix of covariances) performed in the area of the INF using bands 4 and 7; finally picture (e) shows the ratioed image of TM4 and TM7 (ratio47). According to Chavez *et al.* (1989) this second PC (which represents the difference between the regions of the spectrum represented by the two images) could also be compared with the image created by the ratio between the two involved bands. The author mentioned that the ratio of two images is directly related to the slope between the two spectral regions covered by the two images. Mather (1999) mentioned that ratio techniques are used to detect the magnitude of the differences between spectral bands. In this study, this refers to a

comparison between PC2 image of TM4 and TM7 versus ratio47. In the study area, they were, in fact highly correlated, but ratio image appears to be saturated in some areas (**Figure 5.9 e**). All images have been subjected to a histogram equalisation contrast stretch.

5.3.2 Tasselled Cap (Kauth-Thomas) transformations

Kauth and Thomas (1976) developed a method, which was intended to deal with the problem of identifying some agricultural crops systematically on a large scale, based on RS. According to Campbell (1996), the original approach called tasselled cap (TC) by Kauth and Thomas (1976), is a linear transformation of Landsat MSS data that projects soil and vegetation information onto a single plane in multispectral data space. It intends to define a new (rotated) co-ordinate system in terms of which the soil line and the region of vegetation are more clearly represented (Mather 1999). The reasoning was extended by Crist and Cicone (1984) to Landsat TM data. The concepts involved are: a) greenness is a band that conveys information concerning the abundance and vigour of living vegetation, derived from the contrast between visible and near infrared bands; b) brightness is a weighted sum of all bands (four in the original MSS approach and six in the modified procedure, what includes all TM bands but the thermal infrared one); and c) wetness is the third dimension introduced by Crist and Cicone (1984) and is a contrast between the mid-infrared bands (TM5 and TM7) and the red/infrared bands (TM3 and TM4). The images' greenness and brightness were also used in the process of radiometric rectification, as detailed in **Chapter 3**. **Table 5.5** presents the coefficients for the tasselled cap functions as derived by Crist and Cicone (1984).

Table 5.5 Coefficients for the tasselled cap functions for TM bands 1 to 5 and 7.

Band	TM1	TM2	TM3	TM4	TM5	TM7
Greenness	-0.2848	-0.2435	-0.5436	0.7243	0.0840	-0.1800
Brightness	0.3037	0.2793	0.4743	0.5585	0.5082	0.1863
Wetness	0.1509	0.1973	0.3299	0.3406	-0.7112	-0.4572

The negative weights of greenness on the visible bands tend to minimise the effects of the background soil while its positive weights on the infrared bands emphasise green vegetation signal. In this study, the greenness image, defined as a contrast between visible and near-infrared regions, is dominated by TM4 with great interference of topographic conditions, despite the INF area is being relatively flat (only 200 m. of difference between altitude extremes). Vegetated versus non-vegetated (soil, water) areas are discernible with prominence given to *Araucaria angustifolia* plantations, which appeared in the darkest tone. It produced the same and expected results as any tested vegetation index, even if in this case it included a weighted approach, with prominence to TM3 and TM4. The brightness image is influenced to a small amount by TM5. It separated coniferous and deciduous vegetation but it did not differentiate pine from paraná-pine plantations. It

was also influenced by the topography and was not considered more helpful than the original TM5 by itself. The wetness image was designed to be an interesting contrast between visible and infrared regions. It discriminated between vegetated and non-vegetated areas and more specifically between coniferous and deciduous vegetation, becoming the best performance of the three, as expected. Even if it was not capable of separating pine from paraná-pine plantations, it was free of topographical influences. The image was considered a good candidate for the segmentation process. **Figure 5.10** presents a window around 200 by 300 pixels (area of the INF) where the four main forest types are present.

Mather (1999) argued that both the tasselled cap and its TM variation are based on samples of North America soils and mentioned posterior tasselled cap-like functions that uses Gram-Schmidt orthogonal polynomials. The author suggested that the functions should be calculated for the area of study as they may differ from those provided by Crist and Cicone (1984). Campbell (1996) considered that TC is consistent from one scene to another, permitting use of a single transformation to monitor crops throughout a growing season. However, he emphasised its sensitivity to atmospheric turbidity, angle of illumination, and other factors not related to vegetation status.



(a) greenness image (b) brightness image (c) wetness image

Figure 5.10 (a) greenness image; (b) brightness image; and (c) wetness image obtained from Landsat TM5 bands 1 to 5 and 7. The main features of the INF are numbered: stands of slash pine is labelled with (1), loblolly pine with (2), *Araucaria angustifolia* plantation with (3) and *Araucaria* forest with (4).

5.4 Conclusions

This chapter continued the search started in **Chapter 4** to evaluate of the fundamental spectral feature space and dimensionality of the data. After the preliminary steps analysing the relationship between bands and their correlation, as detailed in **Chapter 4**, dimensionality reduction and feature selection could be better understood and were explored in this chapter. Some techniques such as Vegetation Indices (VIs), Principal Component Analysis (PCA) and the tasselled cap (TC) by Kauth and Thomas (1976) were used as aids to feature selection and information extraction. This chapter sought information about: a) general cover-type discrimination, b) coniferous versus deciduous vegetation discrimination; c) discrimination between *A.angustifolia* plantation and the two species of the genus *Pinus* spp most frequently planted in the region, *P.taeda* – loblolly pine and *P.elliottii* – slash pine; d) discrimination between *Araucaria* forest and other types of forest cover composed only of deciduous species (sub-tropical forest or riparian areas); and, finally e) discrimination of clusters of *A.angustifolia* in the *Araucaria* forest at pixel level.

A Landsat-5 TM scene of the INF (Irati National Forest) and surroundings was analysed in conjunction with some ancillary data such as aerial photographs and local maps. Preliminary geometric and radiometric corrections were carried out, as mentioned in **Chapter 3.**

In recent years, several Landsat investigations have focused on the use of image transforms to enhance specific targets. The RATIO being a part of the image ratioing technique, might be one of the first transforms mentioned in the literature. It utilises spectral reflectance data collected in the visible red (0.63-0.69 μm) and near-infrared (0.76-0.90 μm) wavelength regions. The generalisation of the ratio using both bands has been referred as vegetation index and its most utilised variation is certainly the normalised difference vegetation index (NDVI). Both transforms were tested in this study. In the case of RATIO, the results were not useful for the purposes of the present study, but its variation, the ratio between bands TM4 and TM7 presented intriguing results about discrimination of *A.angustifolia* under the natural stands of *Araucaria* forest. Some areas were presumed to have a higher concentration of the species than others. However the results might it be scene specific. In the case of NDVI, the conclusion is that the vegetated areas were really distinguished from the non-vegetated areas. However, the expected ability to minimise topographic effects was not achieved. All the analysed forest types can be considered highly dense (at least within the National Forest), when compared with patches in the adjacent areas and the NDVI discriminated the amount of vegetation only slightly. The conclusion is that apparently the difference between RATIO and NDVI is not great, but NDVI might be more appropriate for use in studies where comparisons over time, what is studied in **Chapter 7**.

The indices mentioned before utilise only two bands. Even after an investigation about which ones should be involved in the process, the other four bands (in the case of the present study) are still judged to be unacceptable, apparently losing the opportunity of inclusion of their available

information in the analysis. On the other hand, multispectral images often have extensive interband correlation and, consequently, the images may have similar spatial structure and contain similar information. PCA is a technique for removing or reducing the duplication or redundancy in multispectral images and for compressing all of the information that is contained in the original channels to their principal components. It was used in the present project with the objective of verifying its usefulness in the process of enhancing and mapping the spectral differences or contrasts between the four main different forest types present in the INF with emphasis to *A.angustifolia*.

The terminology utilised for PCA is sometimes confuse, but it is usefully defined as an application that relies on within-band variance/between-band covariance matrices, eigenvectors which are used to rotate (transform) original image bands such that each rotated band is statistically orthogonal to (independent of) all other bands (Carr *et al.*1999).

Three different approaches were designed to analyse the contribution of the technique to the objectives of this thesis: a) standard PCA (all six bands as inputs) using the non-standardised procedure (covariances matrix); b) selective PCA – approach when high correlation exists between bands; and c) selective PCA – approach when low correlation exists between bands.

Standard PCA revealed that the spectral data could be reduced to three eigenvectors with a loss of just around 1.49% of the scene variance. There is an agreement between these results and the ones mentioned by Crist and Cicone (1984) and Horler *et al.* (1986) who mentioned that the great majority of information could be presented in a three-feature subset of the data. The infrared bands accumulated approximately 88% of total variance and consequently they dominated the highest-order principal components. The largest portion of the variability was concentrated in the PC1 (over 75%, which represented the overall scene brightness and it is also called the "albedo" image). Coniferous stands were clearly darker than the other types of vegetation, in this image, which also discriminated non-forest from forest areas. PC2 presented, generally, no tendency of differentiating among the forest types at all, despite a 21.26% concentration in the variability. PC3 represented actually only around 2% of the variability but the image was strongly dominated by vegetation. Clear discrimination between the two species of Pine (slash and loblolly pine), and the separateness among *Araucaria angustifolia* (plantation) and *Araucaria* forest were achieved. A diminishing influence of topographical conditions was noted. PC3 appeared much more flat then PC1 and PC2 and this effect is desired because it can contribute to the more suitability of the segmentation and classification processes.

The selective PCA was based on two opposite approaches. In the first one, highly correlated bands were united (TM1, TM2 and TM3, group 1) and (TM5 and TM7, group 2), with the objective of dimensionality reduction, which was achieved. The PC1 (group 1) accounted for 96.39% of the variability of the visible region and the PC1 (group 2) accounted for 97.42% of the infra-red region,

excluding the near-infrared, which was left apart, because of its low correlation with the other bands. However, the whole approach did not bring advantages when compared with the standard PCA. In the second approach, poorly correlated bands were united and TM4 played a significant role in the process, being the less correlated band with any other of the data set. Different combinations were tested two-by-two and the most remarkable results were achieved with the union of TM4 and TM7. The second PC of this combination provided discrimination between coniferous and non-coniferous vegetation (tone and texture) and presented differences also in tone and texture between *Araucaria* forest and other natural forested areas. Some areas of the *Araucaria* forest appeared darker than others, indicating different concentration of conifers, which was checked in the analysis of aerial photographs (scale 1:25,000) of the same area and by analysis of the field plots. The conifers in the area are the two species of the genus pine (*P.elliottii* and *P.taeda*) plus the native specie *Araucaria angustifolia*. Considering that it is impossible that the pine species are sharing the same stands with natural forest, the dark dots indicate the presence of *A.angustifolia*. A more detailed analysis was performed in **Chapter 7**, when the multitemporal set were superimposed to the fields plots and compared with information provided by the aerial photographs. No one can deny that this information might be exaggerated (and or saturated) by influences other than just the presence of clusters of conifers in the natural area. However, as a tendency there was an agreement between the aerial photo interpretation and the information contained in PC2 of TM4 and TM7.

The greenness and brightness images from tasselled cap (TC), generated during the process of rectification were considered potentially helpful in the process of forest types discrimination. Because of this, the three images (wetness was included in the analysis) derived by the coefficients suggested by Crist and Cicone (1984) were considered in this chapter. The conclusions are that only the wetness image could be an improvement against the other results already obtained. The suggestion of Mather (1999) of calculating specific functions or the specific study areas is strongly supported by this work and it is recommended that this should be included in future work.

Chapter 6

SPECTRAL MIXTURE MODELLING - REMOTE SENSING METHODS TO DETECT AND IDENTIFY TARGETS AT THE SUB-PIXEL SCALE

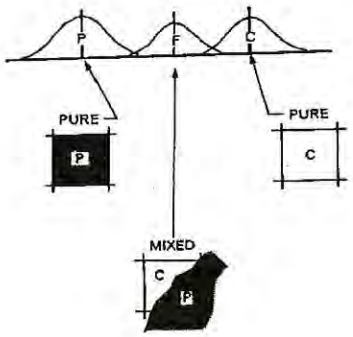
6.1 Introduction

This Chapter analyses the application of Spectral Mixture Modelling (SMM) in general and particularly Linear Mixing Modelling (LMM) and its fraction images derived from Landsat TM in the *Araucaria* region. The method can be used to model the spectral variability in multispectral images and to relate the results to the physical abundance of surface constituents represented by the spectral targets (Tompkins 1997). SMM is based on sub-pixel analysis; its ability to locate *A.angustifolia* as an overstorey indicator species in the natural forest is investigated. This is an essential question in the *Araucaria* forest conservation analysis and *A.angustifolia* is considered here as a target species. *A.angustifolia* is the most harvested species in the biome and some (perhaps the majority) of the forest remnants are nowadays considered just second-growth patches. A method for detecting the presence of the species at sub-pixel level in Landsat TM5 imagery could provide an efficient means for mapping and monitoring the whole biome. As mentioned by Bateson and Curtiss (1996) linear spectral mixture analysis may be preferred to band ratios such as the simple RATIO and the NDVI (as mentioned in **Chapter 5**) for tracking spectrally defined materials because it uses all the bands of data.

6.2 The mixture concept

The techniques traditionally adopted to classify satellite imagery require that each pixel be given a single label (Mather 1999). This kind of approach can be appropriate for large areas of crops or extensive areas with deep clear water but, can be inappropriate for more complex environments. The subject is related to the pixel size and consequently it is clearly a matter of scale. This means the proportions of pixels in an image that are mixed is dependent on the spatial resolution of the imagery (Foody and Cox 1994). It is strongly correlated with spatial resolution because there is a high probability that an image pixel contains information for more than one terrain type. As an example, it is quite unlikely that 1 km x 1 km (AVHRR image) has just one cover type. Even at Landsat TM resolution of 30 m by 30 m what will depend on the desired level of detail. As mentioned in **Chapter 4**, in the *Araucaria* forest, a typical tree of *Araucaria angustifolia* with 40 cm of DBH should have a crown diameter of around eight meters (8 m). Other deciduous species that

belong to the same biome would have, of course, different crown sizes and shapes, but all certainly being smaller than the paraná-pine crown dimensions. With this reasoning, it is a given that until 16 average adult trees appear in one single TM pixel (plus their associated species of the understory), at 100% crown closure. Therefore, pixels contents might be as heterogeneous as is the forest. Because the pixel size in present-day RS systems is large enough to include different types of land cover, the case is known as *mixed pixel* (Shimabukuro *et al.* 1991). **Figure 6.1** illustrates the concept.



In the example, *P* and *C* are areas uniformly occupied. Similar brightness values are averaged and the resulting value does not differ greatly from the original values. In contrast, other land covers are present, the average is composed of more than one value and the resulting pixel value may not accurately represent any of the categories present (Campbell 1996).

Figure 6.1 Representation of mixed pixel when two different land covers are present
Source: Campbell (1996)

The "mixture" concept was first utilised to identify components in chemical mixtures (Bosdogianni 1996). Several investigators have used the idea recently, to detect the presence of minerals, to map differences in soil, and to study vegetation, on a multispectral and multitemporal basis. The pioneer studies came out of the University of Michigan, in the earliest 1970s (Horwitz *et al.* 1971), when researchers started analysing tools to get over the barrier of the limited usefulness of multispectral data due to mixture problems.

The terminology adopted in the different approaches is sometimes confusing and even contradictory. Therefore, in the context of this study, the term Spectral Mixture Modelling (SMM) is used broadly to involve all the procedures developed to detect and identify targets at sub-pixel level.

6.3 Methods for deriving spectral information

Aguiar *et al.* (1999) mentioned that two general directions have been taken in approaching the mixture problem: a) improving the area estimates obtained by conventional classification methods; and b) generating synthetic images, that represents the proportions of each object of the mixture inside of pixels.

Bosdogianni (1996) considered the same concept, within his discussion of methods for deriving information content of the spectral reflectance data. Adapting his classification to the terminology

adopted in the present study, two approaches are used: a) calculation of indices (which could be interpreted as the calculation of image transformations); and b) the spectral mixture analysis, which is broadly called spectral mixture modelling in the context of this study. The author also mentioned that the choice of the method depends on the type of information sought and the complexity of the underlying processes that are depicted by the data. The image transforms approach, including vegetation indices, in the present study is mentioned in **Chapters 5** and **7**.

6.3.1 Spectral mixture modelling (SMM)

According to Bosdogianni (1996), there are two varieties of spectral mixtures: *macroscopic* and *intimate*. When a given photon is reflected (or absorbed) by only one type of material, it is *macroscopic* and the reflected radiances are combined in an additive fashion to produce a linear set of equations, called linear mixture models. This means that linear mixture modelling assumes a single reflectance within the IFOV (Instantaneous Field Of View). According to Schowengerdt (1997), this is an approximation of the reality, in many cases. When individual photons interact with more than one type of material, for example the light might penetrate through the focused component (e.g. salt) and interact with the component beneath. The result is called *intimate* mixtures and gives non-linear equations with their correspondent *non-linear models*.

6.3.1.1 Intimate models

Borel *et al.* (1994) developed a method based on the radiosity procedure, to be applicable to complex surfaces and to volumetric scattering. The authors found a non-linear relationship between vegetation indices and the leaf area index (LAI) using experimental data of cotton plants but recommended further investigation to improve vegetation indices using non-linear unmixing algorithms, mainly at high-resolution level.

According to Bosdogianni (1996), linear approaches have been criticised basically because of their inadequacy to capture non-linear mixing due to backscatter, resulting in a consequent limitation, where the number of end-members is restricted to the number of the imagery bands. He, commented, however, that although non-linear approaches may be more powerful than linear approaches to model far more complex phenomena, they lack the simplicity of the linear models.

According to Roberts *et al.* (1999), for most applications, multiple scattering is assumed to be negligible, although non-linear mixing can become significant for some types of vegetation. Spectral Mixture Analysis was successfully applied to many vegetation studies, but not yet to the *Araucaria* region. However, there is no "a priori" evidence that intimate models might perform better than macroscopic models in dense vegetation (closed canopy) areas like the *Araucaria* forest. In the forested area, in opposition to arid zones (with scarce vegetation), there is no wide variation in soil brightness and, consequently, the best set of end-members might not change profoundly with

canopy density. Therefore, linear mixing is certainly a good modelling approximation for the study area and this thesis will restrict discussion to the linear mixture approaches.

6.3.1.2 Macroscopic models

Under the macroscopic models classification are approaches that involve linear mixing (or mixed) models. Spectral Mixture Analysis (SMA) and the Linear Mixing Model (LMM) are the most common examples.

6.3.1.2.1 Spectral Mixture Analysis (SMA)

According to Lunetta (1999), the SMA was first introduced by J.Gillespie and collaborators in 1990 and later expanded by Adams *et al.* (1995) for use with Landsat TM. The authors, working in the Amazon Region, presented an analytical multispectral and multitemporal approach, to monitor change in areas where the spectrally dominant materials on the ground were well understood and/or could be inferred. Roberts *et al.*(1999) recently published a strategy for using multitemporal SMA change identification, for Eastern Amazonia.

Tompkins *et al.* (1997) mentioned that SMA: a) is a physically based procedure that transforms radiance or reflectance values to physical variables, which are linked to the sub-pixel abundance of end-members within each pixel; b) provides a means to detect and represent components that occur entirely at sub-pixel level; and c) provides quantitative results that can be incorporated into models of the processes governing the distribution of materials within the image scene.

According to Adams *et al.* (1995), SMA differs from many other applications of linear mixing to multispectral images in that it links image spectra to laboratory reflectance data sets. The concept of image end-members and reference end-members is discussed in the **section 6.8**. For now, it is important to stress that the methodology, in the words of Adams *et al.* (1995) is an analytical framework within which images are calibrated to reflectance, and pixels are modelled as potential mixtures of selected laboratory or field spectra. As for the Linear Mixing Model (LMM, mentioned in **section 6.3.1.2.2)**, the linear constraint (fractions of end-members summed to 1 for each pixel) is added to the model.

6.3.1.2.2 The Linear Mixing Model (LMM) approach – basic concepts

According to Shimabukuro *et al.* (1991), the concept of this approach is that a linear relation is used to represent the spectral mixture of targets within the pixel. The response of each pixel in any spectral wavelength can be considered as a linear combination of the responses of each component which is assumed to belong to the mixture. Consequently the idea is that each pixel contains information about the proportion and spectral response of each component within the ground resolution unit. Each pixel can be modelled as a linear combination of a finite set of components:

$$b_1 = a_{11}\ x_1 + a_{12}\ x_2 + \ \bullet\bullet\bullet\ + a_{1n}\ x_n + e_1$$

$$b_2 = a_{21}\ x_1 + a_{22}\ x_2 + \ \bullet\bullet\bullet\ + a_{2n}\ x_n + e_2 \quad \text{OR}$$

...

$$b_m = a_{m1}\ x_1 + a_{m2}\ x_2 + \ \bullet\bullet\bullet\ + a_{mn}\ x_n + e_i \quad \textbf{(1a)}$$

$$b_i = \sum_{j=1}^{n} (a_{ij} x_j) + e_i \quad \textbf{(1b)}$$

or → $Ax = b$ **(2a)** has a unique solution → $x = A^{-1}\ b$ **(3a)**

where:

i = 1,2, · · · m (number of spectral bands); j = 1,2, · · · n (number of components); e_i = error term for i^{th} spectral band; b_i = mean spectral reflectance for the i^{th} spectral band of a pixel containing one or more components; a_{ij} = spectral reflectance of j^{th} component in the pixel for the i^{th} spectral band; x_j = proportion value of the j^{th} component in the pixel.

Two statements are added, being a logical part of the specification of the mixture model. The first one is the linear constraint, since the sum of the proportions for any resolution element is one.

$$\sum_{j=1}^{n} x_j = 1.0 \quad \textbf{(2b)}$$

The second statement is controversial and differentiates whether the approach is to be determined as "constrained" or "unconstrained", as discussed further in **section 6.4**. For the constrained approach the restriction (optional) is:

$$0.0 \leq x_j \leq 1.0 \quad \textbf{(3b)}$$

The maximum number of mixture components must be equal to the number of bands (Mather 1999). When the number of bands (*m*) and components (classes) (*n*) are equal (*m* = *n*), the coefficient matrix A is square. Shimabukuro *et al.* (1991) mentioned that, in the case of the number of components (*n*) being less than the number of the bands (*m*), the system of equations is over-determined and may be solved by any of several techniques. Solving systems of linear algebraic equations involves finding the inverse of the coefficient matrix A. Sohn *et al* (1997) stressed the importance of the inverse matrix (A^{-1}) concept for solving the least-squares problems and Mather (1999) considered the problem of interdependence between the end-members in the unconstrained mixture model, as discussed further in **section 6.4**.

6.4 Constrained versus unconstrained approaches

When the unconstrained approach is adopted, it is important to stress that the model therefore solves the mixture problem equation without considering the constraints **(2b)** or **(3b)** or both of them. Shimabukuro (1995), Sohn *et al.* (1997), NETGIS (1998), Aguiar (1999), and Mather (1999) mentioned some procedures that are proposed to deal with the resultant scenarios. In the first scenario, the values of mixture proportion lie outside the 0-1 range. These values are called *under-shoots* (less than 0) and *over-shoots* (greater than 1). Because these values are illogical, it might

also be illogical to permit that solution. However, as Mather (1999) stressed, the approach (together with residuals values) allows the user to estimate how well the linear mixture model described the analysed data. The author suggested the mapping of these values and the inspection of each fraction image as a method of verifying the fitness of the model.

6.4.1 Singular-value decomposition (SVD)

Another scenario was mentioned by Mather (1999), who focused on the consequences of the application of standard solutions when end-members are correlated. If A is an orthogonal matrix (columns are uncorrelated), then the derivation of the inverse of A is the transpose of A. The traditional solution of the matrix **(2a)** is to calculate the inverse of A, from which x is found from the expression

$$x = A^{-1} b \qquad \textbf{(3a)}$$

If the components (or candidates) are correlated, then errors enter the calculations. Those errors might become significant, if there is high correlation (linear dependency) between end-members; the matrix A is then called *near-singular*. However, if the matrix is *singular* it does not possess an inverse and consequently the least-squares equations cannot be solved. The method singular-value decomposition (SVD) was proposed by Boardman (1989) and suggested by Mather (1999) to solve the problem.

6.4.2 Alternative approaches

More scenarios can also be considered. Algorithms can allow the solution of the model equation, subject to the constraints **(2b)** and **(3b)** as mentioned by Mather (1999) and Shimabukuro *et al.*(1991). Mather (1999) suggested the routine LCLSQ (a Fortran-90 routine from Microsoft IMSL) which allows the constraints **(2b)** to be specified and provides the constraint **(3b)**. He also mentioned routine BVLS, which does not allow for the constraint **(2b)**. Both procedures return the value zero when an under-shoot is calculated. One more method mentioned by Bosdogianni (1996), proposes to substitute the constrained least-squares method. It is called the *unconstrained stepwise regression*, and eliminates components according to the t-statistics.

6.4.3 Constrained and unconstrained least-squares approaches

Shimabukuro *et al.* (1991) presented two methods for estimating the proportions of each component, the Constrained Least-Squares (CLS) method and the Weighted Least-Squares (WLS) method. Both methods can derive fraction images applying the restriction **(3b)** or not. For the purposes of this work, CLS and WLS will refer to the constrained approach and ULS or UWLS will be connect to the unconstrained versions of the method. NETGIS (1998) presented an additional

procedure (PRINCO), a combination between PCA and least-squares, as implemented in the software SPRING. This method seeks to decrease the number of equations in the system by applying initially a PCA transformation, followed by the use of the least-squares method. According to the authors, the method has the advantage of computational speed when the number of components is different from three. However, restriction **(3b)** cannot be applied.

6.4.3.1 The Constrained Least-Squares (CLS) method

The CLS method was presented by Shimabukuro *et al.* (1991) to estimate the proportions of each component inside the pixel by minimising the sum of squares of the errors. They developed solutions for three and four components assumed to be inside the pixel. The derivation of solutions for more end-members is akin to the procedure demonstrated below, with the limitation that $m = n$. The central idea of the method, as mentioned by Sohn (1997) is that general solutions obtained by inversion methods can be mathematically acceptable, but physically unacceptable; in this scope, the negative values are included. Therefore, the restrictions **(2b)** and **(3b)** are imposed to obtain a physically meaningful solution.

The mixture model **(1b)** is then rewritten (in error terms) as:

$$e_i = b_i - \sum_{j=1}^{n}(a_{ij}x_j) \quad \textbf{(4b)} \qquad \textbf{AND} \qquad \text{the function to be minimised is: } F = \sum_{i=1}^{m} e_i^2 \quad \textbf{(5b)}$$

Considering n=3 (number of components) and m=6 (number of TM bands), the function to be minimised is:

$$e_1^2 + e_2^2 + e_3^2 + e_4^2 + e_5^2 + e_6^2 = A_1x_1^2 + A_2x_2^2 + A_3x_3^2 + A_4x_1x_2 + A_5x_1x_3 + A_6x_2x_3 + A_7x_1 + A_8x_2 + A_9x_3 + A_{10} \quad \textbf{(6b)}$$

Considering the constraint $x_1 + x_2 + x_3 = 1$ *or* $x_3 = 1 - x_1 - x_2$ and substituting x_3 into **(5b)**, the function to be minimised becomes:

$$F = A_1x_1^2 + A_2x_2^2 + A_3x_1x_2 + A_4x_1 + A_5x_2 + A_6 \quad \textbf{(7b)}$$

where the coefficients A_1 to A_6 are functions of the reflectance values, a_{ij} is the component reflectance and b_i is the pixel reflectance.

The approach for the solution is to find a minimum inside the area defined by the hatching lines: $0 \le x_1 \le p$, $0 \le x_2 \le q$, and $x_1/p + x_2/q = 1$, where $p=q=1$ (**Figure 6.2**). In the sequence, the partial derivatives are calculated and set equal to zero. According to the outcomes of the unconstrained x_1 and x_2 values, x_1 and x_2 are recalculated if necessary. For an n=3 mixture model, the values of x_3 are determined accordingly. The five possible outcomes are showed in (**Table 6.1**).

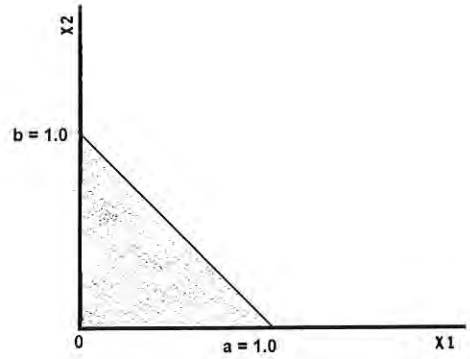


Figure 6.2 Region that satisfies the constraints for n = 3 components
Source: Shimabukuro *et al.* (1991)

Table 6.1 Possible outcomes for the System equation solution

Outcomes	x_1	x_2	Within the region	Values to be recalculated	x_3
1	POS	POS	YES	--	$1 - x_1 - x_2$
2	POS	POS	NO	x_1 and x_2	0
3	NEG	POS	NO	x_2 ($x_1 = 0$)	$1 - x_2$
4	NEG	NEG	NO	$x_1 = x_2 = 0$	1
5	POS	NEG	NO	x_1 ($x_2 = 0$)	$1 - x_1$

Source: Shimabukuro *et al.* (1991)
When: POS=positive and NEG=negative

The CLS method, as implemented in the software SPRING is considered the simplest and fastest of the three available options (CLS, WLS and PRINCO). The restriction **(3b)** can be applied or not, depending on the choice of the user or intrinsic restrictions of the models.

6.4.3.2 The Weighted Least-Squares (WLS) method

This variation of least-squares approach was also introduced by Shimabukuro *et al.* (1991). It is based on an iterative process, instead of CLS, which uses a quasi-closed solution approach. Both methods search for the solution by minimising the sum of squares of the errors. According to Shimabukuro (1991), the WLS method, due its characteristics, seems to be mathematically more precise. Even if the CLS method has the advantage of being faster than WLS, with the improvements in hardware and software this constraint is becoming less significant. According to Shimabukuro (1991), the idea of the method is to give some "weight" to some data values which are considered to be more accurate or more important for the problem. The method relies on fitting a curve to the data with a curve having the form:

$$R = f(A, x_1, x_2, \cdots, x_n) \text{ OR } R = x_1 f(A) + x_2 f(A) + \cdots + x_n f(A) \qquad \textbf{(8b)}$$

Where R is linear with respect to the constants $x_1, x_2, \cdots, x_n$.

The least-squares method can be applied and the solution as in CLS method is mathematically acceptable although, in some cases, meaningless (in the situations of resulting negative values and/or sums higher than one). Becoming a constrained least-squares problem, the constraints equations should be added. In order to solve the problem, it is necessary to apply the concepts of weighting for least squares. To satisfy the restrictions **(2b)** and **(3b)**, four equations are added to **(1b)**, one corresponding to the condition **(2b)**, and another three are corresponding to the restriction **(3b)**. To solve this problem, a diagonal matrix *W* containing weight values associated with the system of equations to be solved is used when the constraints are violated. Initially, the first *m* values of one along the diagonal of matrix *W* mean that the corresponding equations are equally important to the solution of the problem. In Shimabukuro's *et al.* (1991) words, the following weight value represents a very high weight for the equation corresponding to the constraint **(2b)**, which must be satisfied. Hence, if the restriction **(3b)** is also satisfied, the final solution has been achieved. Otherwise, an iterative process is performed, which brings all x_j's inside the interval. The process is to increase the weights (set to zero, initially) assigned to the last three equations, using a small step. The solution is achieved by minimising the quantity, $w_1 e_1^2 + w_2 e_2^2 + ... + w_{(m+n+1)} e_{(m+n+1)}^2$ where w_1, w_2, etc are weighting factors which are real numbers, and e_1, e_2, etc are the residual values for each equation. The WLS method was developed for solving a system of equations. According to Shimabukuro *et al.* (1991), the implementation of this method was based on Gauss elimination and forward and backward substitution algorithms found elsewhere.

6.5 Methods for estimating proportions

Most studies have modelled the remotely sensed signature as a "linear combination end-members" and applied the least-squares approach to decompose the observed signature to component mixtures (Sohn *et al.* 1997). Bosdogianni (1996) added the maximum likelihood estimation (MLE) method as an alternative process of unmixing proportions, noting the restrictions that the method is computationally expensive and a relatively slow procedure compared to the least-squares approach. The author also mentioned that MLE can perform poorly when it is used to classify high-dimensional data.

An interesting procedure was presented by Huguenin *et al.* (1997). It uses the non-parametric Applied Analysis Spectral Analytical Process (AASAP). According to these authors the sub-pixel processing approach used for signature derivation and background suppression yields discrimination performance characteristics generally different from the LMM and from those produced using fuzzy set classification logic. The first step of the process is: a) to extract a pure sub-pixel signature of the material of interest (MOI) – by identifying the background reflectance spectrum. One set of 150 by 150 pixels was used to develop and refine each specie (they were two species of vegetation) spectral signature. The second step in the procedure is b) to remove (by iteration of thresholds) from the signature any units of the pixel that are not material of interest – by subtracting it from the original integrated spectrum, leaving only information about the proportion of

MOI, within the pixel.; and c) extracts and analyses sub-pixel components of each pixel in an image, identifying those pixel components that match the material-of-interest spectral signature. This method presented better results than classical methods of image classification (ISODATA clustering, maximum likelihood, and minimum distance) and was capable of separating species of Bald Cypress (*Taxodium distichum*) from Tupelo Gum (*Nyssa aquatica*), using TM imagery.

Foschi *et al.* (1997) compared two artificial intelligence approaches, i.e., machine-vision and neural-network methods, with the objective of improving classification accuracy of small strips or patches of woody vegetation, using simulated multispectral and panchromatic SPOT HRV imagery. The methods, when compared to standard supervised multispectral classification yield significant improvements in detecting sub-pixel woody vegetation and, in general, the machine-vision approach outperformed the neural-network approach.

6.5.1 Facets of the least-squares approach

Roberts *et al.* (1999) outlined the methods for solution to the linear problems as: a) the least-squares method; b) the singular value decomposition (SVD, already mentioned in the **section 6.4.1**), presented by Boardman (1989) and analysed by Mather (1999); and c) the Gram-Schmidt orthogonalization. They applied a variation of this last method in their study, adopting the unconstrained modified Gram-Schmidt least-squares method, where restriction **(2b)** is applied but not restriction **(3b)**. These restrictions were detailed in **section 6.3.1.2.2**.

Sohn *et al.* (1997) classified the methods that use least-squares methodology as: a) the basic least-squares method; b) Bayesian regression analysis; and c) factor analysis. The basic least-squares method finds the closest point in a given subspace to a given point in function space (Shimabukuro *et al.* 1991, Santos 1999). It was utilised in the present study and the results are to be found in **section 6.10.1.** The Bayesian method assumes a normal distribution for data and noise (Smith 1990). Pereira *et al.* (1998) tested the approach, which calibrates images before the selection of end-members. The procedure is based in the use of pure end-members (reference end-members, obtained from a spectral library) as related to the image end-members. The process is divided into four steps: 1) linear regression - define some pure pixels in the image (image end-member), and compare them with data obtained in laboratory or in controlled field experiments (reference end-members), regressing both data sets; 2) spectral evaluation - the gain and offset of the combinations with r^2 higher than 0.95 are used to convert the image to initial reflectance and the pure components are then subtracted from the reference set; 3) selection of pure reference components - the selection of the pure components are made by modelling the components initially calibrated. A good candidate must have proportions equal to 100% for a pure component of the image and low proportions (or negative) for the others; 4) regression of the mixture – the solution is based on the square root of the mean error, which is calculated for each band and used to calibrate the mixture, in an iterative process. Finally, the procedure factor analysis was also called (by Mather 1999) "image factor analysis". He mentioned that the procedure is actually a principal

component analysis what is concerned with the estimation of the parameters of the factor analysis model. A variation of the procedure was also included by Mather (1999) in his topic about possible alternatives for end-members selection. It was developed by J.Imbrie in 1963 and it is called Q-mode. The method is recommended to be included in future analysis using the same data set of this study.

6.6 Error assessment

The different approaches are supposed to be evaluated as to how well they account for the actual spectral variation in the acquired data. It is important to determine if the model does fit the data or not and to define a strategy to determine the appropriateness of chosen end-members. Evaluation of the results of the mixture analysis can be broadly divided in three ways: visual analysis, calculation of the mean error, and computation of fraction overflow.

6.6.1 Visual analysis

Visual analysis of fraction images that result from the mixture process allows the analyst to determine whether they are consistent with other ancillary information, such as field data, aerial photographs or maps. If the patterns do not correspond with the so called "ground truthing", then the constructed model may not be correct (Kressler *et al.* 1999).

6.6.2 Calculation of the mean error

The error is a measure of the spectral residue that cannot be explained by the mixing model (Adams *et al.* 1989). According to Mather (1999), there is one residual value in each spectral band, representing the difference between the observed pixel value and the value computed from the linear mixture procedure. The objective of the error calculation is to help in the analysis of mixture model adequacy for a certain scene. For each band i, the term of error e_i is given for:

$$e_i = \sqrt{(b_i - \sum_{j=1}^{n} (a_{ij} x_j)^2} \quad \textbf{(9b)} \qquad \text{AND} \qquad \text{Mean total error} = (\sum e_i)/m \quad \textbf{(10b)}$$

where:

b_i = spectral reflectance in the i^n spectral band of a pixel (i.e., pixel value in the band i, converted to reflectance).

a_{ij} = known spectral reflectance of j^n component in the i^n spectral band.

x_j = proportion value of the j^n component to be estimated inside the pixel;

i = 1, . . ., *m* bands

j = 1, . . ., *n* components

It is possible to calculate the error by band **(9b)** and the mean total error **(10b)** by taking the error values for each pixel. A small error is an indication of a mathematically good model (Kressler *et al.* 1999). Additionally, generation of error images can be materialised, which helps in the visualisation

of the fitness of the model. It represents the spatial distribution of the errors. The values of these images are obtained by the multiplication of the absolute value of the e_i by the factor of scale 255. According to NETGIS (1998), because the error values are usually very low, enhancement of those images is eventually necessary to facilitate the visualisation of the spatial distribution of the errors.

6.6.3 Computation of fraction overflow

As mentioned before, under-shoots and over-shoots are normal in the process, when unconstrained procedures are adopted. The constraint **(3b)** or both of them [**(2b)** and **(3b)**] are consequently not applied and some pixels will lie outside the logical range of 0-1. According to Mather(1999), a number larger than 5% of pixel under- or overshoot indicates that the model does not fit well. A procedure should be performed to map this fraction overflow and to include the results in the decision process regarding the fitness of the model. The approach was implemented in this study and the results are detailed in **sections 6.10.1.4.1 and 6.10.2**.

6.7 Applications of SMM in vegetation studies

Although SMM has been used for estimating sub-pixel components in various fields of earth and lunar resource applications, mapping vegetation is still one of the most frequently used techniques. Specifically in Brazil, it was first introduced in the work of Y. Shimabukuro and the INPE team, in the 1980's. They used the Linear Mixing Model (LMM) to discriminate species of pine (*Pinus elliottii* and "others") and *Eucalyptus* spp. The shade image was found to be related to the forest canopy structure caused by differences in ages of the eucalyptus plantations (Shimabukuro *et al.* 1991; Shimabukuro *et al.* 1995). In the same broad region, Sousa (1997) first utilised synthetic images provided by LMM in the evaluation of pine volume changes detection (using TM data) and then Sousa (1998), applied the mixture model (also using TM data) on an area of hilly terrain, with strong anthropogenic action (State of S.Paulo), to discriminate different land covers. In Central Brazil, Holben *et al.* (1993a and 1993b) applied the LMM to a "cerrado" (open woodland) area, using AVHRR data and Aguiar *et al.* (1999) compared mixing model and some traditional feature reduction methods like PCA and Canonical analysis (using Landsat TM imagery) over an reforested area with *Eucalyptus* spp. They concluded that the best results were obtained through canonical analysis and the mixing model procedures.

Some of these procedures were performed in the Amazon Region, using Spectral Mixture Analysis (SMA). One such pioneer effort was the study performed by Adams *et al.* (1995), using Landsat TM. The SMA-based classification was tested on a time series of Landsat TM images and the authors selected four end-members to represent the scene: shade, green vegetation, non-photosynthetic vegetation and soil. They applied the methodology to monitor changes by classifying the images based on those end-members. The conclusion was that the method is a powerful way of monitoring detailed changes in land cover. Roberts *et al.* (1999) applied SMA in a multitemporal scale in the region of Eastern Amazonia.

Yet in the Amazon region, the Linear Mixing Model (LMM) was used by Bernardes (1996), while working with Landsat TM imagery. He compared vegetation indices images and fraction images to characterise primary rainforest and successional stages in the area under the influence of the Tapajós National Forest. He concluded that fraction images were capable of describing the evolution of biomass during the succession process. Biophysical estimation was the objective of Santos *et al.* (1999), who used the LMM to estimate forest and savanna biomass at transition areas in Mato Grosso State, at the border of the Amazon area. These authors concluded that the fraction images (vegetation, soil and shade) brought a significant contribution to the discrimination of the vegetation types and also contributed to the monitoring of change processes due to land occupation in transition zones.

Although some of the studies mentioned were performed in southern Brazil, they focused on the exotic species of the genus Pine and Eucalyptus. The present study is the first to introduce an analytical approach which aims the dimensionality reduction of the feature space using PCA or Tasseled Cap and the SMM in the South Brazil natural forests, and specifically in the *Araucaria* region.

6.8 The selection of end-members

The main objective of SMM is to separate the spectral contribution of each object within a pixel, using a set of pure components of the image as the starting point. Here resides one of the determinant factors of the approach and, as will be discussed later, the impact of finding the "right" end-member and the typical reflectance curve for that end-member is enormous. After the decision about how many end-members are going to be involved in the process and which ones will be included, a question still remains regarding to the way to catch the spectral information for each end-member.

According to Aguiar *et al.* (1999), for forest targets, the components should represent the structural differences of the canopy but in complex environments the task is often very complicated. According to Tompkins *et al.* (1997), the most difficult step in the SMA approach lies in the selection of the spectral end-members, which are chosen to represent surface components. Actually, at this point the issue gives rise to some practical questions such as: a) how many end-members will be involved in the process ?; b) which ones will be included ?; and c) how will the spectral information for each end-member be collected ? It is a complex matter and, as mentioned by Sohn (1997), in many related studies the necessary steps have not been fully considered. Modelling of spectral mixtures and the procedures for choosing end-members appear arbitrary. As mentioned by Mather (1999), in the final analysis, some authors do not give enough details for their readers to discover whether or not their model provides a good fit. This means that questions *a)* and *b)* above are intrinsically related to knowledge of the area and to the quality of ancillary

information that might help in the decision process or in the choice of methods that do not require previous information about the area.

It seems that the candidate end-members can be selected (Bateson *et al.* 1996): a) from the spectra corresponding to pixels within an image - called *image end-member* (derived from training data on images). They can be a result of a deep analysis of spectral characteristics of the representative forest types and the other features present in the scene or be defined empirically, accepting, for example, components like vegetation and soil as definitive candidates; b) from libraries of end-member spectra - called *reference end-members*, which are characterised as invariant laboratory or field reflectance spectra of the focused materials (Sohn *et al.* 1997); c) with a two-step method that models the spectral data with end-members selected from the image and modelled as mixtures of library end-members , as defined by Adams *et al.* (1995). The spectra of forested areas were adjusted by mixtures of laboratory or field spectra (reference spectra) of green vegetation, non-photosynthetic vegetation, and shade. Image end-members themselves were redefined as mixtures of reference end-members, derived from a set of laboratory or field spectra measured in units of reflectance; d) by PCA, suggested by Murphy *et al.* (1994) as a method to identify candidate image end-members, assuming that the most significant variability will be find in the first and second components; (Smith *et al.* 1990); e) with automatic methods of constructing end-members from higher order PCA eingenvectors (Ehrlich *et al.* 1987); f) by a mix method of automatic and manual selection, as suggested by Bateson *et al.* 1996; g) by other alternatives, mentioned by Mather (1999) as a variation of factor analysis. The procedure is called Q-mode and it was developed by J.Imbrie in 1963. This method considers the similarity relationships – measured by the cosine-theta (cos θ) coefficient of similarity between k pixels over m images. A variation of the method was developed by Sohn *et al.* (1997) for mapping desert shrub rangeland; and finally h) the work of Aguiar *et al.* (1999), postulating that the selection of an appropriate feature subset cannot be done in an indiscriminate manner. They suggested performing a comparison between the candidate subsets based on the classification performance. They considered that ideally the criterion should be based on the error probability, recognising, however, that its computation is usually very difficult.

The procedures mentioned in items *b), e) and g)* can be considered as methods that do not require prior knowledge of the area. Some other approaches for selecting mixture components that should be mentioned later are classified under the non-linear procedures, which are past the scope of this study.

Considering the approaches mentioned in the items *a, b,* and *c*, and also starting with the first question asked before, it is possible to summarise that as a rule and as a limitation, the number of mixture components cannot exceed the number of bands (Mather 1999). This should be interpreted as a maximum of six end-members in the case of Landsat TM, although the majority of studies on vegetation advocate the selection of three or four end-members. Because of the natural correlation

regions, desert areas or even in small strips or patches of vegetation. According to Holben *et al.* (1993a), this fraction might be useful for deforestation studies since it contains information about bare soil proportion within the pixels.

Component shade – before detailing the importance of shade, it is important to mention that, actually two types of shade can be distinguished: the first type is attributed to the terrain morphology, as slopes. It can be estimated using appropriate elevation models (DEM) and forms a shade image that should be subtracted from the original one (Bosdogianni 1996). The second type of shade is related to vegetation. Shimabukuro *et al.* (1991) proposed the use of the image formed by the shade proportion in each pixel as indicative of variations in the structure of the forest, that is, the estimation of the shade proportion indicates variations in age, type and forms of the trees canopy. According to Adams *et al.* (1989), as an end-member, shade can mix in all proportions with each of the other end-members or with their mixtures, therefore modelling the spectrum of the end-member material when it is not fully illuminated (shading). To summarise, shade accounts for variations in lighting (at all scales) due to changes in incidence angle and variations (at all scales also) that are caused by shadows. Under the term "shadow" topographic features and the ones cast by trees, bushes or other objects having roughness and texture might be included. The choice of which feature best represents shadow is normally not easy, but dark areas, such as a blacktop road or a body of clear water under some circumstances, can be a proxy for shade (Adams *et al.* 1989). However, it is important to stress that any feature used as measure for shade will always be a surrogate, even if it is calculate. In that case, it can be calculated by using radiative transfer calculations to estimate the amount of light transmitted through the canopy or scattered onto the background from the atmosphere or adjacent canopies.

Component vegetation - this is a very important end-member for monitoring purposes. When an image end-member is used, normally the choice is over the predominant vegetation type. However, as is discussed later, a previous analysis of the spectral signatures of forest types is necessary when discrimination between vegetation types is the objective, mainly when image end-members are being used.

6.9 The pure pixel (end-member) and training data

As mentioned before, two traditional methods of capturing mixture information are: image end-members, normally using training data and/or reference end-members. End-members must be defined from laboratory or field reference spectra when: a) in the image there are no pure pixels that represent the desired materials, or b) if every pixel presents all the desired materials, according to Adams *et al.* (1989) and Sohn *et al.* (1997). As mentioned by Tompkins *et al.* (1997), the use of a reference end-member depends on having a well-calibrated image, and on the assumption that the reference material adequately represents the spectral variability of the end-member in the image scene. According to the author, without extensive fieldwork, appropriate

reference end-members may not be available. The problem is particularly crucial in region when the target is sparse, like vegetation in arid zones or when all the pixels actually contain two or more of the candidates to end-members.

In some circumstances, in any case, it is possible to extract the information directly form the image with reasonable success, using the image end-members approach. In those cases, the concept of pure pixel is especially important. However, pure pixels exist just as a conceptual convenience. Even in extensive portions of deep clear water (uniform and regular distribution of the feature) the sensor noise and within-class variability will be present (Schowengerdt, 1997). Ideally the class should be represented by a pure pixel, composed of a single land cover type, that determine the characteristics of all other image pixels (Mather 1999). The next important aspect to be mentioned is the type of *the image training data*. According to Bosdogianni (1996), information can be captured via *point data* or by *sets of data*.

6.9.1 Sets of data

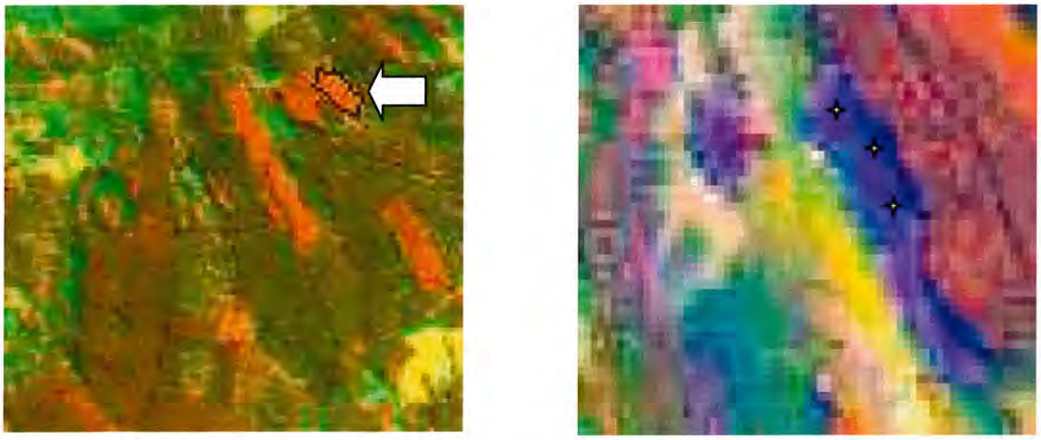
The use of sets of data signifies that means of the selected feature will define the signature of the selected end-member, instead of a point in the space. Several measures are taken in on area that is known by containing homogeneous amount of the desired feature. In the approach adopted by Adams *et al.* (1995), they firstly defined the areas that corresponded to known types of land cover. In the second step, based on field observations, measures of the mean values of the selected four end-members using representative areas in the images were taken.

Bosdogianni (1996) mentioned that the use of *sets of data* is not very common, although the quality of the results of this approach depends on the balance of constituent components in the mixture. However, the author also mentioned that the *point data* produce more inaccurate (biased) results for extreme population ratios in the mixtures. Ouaidrari *et al.* (1996) solved the mixed system using the sets of data procedure and developed a method that is similar to a splitting window. The idealised blocks are square and receive the values equal to the mean reflectance of the pixels inside the block. The mixing system is solved for the first three blocks in the left corner of the image and then the window is moved one position. In the second round, the system is solved again to the new group and so on, until the whole image has been explored.

6.9.2 Point data

Point data are by far the most popular procedure to catch end-members. Each end-member is a point in the space defined by the multispectral bands (Bosdogianni 1996), or it is the value of a pixel known to contain a "single" class. This approach was adopted by many authors (Shimabukuro *et al.*1991; Shimabukuro *et al.*1995; Sousa, 1997; Santos *et al.* 1999; Aguiar *et al.*, 1999). Actually, the idea is to decide by a "good" candidate of pure-pixel and pick it up directly from the image,

accepting that it is "pure" enough to represent the end-member. **Figure 6.4** exemplifies the information collection for *sets of data* and for *point data*.



(a) sets of data **(b)** point data

Figure 6.4 Capture of spectral information in **(a)** *sets of data and for* **(b)** *point data*.

6.10 Experiments description

In the present study, the linear mixing model was applied with the objectives of: a) digital enhancement of phytophysionomics characteristics from INF; b) analysis of its capability for detecting the species *Araucaria angustifolia*, in plantation and natural stands; c) analysis of its usefulness in the change detection process, which is discussed in **Chapter 7**, which deals with the multitemporal aspect of the problem. For the use in this chapter, the methodology can be divided into: a) selection of imagery and previous image pre-processing, detailed in **Chapter 3**; b) generation of fraction images corresponding to the different forest components; c) digital enhancement of fraction images using RGB coloured compositions.

Two approaches were performed: a) the analysis using the six untransformed TM bands considering two end-members of vegetation (*Araucaria angustifolia* and *Araucaria* forest); and b) the same data set performing analysis considering one additional end-member to represent soil.

6.10.1 Experiment 1 - Linear Mixing Model: three candidates as end-members

As mentioned by Mather (1999), in order for the components to be computable, the number of end-members must be less than the number of bands. Even if six end-members can be included in the model, according to Bosdogianni (1996), adding end-members may lead to unstable proportion derivation due to progressive lowering of the spectral contrast between end-members. The author commented that it is commonly accepted that the six reflective bands of Landsat TM are correlated

and they only span a three dimensional spectral feature space, so the number of different end-members is limited. In this experiment, the spectral responses were extracted using the six TM reflectance bands that were pre-processed as detailed in **Chapter 3**. According to Holben *et al.* (1993b) the method adopted in the present study can be used for three or four components assumed to be inside the pixel. Bateson *et al.* (1996) mentioned an approach for determining the number of end-members, which is to view the spectral data as points in the space of the complete set of spectral bands and to find the PCA eigenvectors or directions accounting for most of the variance in the data. In the study area, the preliminary analysis using PCA (**Chapter 5 - Table 5.2b**) showed that the first two components (PC1 and PC2) represented 96.47% of the whole variability. Bateson *et al.* (1996) also mentioned that if there are N eingenvalues that respond for the most variability (two, in the case of the present study), there are N+1 end-members (three, in this case) in the space span, which will be referred to as the mixing space.

6.10.1.1 The selection of end-members

An empirical method should be used in this case, as the one advocated by Murphy *et al.* (1994), who advocated a graphic method to identify candidates image end-members. According to the authors, the purest pixel of each component within the image should be in the place (vertices of the polygon) that bounds the data space of the PCs which contains the information". According to their reasoning, other pixels, which appear within the body of the polygon or between the end-members, are mixtures of one or more end-members. For the present study, a scatterplot with the location of the forest four main forest types, soil and shade is presented in **Figure 6.5**.

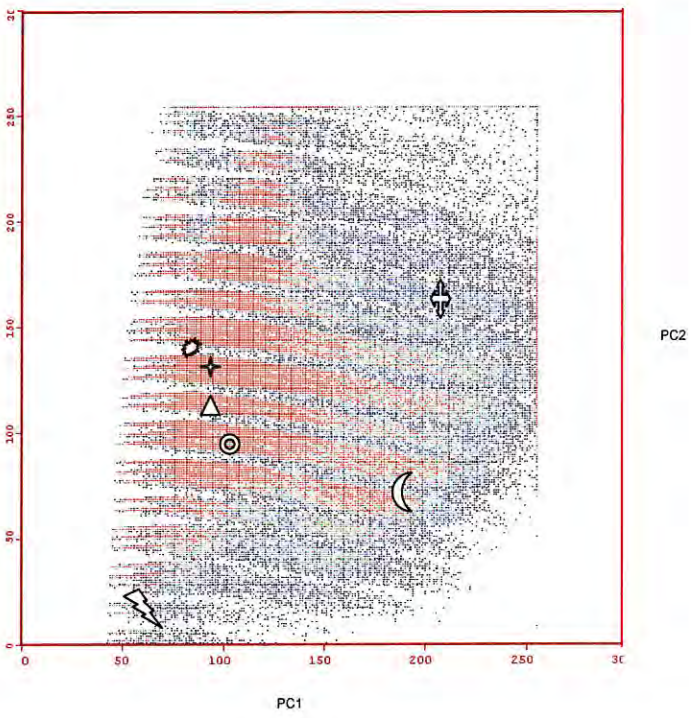


Figure 6.5 Principal component 1 (PC1) - *x* axis and principal component 2 (PC2) - *y* axis. (Area around INF).

In order to have the spatial representation of each feature, the symbols below were adopted to indicate the typical value for each of the main forest types: *Pinus elliottii* ( ), *Pinus taeda* ( ), *Araucaria angustifolia* plantation (△), and *Araucaria* forest (◎). The typical location for soil ( ) and shade ( ) was also determined. A symbol ( ) was included in the figure to facilitate the visualisation of the different types of crops.

Murphy *et al.* (1994) suggested a visual inspection of the graphical display. Actually, the visualisation of the pixels concentration for each feature helps in the spatial representation of the scene, but is not a straightforward methodology. As mentioned by Mather (1999), if PCA is based on covariances instead of correlation matrix, the spectral bands may contribute unequally to the total variance analysed. However, it might be very useful when hyperspectral sets of images are being used. The methodology is not an automated approach of constructing end-members; spectral plots are used to identify each end-member on the basis of its spectral curve shape and absorption features, but field experience and ancillary data are fundamental (Murphy *et al.*1994). In the next step, Murphy *et al.* (1994) used statistics (means and covariance matrix) to generate a model based on the least-square fit mixture. The analysis performed previously and presented in the **section 5.3.1.1**, showed that PC3 mapped the different forest types well. This was the reason for drawing the scatterplot of PC1 and PC3 (vegetation band), besides PC1 and PC2, as presented above. Of course, PC3 responded for much less of the variability in the data set (only 2.04%). However, as mentioned by Mather (1999), significant variability may remain in the lower-order principal components and but when PC3 is added to PC1 it summed 77.15% of the variability.

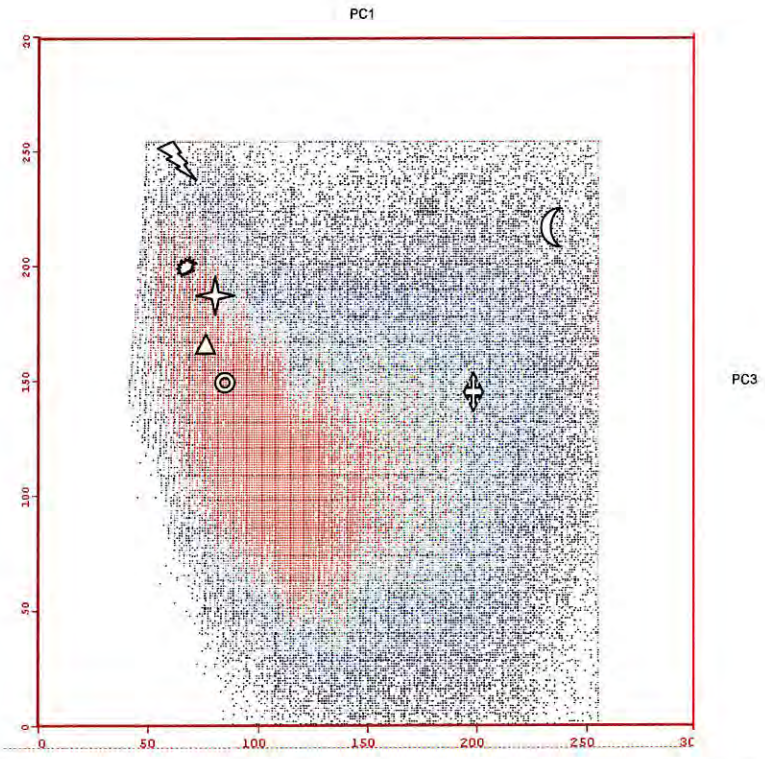


Figure 6.6 Principal component 1 (PC1) - *x* axis and principal component 3 (PC3) - *y* axis. (Area around INF).

It seems that when using the composition provided by **Figure 6.6**, locating the purest pixels becomes easier. However, it is scene related, very much influenced by the analyst's choices and, in the case of the present study a bit more complicated, considering that the area of interest is composed by two types of vegetation. PC3 presented very good results differentiating coniferous from deciduous, and this ability was considered important in the case of natural stands of *Araucaria* forest. It was also very sensitive to the water contents. Actually, moisture acts in a special way, appearing very dark in PC1 and very light in PC3. Considering that bands from the middle-infrared region are sensitive to the vegetation moisture content, they might be responsible for the differences in tone between the two images. It appears that in the upper right corner of the graphic is where the different spectral signatures for agriculture reside. Potato and soybean are common crops in the region as market-oriented farming with high technique pattern for pasture. Maize, beans and tobacco are cultivated by small farmers with subsistence-oriented farming.

Due to the complexity of forest classes present in the area and their similar spectral response, the additional efforts in the search for the best number of primary components were performed following the *image end-member* approach, (data derived from training data on images). The end-members were defined as a result of a analysis of spectral characteristics of the representative forest types and the others features present in the scene (**Chapter 4**). The spectral curves defined in **Chapter 4** were used as a guide. Those values were then considered the minimum limit and were compared to the values obtained in the stands (via spectral plots) that presented extreme circumstances and were the purest pixels might reside. Extreme circumstances were defined as: stands with the highest volumes, or more percentage crown closure, which should contain pure information about the spectral behaviour of each feature. Those values were then adopted for all the candidates to end-member.

6.10.1.1.1 The end-members vegetation and shade

In this study, image end-members were used. All six reflective bands were included and image end-members consisted of only the spectral features that actually participated in the mixing. The species *Araucaria angustifolia* (by itself) and other deciduous species mixed with *A.angustifolia* (*Araucaria* forest) composed the most important mixture presented in the study area. Thus, these two end-members were firstly chosen to compose the vegetation component of the model. Three primary components (two classes of forest vegetation and shade) were then considered, following the indications of the principal component analysis. They were: a) *Araucaria angustifolia* plantations; b) *Araucaria* forest; and c) shade shadow represents areas not or badly illuminated and some authors preferred to define it as zero (Steinnocher *et al.* (1999).

The shade component has been considered necessary for a reliable decomposition (Sohn *et al.* 1997). According to Gillespie *et al.* (1990), shade is typically represented by the null vector, provided the data have been corrected for atmosphere and instruments effects: i.e., "shade" is the

radiance from an ideal black surface. Hall *et al.* (1995) studied the forest biophysical structure using mixture decomposition and came out with the assumption that the shadowed background and shadowed canopy reflectance are roughly equal to the reflectance of deep lakes. The present study area involves some clear water bodies (distinct and "pure"), represented by a reservoir and other small lakes. Those values, then, should be used as values for shade in order to verify its usefulness as an indicator of variations in the structure of the different types of forest. It was utilised as a measure for shade after some reported successful experiences (Shimabukuro *et al.* 1995; Holben *et al.* 1993a). However, a spectrum of zero reflectance at all wavelengths should be also used as a proxy for shade (Sabol *et al.* 1992). Another alternative for calculating shade should be to use radiative transfer calculations to estimate the amount of light transmitted through the canopy or scattered onto the background from the atmosphere or adjacent canopies. According to Hall *et al.* (1995), shade would be the amount of light transmitted or scattered onto the background, multiplied by the illuminated background reflectance.

6.10.1.2 The estimation of proportions

The method for proportions estimations were: the Constrained Least-Square (CLS, detailed in **section 6.4.3.1**), the Weighted Least-Square (WLS, described in **section 6.4.3.2**) and PRINCO (mentioned in **section 6.4.3**). The unconstrained version of the two first methods were: the Unconstrained Least-square (ULS) and the Unconstrained Weighted Least-Square (UWLS). **Table 6.2** presents the mean errors for each band and the total error, when performing the mentioned methods. **Figure 6.7** presents graphically each candidate to end-member.

Black-and-white and colour composites (RGB) of the fraction images were compared to the ground information (maps, aerial photographs and field data). The evaluation of each resultant sets was based on the generation of error bands (for each band), individuals and total mean error (detailed in the **section 6.6**) and by the generation of images that mapped the fraction images overflow.

Table 6.2 Mean reflectance values for each component, mean error for each band and total error for the model considering the different approaches for unmixing pixels .

Features	TM1	TM2	TM3	TM4	TM5	TM7	TOTAL ERROR
A.angustifolia	18.10	11.98	6.12	33.91	6.88	0	/ / / / /
Araucaria forest	18.10	11.98	6.88	51.00	19.89	4.08	/ / / / /
Shade	19.89	15.04	13.00	1.02	0	0	/ / / / /
	E	R	R	O	R	S	
ULS	0.970	1.598	2.972	1.232	1.735	2.967	1.948
CLS	1.749	2.964	5.757	4.914	13.873	7.991	6.208
UWLS	0.955	1.589	2.715	0.946	1.544	2.824	1.762
WLS	1.762	2.968	5.916	5.110	15.142	8.334	6.539
PRINCO	18.368	12.628	6.970	28.180	9.288	3.261	13.116
Soil	22.44	19.89	22.50	46.41	72.16	35.95	/ / / / /
	E	R	R	O	R	S	
WLS	0.902	1.461	2.495	0.848	0.934	1.054	1.282
UWLS	1.004	1.841	2.081	3.423	5.064	3.479	2.815
PRINCO	20.367	16.125	14.631	33.496	24.730	15.361	20.785

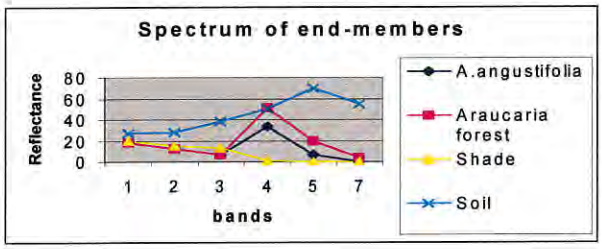


Figure 6.7 Spectra of the candidates as end-members

Comparing the computation times CLS and WLS for methods indicates that WLS takes longer, but as mentioned before, with improvements in technology it is no longer poses such a constraint. Comparing the results between ULS and UWLS (unconstrained approaches) is was noted that the UWLS method presented better results, or smaller mean error per band and smaller total error (**Table 6.2**). According to Shimabukuro *et al.* (1991) the first two methods normally present similar results, but WLS is more sensitive to the scene characteristics on the image. When the constrained methods (CLS and WLS) were compared the CLS method appeared to present slightly smaller mean errors, mainly for band 5. However, total error was less than 5% smaller than the WLS method. The visual results for the WLS method were also better when compared with ancillary data (aerial photographs and maps). More precisely, the fraction images obtained by the CLS method appeared more affected by the topographic effect and by the moisture content of the soil (**Appendix 6.1**). Nevertheless, the constrained approaches do not allow performance of the one of the next steps in the analysis, i.e., visual mapping of the overflows. However, for classification purposes eventually it might suit study goals quite well. PRINCO presented the highest errors and the resultant images were inferior to that produced by the other methods.

Overall, the unconstrained form of the WLS method (UWLS) performed better, presenting smaller errors (per band and total errors) and was also visually superior. Images error for each band and image overflow for each component were then generated. In this experiment, two types of vegetation were included in the model. The fraction image of all scene representing the vegetation type 1 – Vt1 (*Araucaria angustifolia*), the fraction image of the Vt1 focusing on the National Forest (INF), the fraction image (for INF) provided by the vegetation type 2 – Vt2 (*Araucaria* forest), and the shade image, are presented in the **Figure 6.9**.

6.10.1.3 The resultant images

In the fraction images, the abundance of each component is represented from dark grey (low amount) to light grey (high amount). It is demonstrated that differentiation between vegetated areas and non-vegetated areas is visually possible and practically delineates the boundaries of the National Forest, which is of course much more preserved than the neighbourhood. The vegetation fraction image is supposed to provide information about the amount of green vegetation within the

sensor resolution element. The fraction image Vt1 (**Figure 6.9a** and **b**) was obtained by searching areas of *A.angustifolia* in plantations as the "pure" pixel. This means that this species should present the brighter areas in the image. It presented interesting results directly related to the objective of this research. *Araucaria angustifolia* (the focus of the image) plantations (labeled as 3, **Figure 6.8**) can be visually separated from *Araucaria* forest (labeled as 4, **Figure 6.8**), because of the tendency of differentiation between conifers (lighter) and deciduous species (darker). Because the planted stands are not currently being commercially managed, (operations like thinning), some areas of *A.angustifolia* plantation presented dark and light tones in the same stand, implying mixed areas. Concentrations (amount) of the species *A.angustifolia* (labeled as 3, in **Figure 6.8**) might be detected in the middle of the deciduous species of *Araucaria* forest (natural area and labeled as 4, **Figure 6.8**). It was noted that lighter tones were characteristic of *A.angustifolia* as a specie (labeled as 3, **Figure 6.8**) and *P.elliottii* (labeled as 1, **Figure 6.8**). They are both plantations of conifers and the similarities in aspect were expected. However, *P.taeda* (labeled as 2, **Figure 6.8**), which is also a conifer presented aspect (darker tone) some times similar to that present by *Araucaria* forest (mixed areas of dark and bright pixels). An analysis in the field database showed that the differences in volume between pine species can be considered insignificant, or that volume for *P.taeda* is within the limits of volume for *P.elliottii* (**Table 4.1.2-Appendix 4.1)**. This means that the differences in tone are not related to biomass. There are very few young *P.elliottii* stands at the INF. Apparently, for *P.elliottii*, the younger the stands, the brighter it will appear in Vt1 fraction image. The Vt1 image, consequently works wonderfully to capture the presence of *P.elliottii* and the amount of *A.angustifolia* (in plantations and in the natural areas), but not necessarily will differentiate *P.taeda* from *A.angustifolia*, when the purpose is forest type classification.

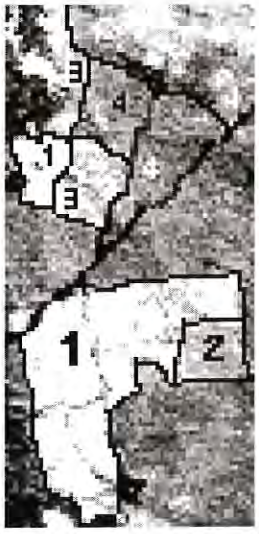
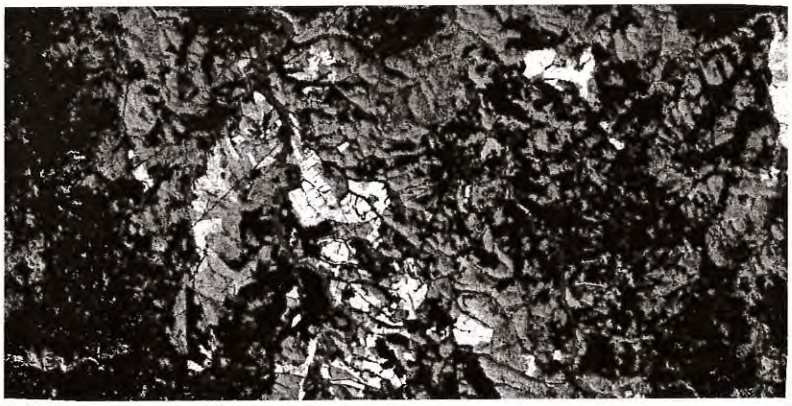
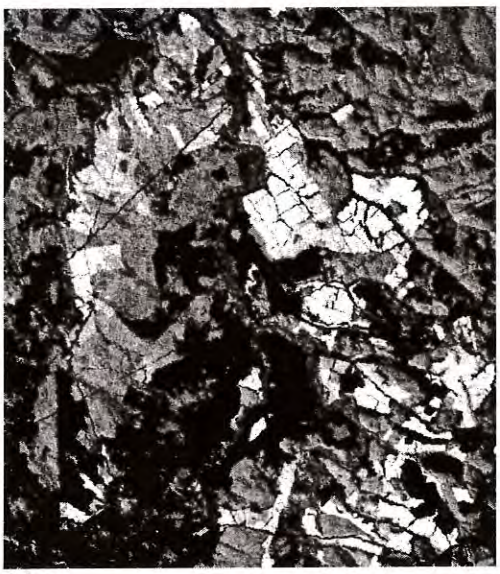


Figure 6.8 Aspect of areas with *P.elliottii* (labeled as 1), *P.taeda* (labeled as 2), *Araucaria angustifolia* plantation (labeled as 3) and *Araucaria* forest (labeled as 4) in the Vt1fraction image



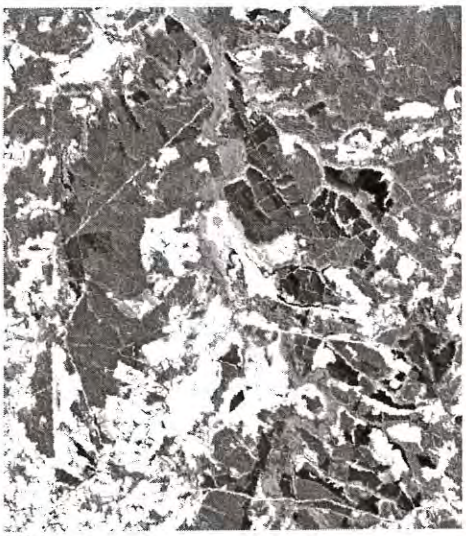
(a)



(b)



(c)



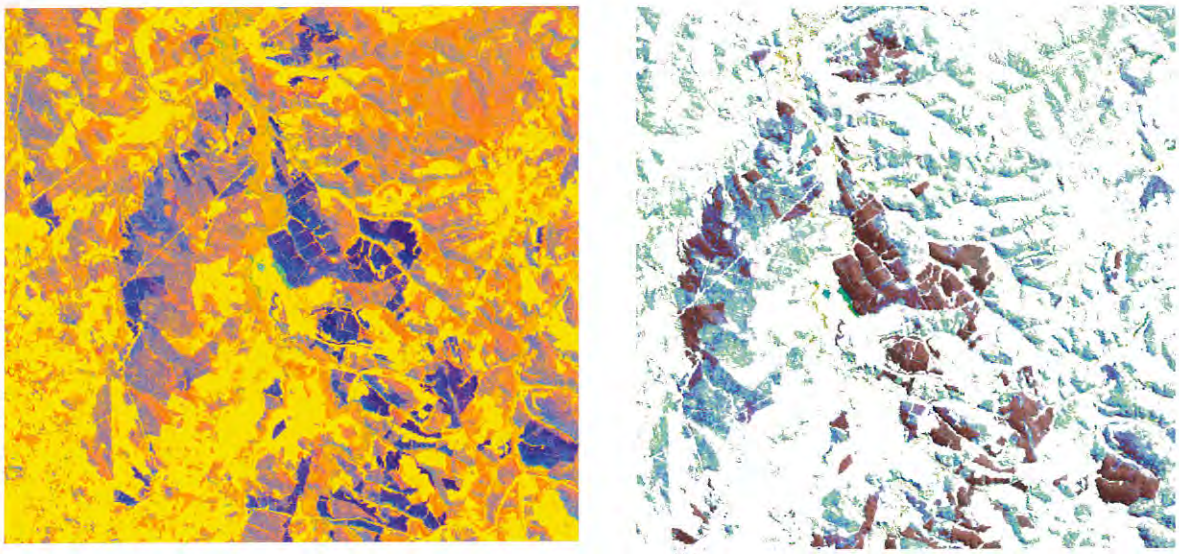
(d)

Figure 6.9. **(a)** Fraction image of Vt1 – Component *Araucaria angustifolia* - INF and adjacencies; **(b)** Fraction image of Vt1 – Component *Araucaria angustifolia* focusing the INF; **(c)** Fraction image of Vt2 – Component *Araucaria* forest; and **(d)** Fraction image of the Component Shade obtained using the Unconstrained Weighted Least-Square method (WLS).

The second Vegetation fraction image (Vt2, **Figure 6.9c**) was obtained by searching the "pure" pixel in areas of *Araucaria* forest. Consequently, it was expected that evergreen areas should be highlighted. In the case, this means all vegetation (including crops and pasture) but conifers. In fact, the image highlighted non-forest areas (very bright) and showed differences between conifers (mainly *A.angustifolia* and *P.elliottii*) from deciduous, but presented the same level of confusion between *Araucaria* forest and *P.taeda* that was found at the image Vt1. The image was affected by topographic effects.

In the third fraction image (**Figure 6.9d**), the shade image, *Pinus elliottii* exhibits darker tone than the other forest types, which means a higher amount of vegetation proportion than other species and, consequently, more shadow (texture). There is tendency of differentiation related to age between the pine stands, which presents different intensities of dark tones. The younger the stand, the darker it is. This was the first of the three images when *P.taeda* presented a different tone from the other conifers and from *Araucaria* forest. Pines are architecturally built as cones, which provide major shadow effect. *P.taeda* has branches in angle more close to the right angle, when compared to *P.elliottii. Araucaria angustifolia* is a conifer and it is the dominant species, but presents a round canopy. Its crown closure (the areal coverage of tree crowns as viewed from above; Campbell 1996) might at least act differently when compared with pines. Shimabukuro *et al.* (1995) used the technique in an area of Southern Brazil and found that the shade fraction images were shown to be sensitive to forest structure caused by age differences for eucalyptus plantation (*Eucalyptus alba* and *Eucalyptus saligna)* and by pine species (considered as *Pinus elliottii* and other pines for the pine forest). In summary, both Vt1 (with limitations) and shade can play an important job in the classification process and the Vt1 and Vt2 certainly can help in the discrimination of the species *Araucaria angustifolia* within the natural stands of *Araucaria* forest. Colour composites were generated to enhance differences between forest types, ages, and the presence of *A.angustifolia* in the natural stands of *Araucaria* forest (**Figure 6.10**).

The first composite (**Figure 6.10a**) is a RGB composition of the three fraction images, where *Pinus elliottii* is intense blue, *Pinus taeda* is also blue, but slightly less intense, *Araucaria angustifolia* plantation presents a tone close to *Pinus taeda* and the deciduous and other forest vegetation (sub-tropical forest patches and riparian areas) are light brown. Finally, non-vegetated areas are yellow. The second colour composite (**Figure 6.10b**) is a HIS composition, which is particularly useful in geological applications (Mather 1999) and was supposed to help in the discrimination of paraná-pine. The whole image enhances the topographic effects, which in this particular case is desired because these effects could be responsible for confusion between the presence of shadow due to solar angle and areas with more concentration of coniferous. The image Vt2 is actually more affected by the relief than Vt1 or shade. In this image, conifers are dark brown. The most highly concentrated areas of *A.angustifolia* plantation or the most concentrated clusters of the species in the natural stands presented the same dark brown tone.



(a) (b)

Figure 6.10(a) RGB colour composite of the faction images, where Vt1(R), Shade (G), Vt2 (B) and **(b)** HIS composition of the same images, where Shade (H), Vt2 (I), Vt1(S)

6.10.1.4 The evaluation of the model fitness

The overall fit of the model should be judged adequate if band errors and/or mean error were low. As mentioned by Adams *et al.* (1995), pixels having high mean error indicate unmodeled compositional variability in the scene. The calculation of the mean error and generation of error images were performed as mentioned in the **section 6.6.** Additionally, based on the preoccupation that overflow images should be one more way to decide about the fitness of the model, a strategy for mapping those elements was developed. It consisted of a program in LEGAL language to visualise the values in three classes. This was chosen because, according to NETGIS (1998), SPRING, for the unconstrained models, allocates pixels with negative values to receive value zero (0) and values bigger than one (1) to receive the value 255. The intermediate values were scaled over the range 1 to 254. The produced overflow images (OI's) are presented in **Figure 6.11**.

Overflow image of Vt1 – The pure pixel was searched in *A.angustifolia* plantations. Results to be mentioned are: a) non-vegetated areas (specifically non-forested areas) were considered under-shoots (red - as expected); b) the stands of *A.angustifolia* (plantations) were included in the legitimate range (dark-yellow, being the only areas in this class, what demonstrates the fitness of the model). However, some highly concentrated *A.angustifolia* pixels were included in over-shoots (blue) group; *P.elliottii* (with emphasis the young stands) plantations were included in the same class. A third group of pixels that belonged to *A.angustifolia* plantations was classified together with areas of *Araucaria* forest (mid-yellow); *P.taeda* was also included in this group; d) light-yellow represents areas that are only deciduous forest. It was considered that the image was modelled quite well, indicating correctly areas of *A.angustifolia* plantation.

6.10.1.4.1 The overflow images (OI's)

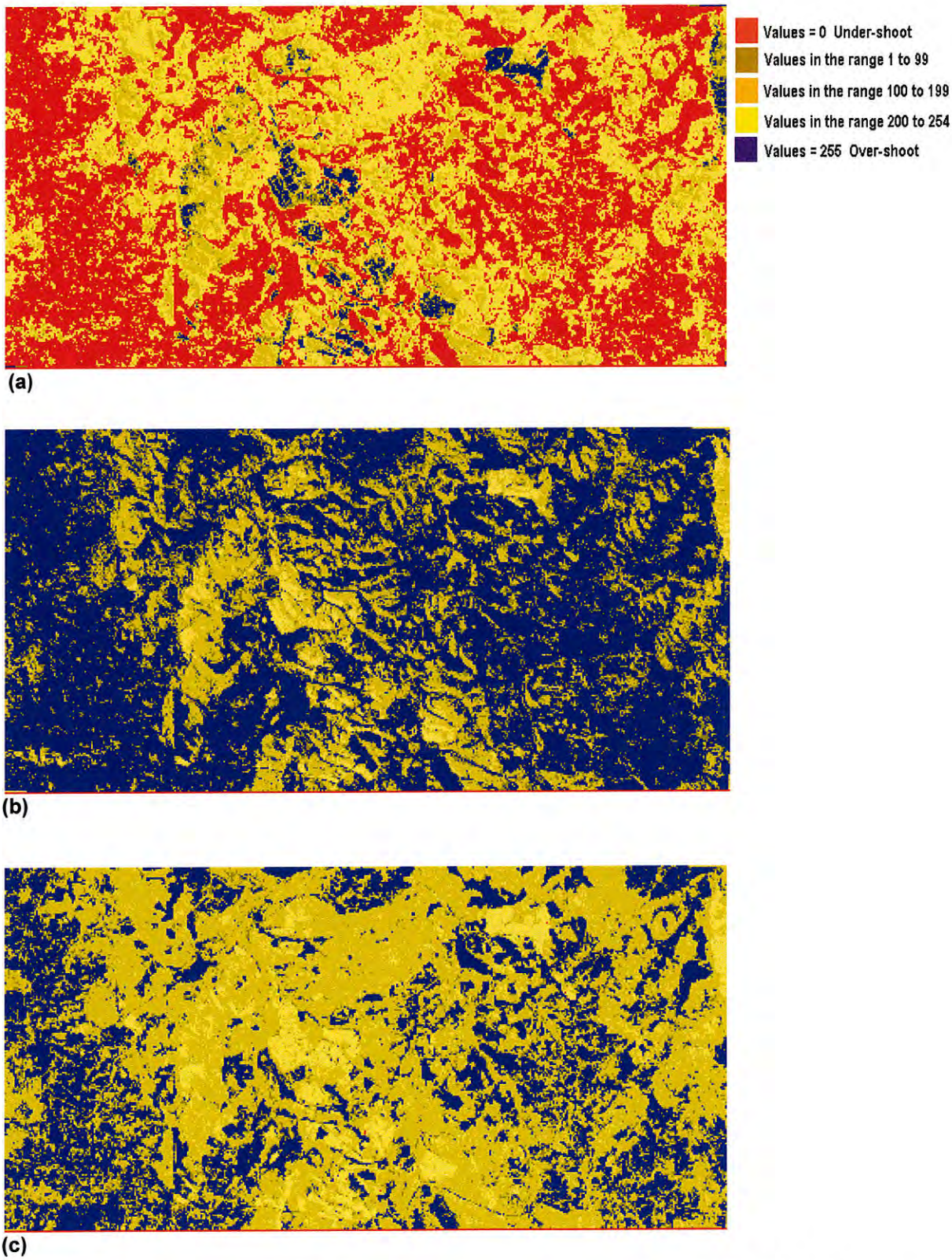


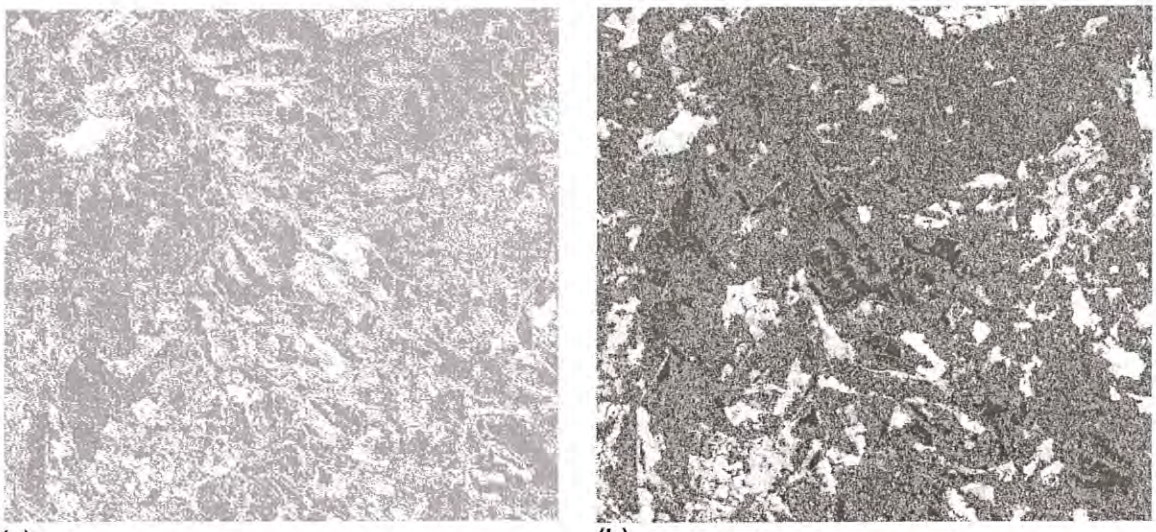
Figure 6.11 Overflow images **(a)** OI of Vt1 – Component *Araucaria angustifolia*; **(b)** OI of Vt2 – Component *Araucaria* forest; and **(c)** OI of the Component Shade .

Overflow image of Vt2 – The pure pixel for this image was searched in areas of *Araucaria* forest. Results to be mentioned: a) There is no dark areas in the image and, consequently, no under-shoots; b) the non-forested areas were classified as over-shoots; b) considering the legitimate range: dark-yellow as more related to dark areas due to topographic effects; as mentioned before the image which this overflow panel was originated was the one most influenced by topography, probably because of the influence of TM band 4. Relief actually reflected in the classification of *P.elliottii*. Some stands, which are homogeneous, appeared half mid-yellow, half light-yellow. *P.taeda* was included in the same class as *Araucaria* forest. In general, deciduous species (including *Araucaria* forest) tend to be mid-yellow and conifers (including *A.angustifolia* plantations) appeared as light-yellow. The image confirms the fitness of the model, as identifies *Araucaria* forest, which was the primordial objective of this end-member.

Overflow image of shade – The pure pixel for this image was searched in order to capture differences in scene spectral responses, according with the different canopies architecture and vegetation size. The analysis of its OI considered that the pure pixel was searched in very dark regions (deep water) and, consequently, it did not included under-shoots. Results to be commented: a) Non-forested areas (urban, crops, pastures, were included in the over-shoot class), what means a good result; b) both pines (that were discriminated quite well) and *A.angustifolia* plantations (when highly concentrated) were classified in the legitimate range (light-yellow). However, some areas with *A.angustifolia* were included in the same area as *Araucaria* forest (mid-yellow); c) dark-yellow appeared very sparsely and probably is connected to small areas with high moisture content; d) some spots of light-yellow (that might be coniferous) can be noticed in the middle of the patches of *Araucaria forest* (mid-yellow). The image completed the analysis about the usefulness of the model and it was considered very helpful in the whole process of deciding the corrected discrimination of the different land covers and, specifically the four forest types studied in the present work.

6.10.1.4.2 The error images (EI's)

The error images were generated as described in the **section 6.6**. The error images show the differences between the observed pixel value and the value computed from the model. As the objective of the error calculation is to help in the analysis of the model adequacy to a certain scene, the inspection of all the images is fundamental in the process. Ideally the errors should not present a pattern in the interest areas, but certainly can be high for the features that are not the selected components. In this section, the error images for TM4 and TM7 are printed (**Figure 6.12**), as the first one represents the smallest error for all bands (0.946) and TM7 represented the biggest value of error (2.824).



(a) (b)

Figure 6.12 Error images (EI's): **(a)** EI of TM4 and **(b)** EI of TM7. The images have been subjected to a histogram equalisation contrast stretch.

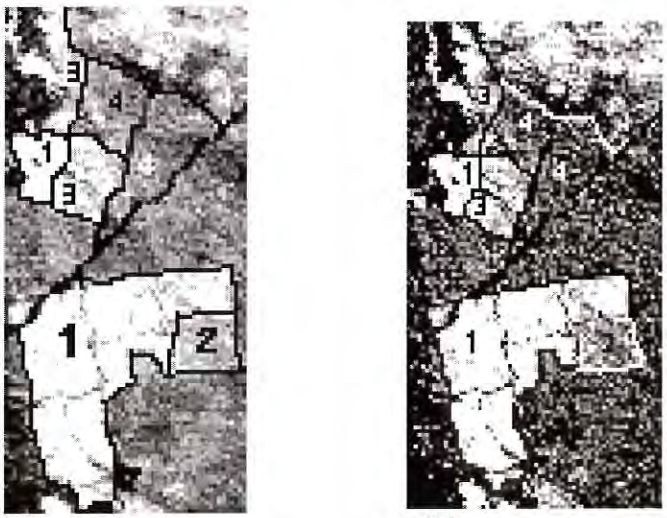
For interpretation these images, the darken the areas, the better the model performed for that component in that band. In the error image of TM4, forest in general presented dark tones but the very young pine that is located in the adjacencies of the INF. The biggest errors were out of the selected components. The area in the upper right corner is a mix of agriculture areas, very small patches of woody vegetation and bare soil. When comparing the EI TM4 to the EI of TM7 it is possible to see that non-forested areas formed a clear pattern in the second image, showing the fitness of the model to the forest features in that band. Coniferous and *Araucaria* forest were well modelled. Because the magnitude of errors is very small and the produced images are very dark in tone, they were subjected to a histogram equalisation contrast stretch.

6.10.2 Experiment 2: the inclusion of the end-member soil in the model

The analysis of PCA eigenvectors in Experiment 1 determined that the number of end-members should be equal to three. The feature under interest is composed by two vegetation types, which means that just one component more should be added and, after many previous analyses, this third element was chosen to be shade. Shadow was preferred instead of soil because the forest stands in the study area are dense, with high percentage of crown closure and less influence of that component. It was considered that, for arid zones or the ones where vegetation is sparse, soil should interfere more consistently than in densely vegetated areas. However, soil is usually an important component in the mixing models and it was decided to include it in the model as a second experiment.

The purest pixel was located using the same reasoning as used for the others end-members. First of all, the mean spectral signature for soil was established by taking the pixel value in various points in the images which were defined by the field work, by the analysis of aerial photographs and scatter diagrams from **Figures 6.5** and **6.6**. Sequentially, the signature was built according to the analysed pattern. Only the method WLS was tested because the method CLS was not developed for the presence of a forth end-members in the model, when using software SPRING. As can be seen in **Table 6.2**, the errors over the bands (and the overall errors) are smaller when the model is composed by the four elements, in both the constrained and in the unconstrained approaches. **Figure 6.14** presents images for the four end-members when modelled together.

The image for Vt1 (**Figure 6.13a**) was obtained by the selection of the pure pixel in the areas of *A.angustifolia* plantations. Results to be mentioned: a) the image discriminates coniferous forest quite well, especially both of pine species. *P.elliottii* (numbered as 1 in **Figure 6.13b**) has a lighter tone than *P.taeda* (numbered as 2 in **Figure 6.13b)**. As presented in the **section 4.4.1**, there is a negative correlation between volume and the band TM5. When analysing the different textures for *P.elliottii*, it was noticed that stands with the highest volume and basal area per hectare have a tendency of appearing lighter and with a more homogeneous texture. *A.angustifolia* plantation (numbered as 3 in **Figure 6.13b)** presented different light tones, according to the density of the stands. The same tendency as pine was noticed, the highest the volume (**section 4.4.2**), the brightest the stand will appear in this image. The vegetation type *Araucaria* forest (numbered as 4 in **Figure 6.13b)** was affected by the inclusion of the component soil. When it is compared with the fraction image for the first component obtained by the three end-members, it turns out that the areas with natural forest are much more rough now than in the Vt1 image obtained by three end-members. Many light dots appear where in the other image it is more uniform. As the light dots as been well connected to conifers, the assumption that they might represent them is possible.



(a) (b)

Figure 6.13(a) Aspect of areas with *P.elliottii* (labeled as 1), *P.taeda* (labeled as 2), *Araucaria angustifolia* plantation (labeled as 3) and *Araucaria* forest (labeled as 4) in the Vt1fraction image for experiment 1 (three end-members); (**b**) same identification for Vt1 fraction image in experiment 2 (four end-members)

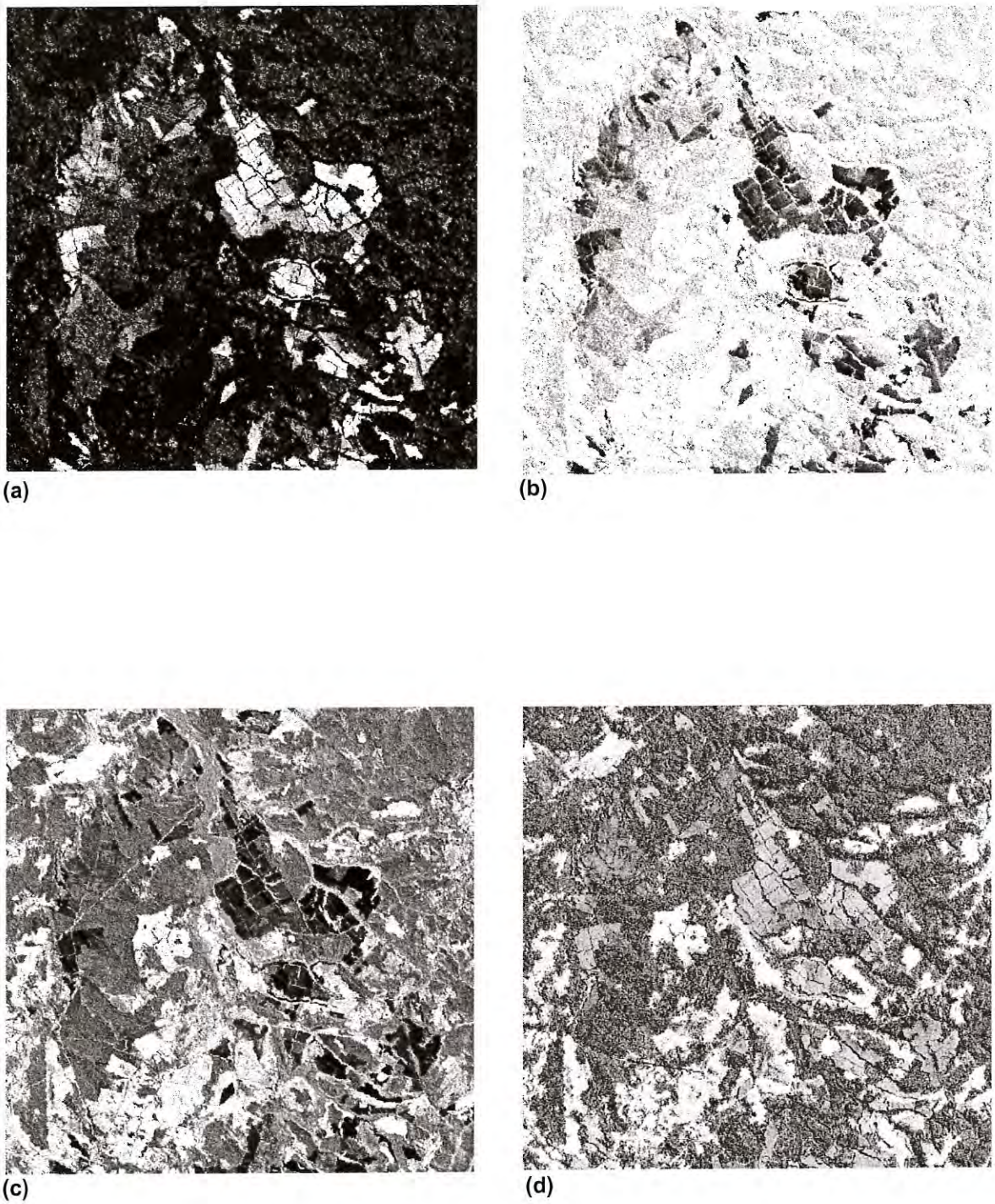


Figure 6.14(a) Fraction image of Vt1 – Component *Araucaria angustifolia*; **(b)** Fraction image of Vt2 – Component *Araucaria* forest; and **(c)** Fraction image of the Component Shade; and **(d)** Fraction image of the Component Soil obtained using the Unconstrained Weighted Least-Square method (WLS).

The image for Vt2 (**Figure 6.14b**) presented similar results to the obtained by Experiment 1 for pines and *A.angustifolia* plantation. However, the inclusion of the soil component in the model did affect the whole appearance of the deciduous areas. In terms of texture, it is rougher, showing the natural non-uniformity (tone and texture) of the patches with natural forest. The image was

affected by topographical conditions, as a result of the contribution from TM band 4, but this effect was minor than at its pair of Experiment 1. Even with this scenario, it seems that some sensitivity to the presence of conifers in natural areas was achieved. A small window of the image obtained by both approaches (three and four end-members) is presented in **Figure 6.15a** and **b** showing a zoomed view of a specific region where the four main forest types are present and the differences in tone for *P.elliottii* (numbered 1) and texture for *Araucaria* forest (numbered as 4) can be noted. *P.taeda* and *A.angustifolia* plantations seems to present the same amount of green vegetation in some areas.



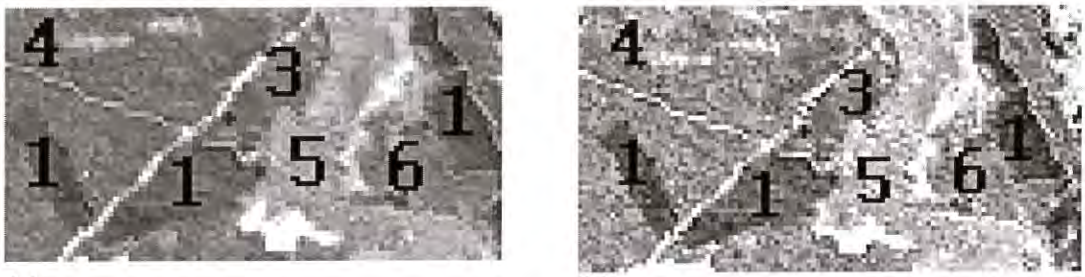
(a)



(b)

Figure 6.15(a) Aspect of areas with *P.elliottii* (labeled as 1), *P.taeda* (labeled as 2), *Araucaria angustifolia* plantation (labeled as 3) and *Araucaria* forest (labeled as 4) in the Vt2 fraction image - experiment 1 (three end-members); (**b**) same identification for Vt2 fraction image - experiment 2 (four end-members)

The biggest (visual) difference between the shade fraction image (**Figure 6.14c**) in experiment 2 and its predecessor obtained in experiment 1 (**Figure 6.9 c**) is the roughness of the second one, which is concentrated on non-conifers areas. Regarding the second image, apparently, there is no gain in information for the features that are the objective of this work. However, small and sparse areas that were characterised as non-vegetated in experiment 1 appear now as vegetated. It seems that the presence of vegetation like crops or pasture was captured with more details in the second experiment. Anyway, the vegetated and non-vegetated (streams and floodplain areas, for example), forested or not are well defined (**Figure 6.16**). Pines are differentiated from each other and young pines presented darker tone.



(a) (b)

Figure 6.16(a) Aspect of areas with *P.elliottii* (labeled as 1); *Araucaria angustifolia* plantation (labeled as 3); *Araucaria* forest (labeled as 4); stream (labeled as 5); and floodplain (labeled as 6) in the shade fraction image - experiment 1 (three end-members); (**b)** same identification for shade fraction image - experiment 2 (four end-members)

Finally, in the soil fraction image (**Figure 6.14d**) very few areas were highlighted as non-vegetated (soil). These areas coincide with the cities of Irati, T.Soares and F.Pinheiros. Conifers were slightly differentiated from each other but no differences in age were captured. *A.angustifolia* plantations and *Araucaria* forest presented different pattern. The image brings no extra information that should be mentioned as important for the purpose of this work. **Figure 6.17** present the Overflow images for the four components, focusing in the INF.

Comparing results from experiment 1 and experiment 2, both OI's for Vt1 are similar in terms of overflow. However, more over-shoots were verified for Vt2, when it is compared with its pair in experiment 1, probably to the presence of soil in some pixels which were not supposed to be detected in the first analysis. The OI for the component shade is very similar to its pair in Experiment1 in terms of over-shoots, but now it presents under-shoots in the region where young stands of pine are installed. The OI for soil did no present any values out of the legitimate range.

The analysis of the error images can be summarised by the fact that this time the bigger error was related to TM3 (2.495) and the smaller one related to TM4. The images for these two bands are presented in **Figure 6.18**. The error images for TM4, TM5 and TM7 are very similar and lead to the conclusion about the good fitness of the model to the selected components.

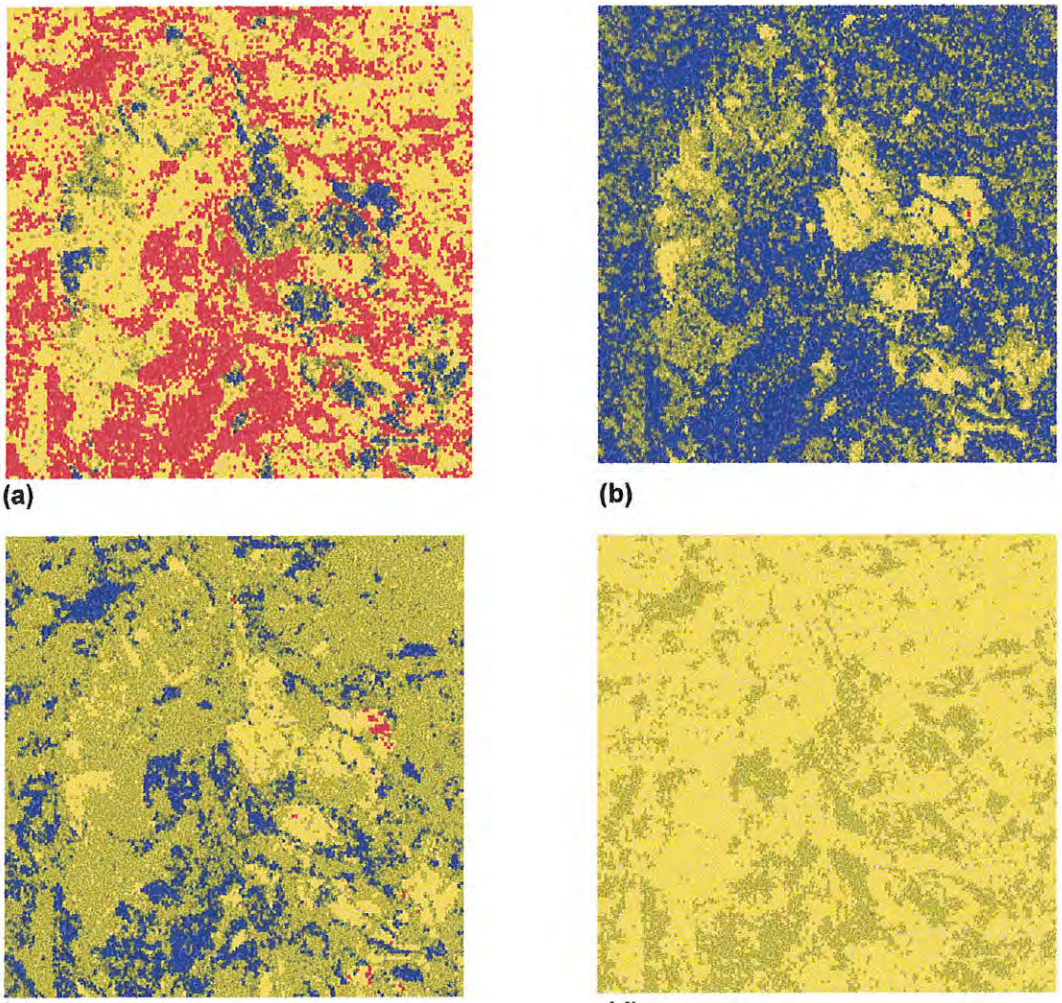


Figure 6.17 Overflow images **(a)** OI of Vt1 – Component *Araucaria angustifolia*; **(b)** OI of Vt2 – Component *Araucaria* forest; **(c)** OI of the Component Shade; and **(d)** OI of the Component Soil.

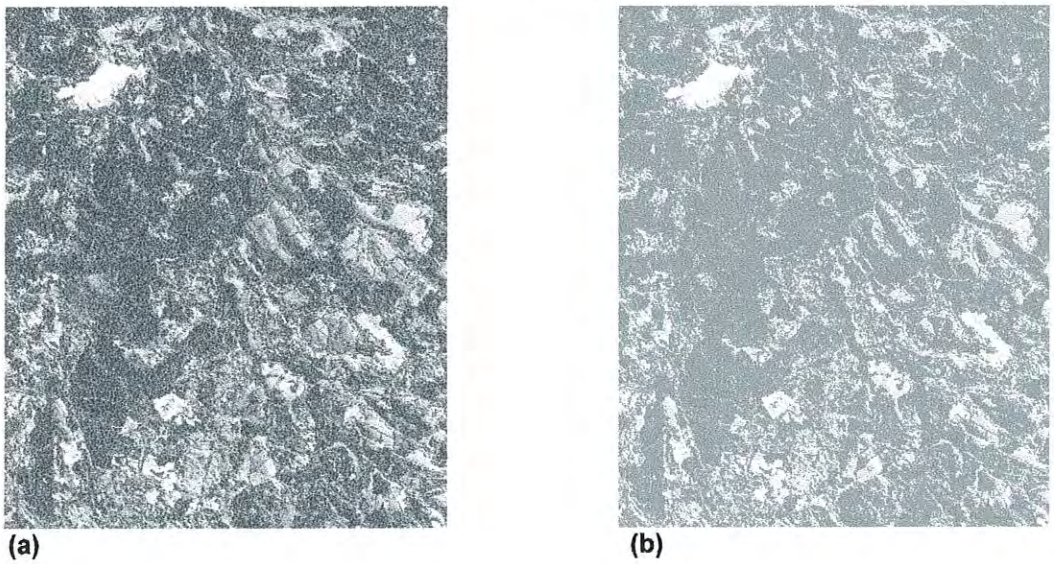


Figure 6.18 Error images (EI's): (a) EI of TM3 and (b) EI of TM4. The images have been subjected to a histogram equalisation contrast stretch.

6.11 Conclusions

Multispectral images have spatial and spectral properties that can be used to detect a specific target. Spatial properties such as shape and size of an object can be included in the process, but the only way to detect a target only by its spatial properties is if it is large enough to be captured at a specific image resolution. If the object is not represented by a group of pixels, it will not be detected. Under certain conditions, however, an object can be detected by its spectral properties (or by a combination of spectral and spatial properties), even if it cannot be detected by its spatial properties. The process is called *spectral mixture modelling* (SMM) and it can be used to detect and in some instances identify targets at the sub-pixel scale. According to Aguiar *et al.* (1999), the concept of synthetic bands obtained by procedures as the linear mixing model, involves the idea that classes are not described as a function of their spectral responses (spectral signatures), but in terms of a physical characteristic (primary components proportions). This means that visual analysis should be facilitated, since physical concepts (proportions) are more easily understood than class spectral signatures.

Broadly, the approaches can be divided into: a) traditional methods for detecting differences between targets (this topic includes procedures analogous to classification); and b) the analysis which hypothesises that variations among the spectral class signatures are mainly caused by structural target characteristics at the sub-pixel level (Aguiar *et al.* 1999). The second approach is the one behind the spectral mixture modelling.

In the present study, the objective for using the technique was to verify if the generated fraction images are capable of discriminating between the different forest types present in the area, with emphasis on *A.angustifolia*.

The number of complete and separable components (end-members) is crucial to the success of the process. In reality, there are two major limitations in the process, which have to be defined before the application of any of the numerous approaches available to unmix the mixture. The first problem is related to the number of end-members and the second one is where and how to capture the purest pixel to be included in the model. Trying to solve the first problem, PCA has being recommended to help defining how many end-members might be considered. The second question is multifaceted since some of the approaches which are being discussed nowadays are not based on the pure pixel approach, take it automatically, or use a mixed process between manual and automatic selection.

The basic principle of the model is that forested areas could be analysed as formed by three main components, nominated as canopy, soil, and shadow. Using the PCA approach three end-members were considered ideal in this study and a fourth one (soil) was also tested. The feature under interest occurs in natural stands together with almost a hundred other species. Because *Araucaria angustifolia* is the only conifer non-exotic vegetating naturally in the studied area the work biome, the study actually focused on the discrimination between coniferous and deciduous woody species. Consequently, the first two end-members had to be the *A.angustifolia* in its pure form (pixel captured in the densest stands) and the signature of the natural *Araucaria* forest. The third element happened to be shade, because of its documented importance in these kind of studies and because the architecture of the features occurring in the area. Plantations of the exotic coniferous *Pinus elliottii* and *Pinus taeda* are vegetating in the same region and are likely to be confused with *A.angustifolia* (plantation) unless the differences in response pattern and canopy architecture could be captured. The process adopted in the present study uses the signature of each end-member (studied in **Chapter 4**) for the definition of the values to be considered as pure pixels.

The number of different approaches to the unmixing problem is extensive. The present project focused in the Least-Squares mixing model, to generate fraction images from TM data. The Constrained Least-Square (CLS), the Weighted Least-Square (WLS) and PRINCO, a combination between PCA and least-squares. For the first two approaches, unconstrained versions were included in the analysis. The evaluation of the fitness of each method was made by the analysis of visual performance of the fraction images, band errors, image errors, and image overflows.

The unconstrained Weighted Least-Square (WLS) method was chosen by the criteria mentioned before. The unconstrained method was preferred to the constrained approach, not only because of its visual performance, but also because it could allow the derivation of overflow images to evaluate the fitness of the model. The inclusion of the end-member soil diminished the error bands but presente more overflows and did not improve the quality of the images at least for classification purposes.

It was possible to detect some clusters of conifers mixed in the natural stands of *Araucaria* forest, which was considered to be *A.angustifolia*. The other species whose signature is very close is *Pinus taeda*, which being an exotic species cannot be in the natural stands. The presence of paraná-pine was confirmed by the analysis of the aerial photographs. Differences in amount of vegetation in *Araucaria* forest were easier to be detected in the first two images (Vt1 and Vt2) of experiment 1, when compared with their pairs in experiment 2. The methodology developed in this thesis to derive the overflow images was considered very useful in the analysis and it is recommended that the procedure be included in the software SPRING.

The fraction images derived from Linear Mixing Model are also presented in **Chapter 7**. These images are useful in the process of change detection and can be utilised to generate difference images with the objective to detect the green biomass variation in the study area. These images also can be included as additional channels for image analysis in the segmentation process, prior step of image classification.

Chapter 7

CHANGE DETECTION IN *ARAUCARIA* FOREST

7.1 Introduction

Forest cover monitoring is a means of obtaining information over time. It is a broad term, including several different approaches. Before the definition and development of a methodology to achieve any goal, it is important to define and specify which information is required and at what level. The identification of areas where monitoring is most required is a crucial step. As mentioned in **Chapters 1** and **2**, natural forests in Southern Brazil (*Araucaria* forest and Atlantic forest) need to be monitored since forest loss though human exploration is significant.

In this context, the objectives of this chapter are: a) to determine the magnitude of any changes in *Araucaria* forest patches in the period from 1987 to 1997 within and in areas adjacent to INF; and b) to develop a method by which future changes in *Araucaria* forest can be monitored.

The increasing threats to forests in general stress the need for conducting repetitive forest inventories as a basis for long term management of sustained yield. However, multitemporal forest monitoring only by conventional methods is often impractical. In the scenario of changes in land use and declining forest cover caused by an expanding human population, remote sensing can contribute as an important tool for the monitoring of deforestation, providing essential tools to foresters. Because of the nature and characteristics of satellite imagery, repetitive surveys are best performed by remote sensing methods, which can present good results when connected to field inventories. The availability of archived remote sensing data (aerial photographs and satellite images) is essential to the assessment and mapping of forested areas changes. Assessing and mapping changes, as well as determining their pattern and rate is fundamental for understanding their causes and for proposing strategies for protective or mitigating measures.

This work based the detection of changes in forested areas on the image differencing method. The first section tested approaches involving TM single band and vegetation indices for detecting positive and negative changes. The number of available indices and the possibility of combination between bands and their mathematical variants is very large and it is virtually impossible to analyse all the possible combinations. The strategy adopted in the study was to select band 5 (middle infrared region, 1.55 to

1.75 μm) and NDVI to perform the analysis. Band 5 was chosen because it provided the better discrimination of the forest types analysed in this work, when compared with the other bands, as characterised in **Chapter 4**. NDVI is certainly among the most used approaches for change detection studies in densely vegetated areas. In this work, the NDVI represented the family of slope-based VI's. Distance-based VI's (derivatives of the Perpendicular Vegetation Index – PVI) were not included in the study. They should be more useful in areas with sparse vegetation. Therefore, the indices included in preliminary tests performed poorly when compared to NDVI (**Appendix 5.1**).

A second approach was to test the classification of the product of image differentiation (TM band 5) in the process of categorising losses of vegetation. Supervised classification was selected for this purpose. In the third approach, image differentiation was tested over the fraction images derived using the Linear Mixing Model, as described in **Chapter 6**. Both supervised and unsupervised classifications were tentatively applied. The aim was to develop and suggest an alternative methodology to detect changes in *Araucaria* forest. The vegetation and shade fraction images were considered in the process. The shade fraction image has been suggested (Shimabukuro 1995; 1998; Aguiar *et al.* 1999) as very promising in densely vegetated areas such as the Amazon Region.

As stressed in **Chapter 3**, every effort must be made in the pre-processing and processing phases to guarantee that the data sets are comparable. Mis-registered images might result in a difference image with changes that result from problems in the original data sets. Differences in atmospheric conditions and illumination angle have to be considered in the decision process about which methods should be adopted, as previous steps or as a part of the change detection procedure.

7.2. The impact of mis-registration on change detection

Almost all published texts that involve any kind of change detection studies remind the readers about the importance of being sure that the detected changes are based on reliable sources. This importance is increased because of the current interest in the nature and impact of global climate change, which demand long-term data sets so that changes can be reliably documented (Townshend *et al.* 1992). Chavez *et al.* (1994) considered the problem and developed a procedure, which includes multitemporal image calibration using a hybrid method prior to application of any other change detection approach. According to Milne (1988) two sets of images must be registered within an accuracy of one pixel or less. Lillesand and Kiefer (1994) recommended registration to within ¼ to ½ pixel. The accuracy is critical in ensuring that the results of the adopted method are not an artifact of the image processing steps instead of real changes (Mouat *et al.* 1993).

This work used two sets of TM imagery (from 20/06/1987 and 18/08/1998), bands 1 to 5, and 7, as detailed in **section 3.4.1.2**. Images were rectified and georeferenced to the UTM projection and

resampled using a bilinear interpolator. The study presented a calculated RMS of 0.852 pixels (approximately 25.56 m). In the geometric co-registration, image 1987 was used as the reference. The resultant RMS was smaller than one pixel (0.486 or 14.58 m approximately). This accuracy is directly related to the accuracy of the whole process of change detection (Igbokwe 1999). In the radiometric corrections, DN were converted to reflectance (and multiplied by 255 to keep the same amplitude of DNs from raw images). Atmospheric effects were also considered a great problem in satellite imagery calibration as they affect the temporal sets differently. One of the two following approaches are normally considered, when multitemporal analysis are performed: a) the change detection procedure includes multitemporal calibration as a part of the method (Chavez *et al.* 1994; Adams *et al.* 1995); or b) previous calibration and posterior application of the change detection method. The present study adopted the second approach and the subject was considered in **section 3.4.1.3.2.2a.**

7.3 Detecting positive and negative changes of biomass

Direct comparison methods for change detection such as image differencing and ratio using the original bands or vegetation indices were included in this topic. Subtraction and ratio (or division) enhances spectral differences. Ratio theoretically removes (or diminishes) topographic effects and enhances differences between soil and vegetation. Most of the developed indices are based on the interactions between vegetation and electromagnetic energy in the red (band 3, in Landsat TM) and near-infrared (band 4) wavelengths. In this study, the reflectance values from two dates (1987 and 1997) were used to analyse where changes in forest areas have occurred. According to Morisette *et al.* (1999) different change-detection algorithms can be considered different models. Change is the output variable on one side of the equation and the reflectance values from two different times are the input variables.

Mathematically, the general form of change detection can be written as:

$$\text{Change} = f(x)$$

Where: x is a vector of reflectance values from the two images; change is a polynomial response, where each value represents a particular type of change. f will depend on the method used for change detection (algorithm).

7.3.1 Image differencing

Subtraction is a basic image transformation, which applies simple arithmetic operations to the image data. Image differencing is often used to identify changes that have occurred between images collected on different dates. According to CCRS (1999), this type of image transform can be useful for mapping changes in urban development around cities and for identifying areas where deforestation is occurring. Typically, two images which have been geometrically registered (see **section 3.4.1.2**), are used with the pixel values in one image being subtracted from the pixel values in the other.

$$\Delta p^{m}_{(i,j)} = p^{m}_{(i,j)}(t_2) - p^{m}_{(i,j)}(t_1) + c$$

where:

$\Delta p^{m}_{(i,j)}$ = difference pixel value ; $p^{m}_{(i,j)}$ = pixel value at line i, column j, and band m; t = time and c = constant

In the **section 3.4.1.3.1**, it was mentioned that, in the present work, DN were converted to reflectance and the obtained reflectance values were multiplied by 255, to present the same comprehensiveness of the original digital values. In this step, because a subtraction procedure is being performed, negative values can emerge. Assuming both images have a potential dynamic range of 0-255, the maximum negative difference is –255 (0-255) and the maximum positive difference is +255 (0+255) (Mather 1999). If the system has a fixed, 8-bit resolution, the results cannot be displayed. The solution is the adoption of some scaling factor (constant = 127, for example), which does not change the proportions, but reduces the dynamic range to 0-255. A typical example of negative values should be, for example, the case of conversion of a pine stand (average reflectance = 7.1) to non-forest areas (pasture, which average reflectance = 35.5). According to Mather (1999), some arithmetic operation is normally incorporated to the formula of operations like addition, subtraction or ratioing to reduce the dynamic range to 0-255. The author also commented about the eventual suitability of a linear contrast stretch by subtracting an offset and using a variable divisor, which values should be determined on the basis of the user's experience.

The techniques that compare two different dates are sometimes called pairwise comparisons and the procedures concerned with the analysis of multiple images are named time series analysis. The present study deals with the first category. Any pair of bands or transforms can be involved in the process. In this study, the basic overlay using the middle-infrared band (TM band 5) was initially considered.

Being the most intuitive and basic change detection technique, image differencing has been used in many studies involving vegetation. Sunar (1998) reported the use of the method in detecting land cover changes in Turkey. In Brazil, Soares *et al.* (1996) working with data from *Eucalyptus* forest plantations found image differencing very effective in the discrimination process of several deforestation and reforestation levels.

7.3.1.1 Simple image differencing - Overlay

The basic version of image differencing involves the subtraction of one band of each date and it is one of the overlay techniques. The selected bands are normally those that had better discrimination of the differences or changes in vegetation between the analysed periods of time. Band 5 has been selected for this purpose in the Amazon region (Alves *et al.* 1998a) and for monitoring changes in plantations of *Eucalyptus* sp. in Central Brazil (Soares *et al.* 1996). By visual analysis of the six available bands for

this work, the region of the middle infrared was also considered the best choice and the Difference image (Dif.image5) was generated and is presented in **Figure 7.1**.

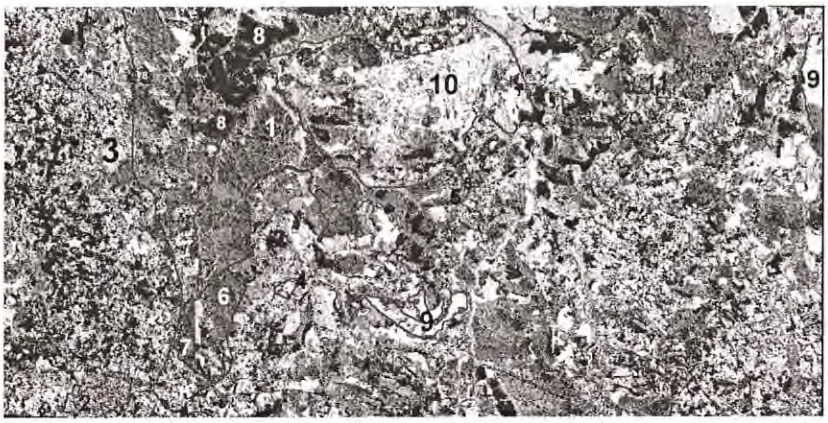


Figure 7.1 Simple image differencing image from TM band 5 (Dif.image5). The image has been subjected to a histogram equalisation contrast stretch for printing purposes.

In the next steps of this study, Dif.image5 was chosen to be a reference image and also was involved in the analysis of the additional techniques. It was expected that the additional approaches presented some improvement in the change detection procedure. The purpose of using the overlay technique as the first change detection approach of this work was: a) to test its overall capability of detecting major changes; b) to present a preliminary idea of the environmental complexity and a visual perception of the relationships between the protected area (Irati National forest - INF) and adjacencies. The INF contour is almost auto-delineated by itself, showing how altered and fragmented the area is. To mark its area, a layer with its perimeter is shown in (1). In the lower left corner is the Irati town (2), the main urban concentration in the region. The left part of the window (3) is dominated by the Irati urban and semi-urban area (including some small farms). In the same image two other smaller towns are present (Fernandes Pinheiro (4) and Teixeira Soares (5) and other small villages (11). The area with *Araucaria* forest below the INF (6) belongs to the Agronomic Institute of Paraná (IAPAR), a State of Paraná research institution. The area labelled with (7) is a private area which belongs to a brickyard (Olaria Van Laars). The areas labelled as (8), (9) and (10) will be discussed below in **section 7.3.1.3**.

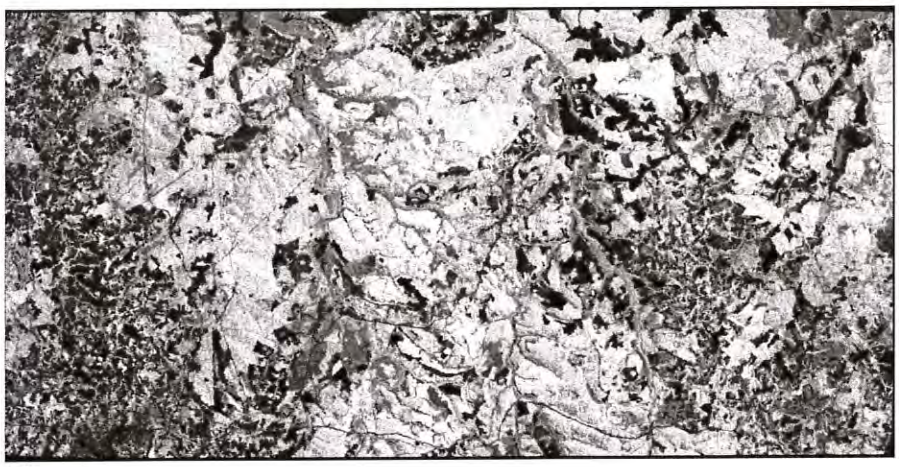
The Dif.image5 shows some changes outside the National Forest. Within the INF, however, apparently no changes were detectable by the technique, in the period of 1987 to 1997. The technique was applied to a single band, but transforms such as vegetation indices and fraction images should also be considered. The reason for this is their inherent multispectral ability; it is expected that the use of transforms should improve the change detection analysis. This is logical; the inclusion of other bands in the process can offer other dimensions to the analysis, which is the focus of the next **sections**.

7.3.1.2 Simple image differencing using NDVI

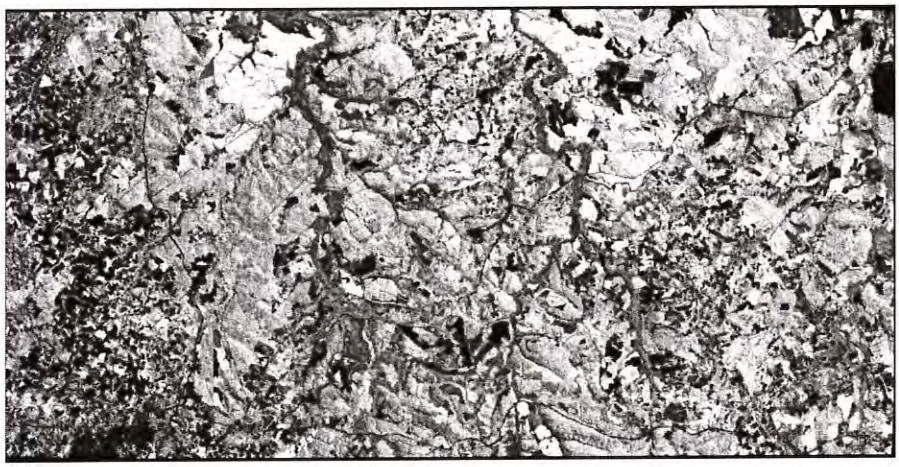
As mentioned before, there are many different vegetation indices suggested in the literature; some of them are discussed in **Chapter 5**. In change detection analysis, NDVI was selected because it had been widely applied with success. Some of its strengths are: a) it is an improvement to the use of only one band; b) it is supposed to be less affected by topographic features; c) it is supposed to normalise the differences in irradiance; d) it requires only the red and near infrared portions of the electromagnetic spectrum (which can be calculated by virtually any image processing software); e) it is a better choice for forest canopy change detection since it emphasises the vegetation spectral response differences (Lyon *et al.* 1998); and f) it has a strong relationship with green biomass (Anderson *et al.* 1993); Ponzoni *et al.* (1998). Some of its weaknesses must also be mentioned, including: a) it carries only part of the available information in the original channel reflectances (Govaerts *et al.* (1999); b) it presents saturation at relatively low vegetation biomass value (Epiphanio *et al.*, 1997).

The process of subtraction can result in positive or negative values, representing different directions of change. Zero values show areas of no change. A constant *c* is then normally added (255 is the standard scaling but 100 or 127 are also mentioned in the literature, and depends on the analyst's choice in which range he/she intends to have the data). According to CCRS (1999) scaling the resultant image by adding a constant 127 to the output values will result in a suitable "difference" image. In such an image, areas where there has been little or no change between the original images, will have resultant brightness values around 127 (mid-grey tones), while those areas where significant change has occurred will have values higher or lower than 127 - brighter or darker depending on the "direction" of change in reflectance between the two images. This suggestion was adopted in the present study. **Figure 7.2** presents: (a) the original NDVI image from 1987 (NDVI87); (b) the original image from 1997 (NDVI97); and (c) the resulting differencing image (Dif.image).

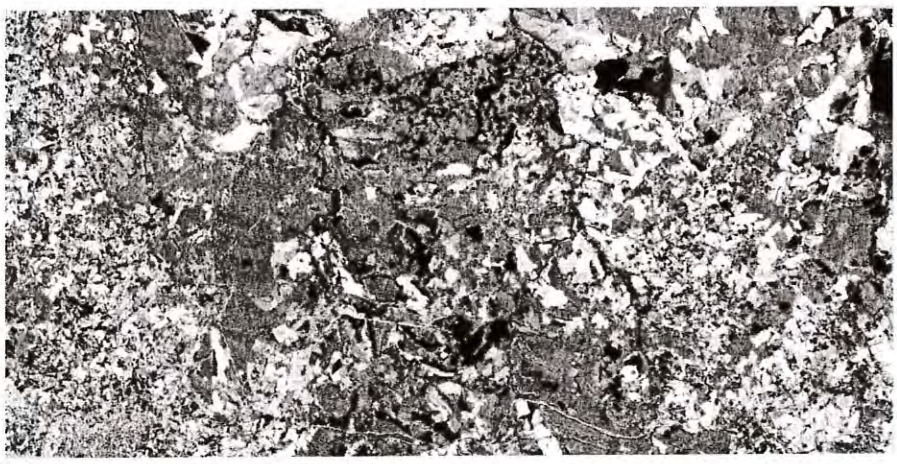
By visual analysis, it seems that the interval of 10 years resulted in an area with an environment which is more fragmented in 1997. However, the method presents a major problem, related to its capability of distinguishing between true change and natural variability (Thiam 1997). In the present work, the analysis was performed using NDVI; therefore, every feature that numerically decreases can be interpreted (sometimes erroneously) as vegetation cover decreases even when they are bare soil or deep water, for example. As mentioned by Sohl (1999), the method is straightforward but its weakness is that it fails to provide adequate qualitative information regarding the nature of change. As a conclusion, it should be mentioned that the method could just provide a broad idea about the alterations in the landscape, being inefficient on the differencing changes from no-changes. A refinement of the method is then proposed, using threshold values, as mentioned in the next **section**.



(a)



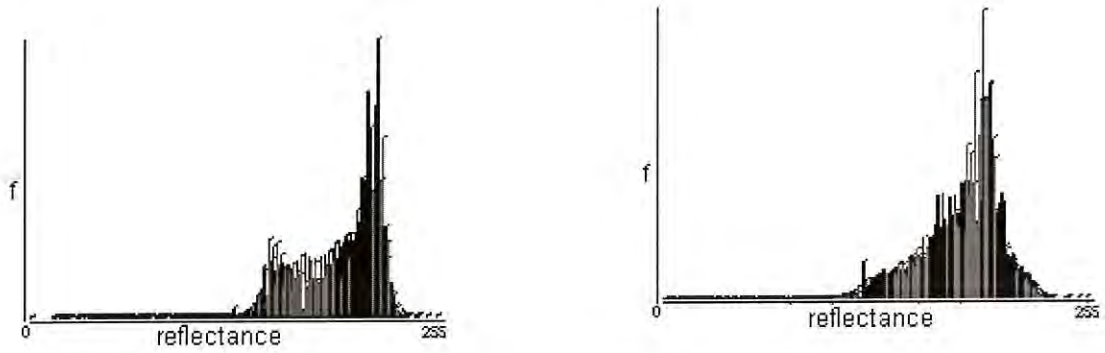
(b)



(c)

Figure 7.2 (a) NDVI87; **(b)** NDVI97; and **(c)** simple image differencing image from NDVI97 – NDVI87 (Dif.image). All images have been subjected to a histogram equalisation contrast stretch for printing purposes.

Figure 7.3 presents the histogram of both images (a) NDVI87 and (b) NDVI97. As mentioned before, data were converted to reflectance and the results multiplied by 255 during the pre-processing procedures.



(a) **(b)**

Figure 7.3 Histogram of grey tones of the images (a) NDVI87 and (b) NDVI97 (f=frequency)

7.3.1.3 Simple image differencing with thresholding

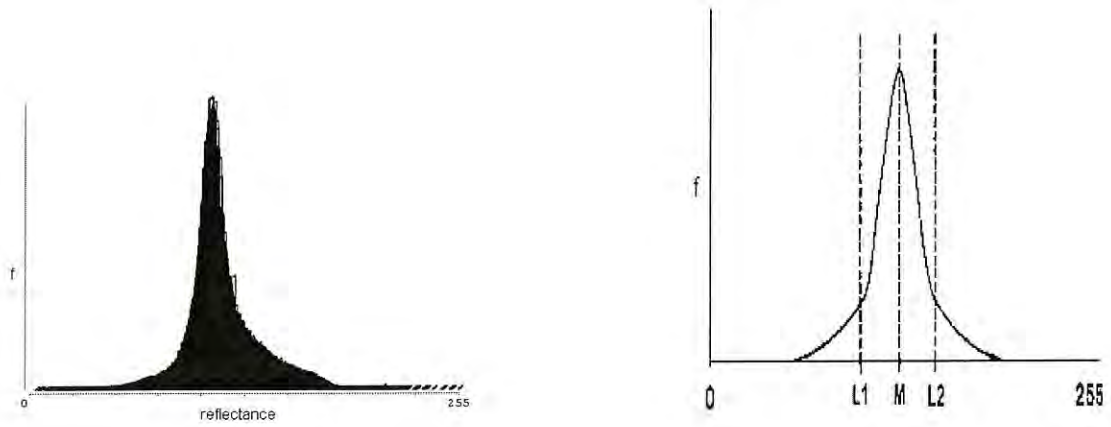
The objective of the method is to design an "automatic decision process" to define what is 'change' and what is 'no-change', based on some pre-defined criteria, involving limits for each class or category. According to Eastman (1997), the difference image needs to be thresholded, to distinguish change from natural variability. This involves a decision regarding which portion of the dynamic range will belong to each category. The weakness of this method is in accounting for the decision regarding where to place the threshold boundaries between "changed" and "no-changed" pixels (information may be lost if the threshold value is not accurately determined (Mispan 1997)). The choice can be arbitrary. Normally these values would be fixed by a trial-and-error procedure (Mather 1999). In this work, the term 'no-change' is used to characterise natural changes, where few differences are found between the pairwise comparison, such as in natural regeneration. Human-induced changes can be characterised as positive or negative changes, related to the augmenting or diminution of the selected feature, in this case, vegetation.

Some authors reported studies where thresholding was adopted as a criterion to classify different features. In Brazil, Alves *et al.* (1998a, 1999) reported change characterisation in the Amazon region and, by thresholding band 5 (Landsat TM), obtained information on the evolution of the deforestation process between two different dates.

Intending to fix the inability of simple image differencing to separate true change from random change, Thiam (1997) mentioned a principle that assumes that the values of the resulting difference image follow a normal distribution around the mean. The standard deviation (std.dev.), indicating the

distribution of the data, is used to establish the lower and upper thresholds, in combination with the mean. The number of standard deviations being used with the mean depends on the degree of suspected change. In the present study, this next step of image differencing was performed on the image depicted in **Figure 7.2c** (Dif.image).

Figure 7.4a presents the resultant histogram of Dif.image. Due to its symmetrical distribution, it was assumed that the distribution follows a normal distribution thereby allowing a statistical definition for the threshold chosen (Eastman *et al.* 1991). **Figure 7.4b** shows schematically the approach adopted in the present study.



(a) **(b)**

Figure 7.4 (a) Histogram of Dif.image; and **(b)** Schematic histogram of a difference image, indicating candidate areas to be thresholded.

(b) adapted from Mather (1999), where: L1 and L2 are the thresholds for the lower and higher tail of the distribution; M is the mean where the no-change category resides.

As mentioned by Mather (1999), a difference image tends to have a histogram that is normal in shape with the peak at 127 (using standard scaling). The region (M) in **Table 7.4b** represents the 'no-change' area, while the changing areas are located in the tails. **Table 7.1** presents the basic statistics for the original images (NDVI87 and NDVI97) and for the resultant different image (Dif.image).

Values mentioned in **Table 7.1**, were used to prepare a program in LEGAL language to spatialise a threshold applied to three categories, called 'positive change', 'negative change' and 'no-change', or natural variability (Eastman 1997). With this way of evaluating the results on the image display, various

thresholds were set interactively and the best threshold, defining areas of change, was selected; it was based on the criterion that the INF was located in the 'no-change' category. For this investigation, it was found that the mean, plus or minus one standard deviation, isolated negative and positive changes. Even with reasonable results, Eastman (1997) suggested caution when using the method because it is eventually image-specific. **Figure 7.5a** and **b** present respectively the three categories mentioned below in one thresholded synthetic image and a colour composite (NDVI87(R) and NDVI97(GB)), showing the visual matching between the categories created when the mean plus or minus 1 standard deviation are adopted and noticeable changes occur in the colour composite.

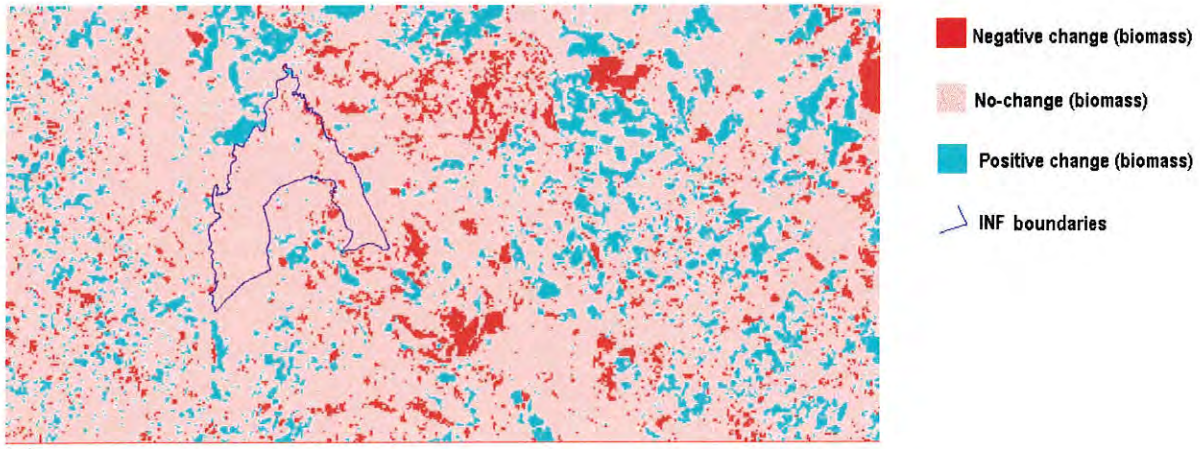
Table 7.1 Basic statistics of the NDVI87, NDVI97 and the difference images (Dif.image, Dif.imageR and Dif.imageN)

Band	Mean Reflectance(*)	Std Dev.	Var Coef.	Min reflectance(*)	Max
NDVI87	193.35	30.55	0.16	0	248
NDVI97	182.67	26.56	0.15	0	249
Dif.image	116.24	24.92	0.21	0	254
Dif.imageR	133.01	18.73	0.14	0	254
Dif.imageN	198.63	8.94	0.05	65	250

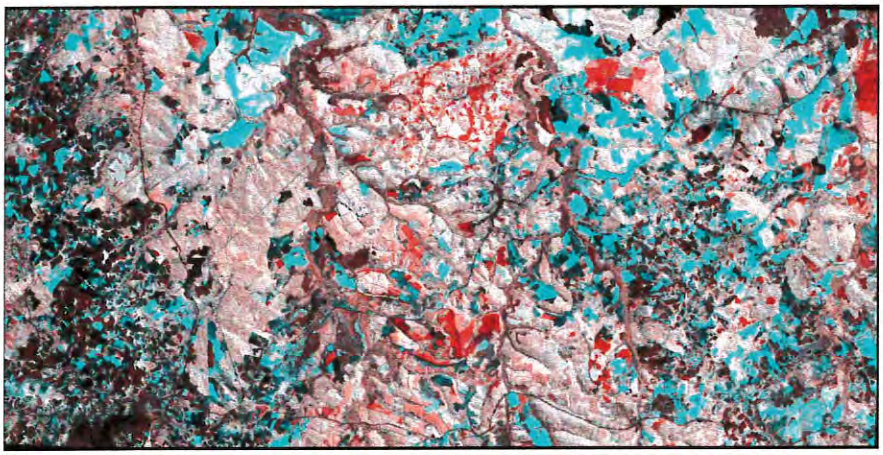
(*) Reflectance values were multiplied by 255

In **Table 7.1**, it can be noted that the range of the original NDVI87 image and NDVI97 image reflectance values did correspond and the mean values were similar (approximately 5% of difference) in the mean value.

Once more, it seems that the broad categories ('positive changes', 'negative changes' and 'no-change'), when visualised in the synthetic image, presented the same visual pattern of change, in comparison to a simple colour composite. The patches tend to be of the same size and shape. However, even if the threshold is well determined, another constraint must be mentioned. According to Mispan (1997) the image does not represent the quantitative nature of change, because different kinds of change may have identical results. If the operation results in a difference equal 30 (pixel value = reflectance * 255), for example, it could be the result of the subtraction of 230-200 or 130-100. This does not give the relative magnitude and direction of change, thereby complicating the analysis of the type and nature of change occurred during two different times. However, even presenting some restrictions, the method is an improvement when compared with the image differencing itself.



(a)



(b)

Figure 7.5 (a) Synthetic image (SDif.image) derived from Dif.image, presenting the three categories which represent 'negative changes' (red), 'positive changes' (light blue) and 'no-change' (light pink). The boundary in navy blue represents the INF; and **(b)** Colour composite of NDVI87(R) and NDVI97(GB).

Analysing the changes in the 10 years period, some preliminary comments emerge about "positive and negative changes". From the agriculture point of view, change is normal and it is expected, as rotation of crops is the usual practice. Using the image change detection procedures, this kind of change will emerge as a consequence of the alterations in texture and tone due to the crop replacement. However, non-forested areas are not the goal of the study and will not be discussed or detailed. Another source of positive change is the substitution of crops by forest plantations (typically species of the genus *Pinus* sp. or *Eucalyptus* sp.) which appear with a dark colour in the difference image (example labelled as 8, in **Figure 7.1**). The third source of positive change occurs as a result of the natural forest succession dynamics. However, some of the detected 'negative changes' are the result of the usual practice of

clearcutting commercial exotic man-made forests (example is labelled with (9) in **Figure 7.1**). The practice is very common in the region but they are also not the main focus of this study. The analysis is centred on the verification of eventual changes in the natural areas by selective or drastic exploitations in patches of *Araucaria* forest, which are adjacent to the INF, but in private ownership. A clear example of this environmental alteration is an area located at the right side of the INF ((10) in **Figure 7.1**). It appears lighter than the others in the image differencing process as a consequence of intense deforestation, as detected in the field campaign. Although a mask was drawn around the area (around 3000 ha) to facilitate its recognition, its contour is visible to the naked eye. Visual analysis shows: a) the differencing process depicted the exploited area (10) and b) showed no changes in the INF (1) area or in the areas labelled with (6) and (7).

7.3.1.4 Linear regression differencing

Linear regression is a statistical procedure that presumes linear relationships between two or more variables. The traditional statistical procedures is applied with the equation:

$Y = b_0 + b_1 X,$

Where:

Y = dependent variable
X = independent variable
b_0 and b_1 = coefficients

In the case of image processing, a special meaning is related to the intercept (b_0) and slope (b_1) coefficients. The intercept is called 'offset' and, according to Eastman *et al.* (1991), offset effects are not severe and images can be equalised easily. The slope coefficient is called 'gain'. It is related to non-uniform differences derived from atmospheric differences or variations in the calibration of measuring instruments. Image regression is supposed to deal with both offset and gain. As a matter of fact, using the technique, the image at date 2 is considered to be a linear function of image at date 1. **Figure 7.6a** and **b** illustrate the theoretical interpretation of the coefficients.

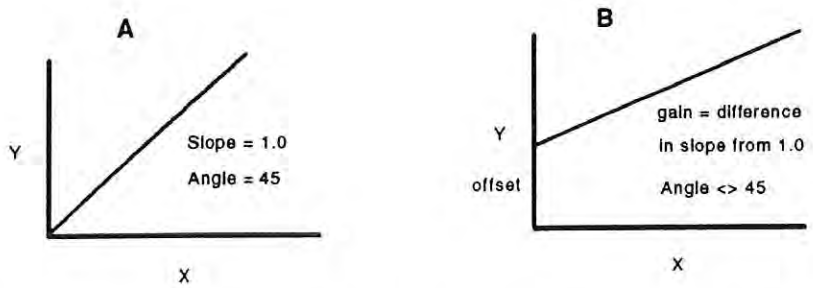


Figure 7.6 Interpretation of offset and gain in the linear regression differencing
Source: Eastman *et al.* (1991)

In graphic (A), the intercept would equal zero and the slope would equal 1, meaning no presence offset or gain. Graphic (B) shows that, with offset, an intercept other than zero will be found and with gain, the slope will be different from 1.

By this method, in the context of image processing, the earlier image (A) is used as an independent variable in a linear regression; the later image (B) is the dependent variable (Eastman 1997). By definition of the technique, the first image can be predicted by combining the second with the intercept (offset) and the slope (gain). After this first step, the method advocates that the actual and predicted images are then either differenced or ratioed to examine change, eventually using some kind of threshold to separate local variability from true change (Eastman *et al.* 1991). According to Thiam (1997), it is one of the strongest environmental change detection approaches. According to Eastman *et al.* (1991) image regression can also be considered effective as a pre-processing technique in change analysis. There is a second positive aspect to the method, besides its capability of dealing with differences caused by environmental factors (offset) and calibration differences between measuring instruments (gain). It is its ability to predict one image through the use of another. In applying the method to the study area, a linear regression was carried out between the 1987 image (NDVI87 - independent variable) and the 1997 (NDVI97 - dependent variable) NDVI images. The linear regression model is written as: NDVI97 = a + b * (NDVI87)

Figure 7.7 presents the resultant graphic scatterplot and trend line that expresses mathematically the character of the relationship between the images. The resultant equation was:
$Y = 0.304337 + 0.262885 * X$

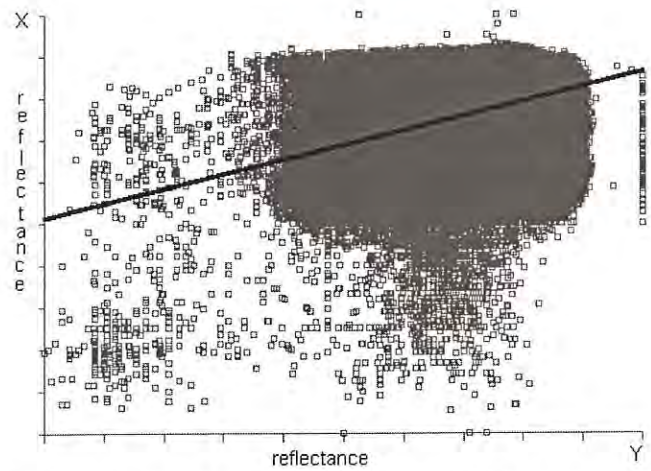


Figure 7.7 Graphic scatterplot of the NDVI87 (X axis) and NDVI97 (Y axis) and the linear regression carried out between the independent variable (NDVI87) and the dependent variable (NDVI97).

Where: coeff.of det. (r^2)=9.85%; std dev.of x (Sx)=0.1990724; std dv.of y (Sy)=0.1667707;
s.e.of estimate=0.1583467; s.e. of beta=0.0007883; t stat for beta=333.4854;
n (apparent degrees of freedom)= 1018162

In the first step of the method, it is assumed that the image of 1997 is a function of the one of 1987. In other words, "it is the way it is largely because of the way it was in the past" (Eastman *et al.* 1991). According to Eastman (1997) this new image can be thought as an adjusted version of the earlier image; this can be differentiated or ratioed as usual with the later one.

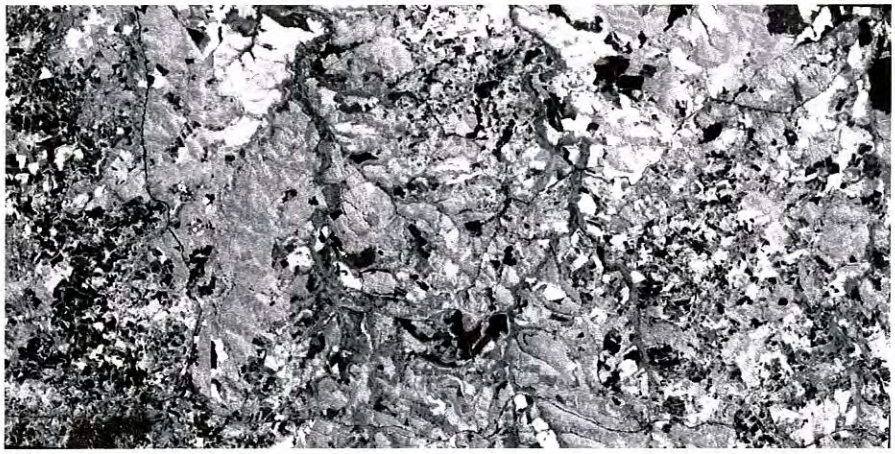
It is important to stress that the first stage of the regression method basically is intended to "produce" a new image of time 2 as an "adjusted" time 1 image. The interpretation of the determination coefficient is related to the similarities between the two dates, not an indication of the fitness of the model. Analysing the regression, it is possible to verify that little of the 1997 situation could be predicted by the situation in 1987. The low correlation coefficient of determination (r^2= 9.85%) indicates that much of the latest image is unusual. This was expected because the regression was performed over the window that involves the INF and surroundings. If only the INF area was highlighted, 1997 could be more easily predicted by the image from 1987, if a linear correlation is assumed. This is because it is a protected area where only natural processes happen (in general) instead of the adjacent areas, where anthropogenic processes affect the landscape in a major way. Therefore, the alternative approach would be to run the regression over an isolated (masked) area with expected little natural change (for example, inside the INF). Even when this area is identified, however, another problem surfaces, which is related to the theoretical assumptions of the regression model. It is difficult to achieve a representative sample of independent measurements (Eastman *et al.* 1991) because geographic data tend to be largely autocorrelated (strong spatial dependence).

When analysing the results of this first step, it was found that: a) the procedure may be a good option to address problems of non-comparability due to sensor changes in imagery that did not suffer any calibration but not necessarily will be advisable when radiometric restored data is available; b) for future analysis regarding the usefulness of image regression, it is strongly suggested that, for this first step, samples of data not corrected radiometrically should be compared with pre-processed data; and c) samples involving only areas within the INF should be tested (this should reduce the effects of autocorrelation) and then compared with the procedure that included the whole area.

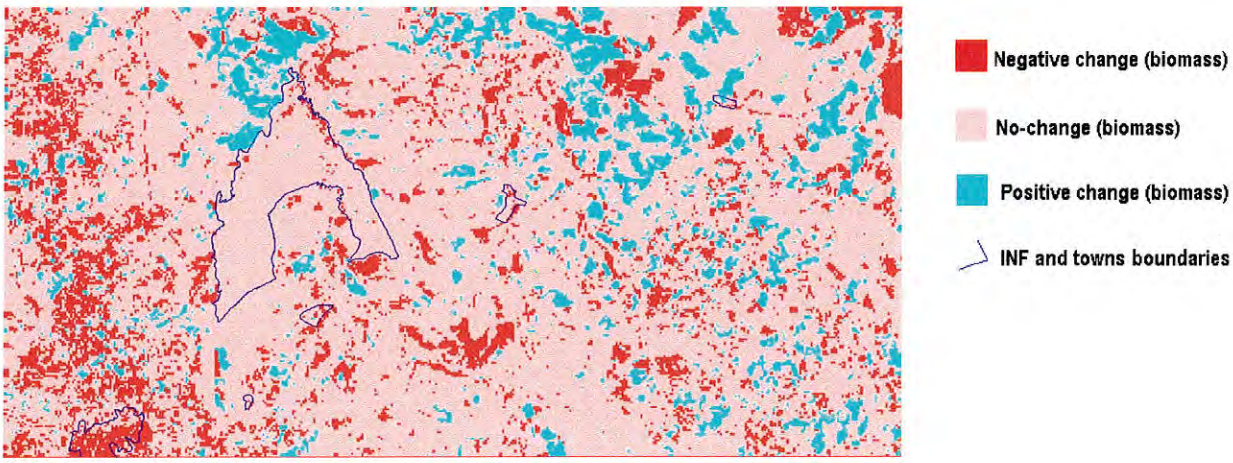
In the second step of the method, the dependent image was adjusted with the regression coefficients obtained previously. A new image was thus obtained; the subtraction process was then repeated. Basic statistics of the image are presented in **Table 7.1**.

A third stage involved the definition of values for the thresholding procedure, using the means and standard deviation of the new difference image (Dif.imageR) (**Figure 7.8a**). A program LEGAL was used to generate the new values and an output was again generated (Sdif.imageR), presenting the same three categories mentioned previously in the thresholded synthetic image, which can be seen in **Figure 7.8b**.

The image differencing was applied in order to isolate the unusual values (Eastman *et al.* 1995; Thiam 1997). It presents some differences when visually compared to the difference image after thresholding in **Figure 7.5a**. Both synthetic images were generated using the criterion of mean plus or minus 1 standard deviation, for definition of 'negative changes', 'no-changes' and 'positive changes'.



(a)



(b)

Figure 7.8 (a) Simple image differencing image (Dif.imageR); and **(b)** synthetic image (SDif.imageR) presenting the three categories which represent negative changes (red), positive changes (light blue) and no change (light pink). The contours in navy blue represent the INF and towns (or villages) in the region. The image (**a**) has been subjected to a histogram equalisation contrast stretch for printing purposes.

In the second approach using regression differencing (SDif.imageR), the region at the left side of the image ((2 and 3), in the **Figure 7.1**), which corresponds to urban, semi-urban and crops areas, presented more widespread patches being discriminated as 'negative changes' than the simple differencing image (SDif.image). The only difference between the approaches is that in the image regression, a new image was "adjusted" by the coefficients (gain and offset) determined in the linear

regression performed with the earlier 1987 image and 1997 image. By the techniques used in the present work, however, it was not possible to detect if that adjustment collaborated to determine the best pair of images for forest change detection. It was also not clear if the differences in those areas (2 and 3) occurred due to the impact of the coefficients in the adjusted image or the impact of applying 1 std.dv. as threshold. This threshold might be adequate; there was no differences (comparing with the previous method, in **Figure 7.5a**) in the area of INF, considered as a "reference" area but it is still questionable and, consequently, further work is suggested in this area. Consequently, no critical comparison can be made about the differences in the results. Urban expansion in the area is a fact; the region is now dominated by a relatively large urban area, the Irati town, which had approximately 42,000 inhabitants in 1987 and approximately 51,000 in 1997. According to Prefeitura Municipal de Irati (1999), the population is currently concentrated in the urban area (65%); the rural population is comprised by small farmers (35%). Its economy is based on expanding agriculture (potato, beans, corn, and soy). Due to the lack of spatialised ground data regarding urbanised areas in the ten years span, it is recommended that further research be conducted in this area. It was also apparent that the 'negative changes', in the region with natural forest (the central upper area of the image, labelled as (10) in **Figure 7.1**), were less remarkable in number and size of patches, when using the second approach. This again might be result of the image adjustment and/or thresholding limits.

7.3.1.5 Normalised Image Differencing (NID)

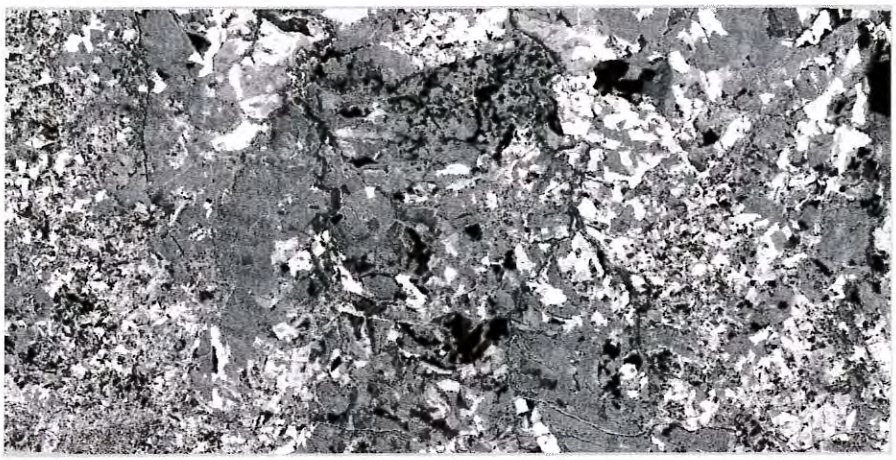
The NID technique is a modified version of image differencing. More specifically, it is a combination of image differencing and image ratioing. Mathematically it is expressed as:

$$\Delta p^{m}_{(i,j)} = \frac{p^{m}_{(i,j)}(t_2) - p^{m}_{(i,j)}(t_1) + c}{p^{m}_{(i,j)}(t_2) + p^{m}_{(i,j)}(t_1) + c}$$

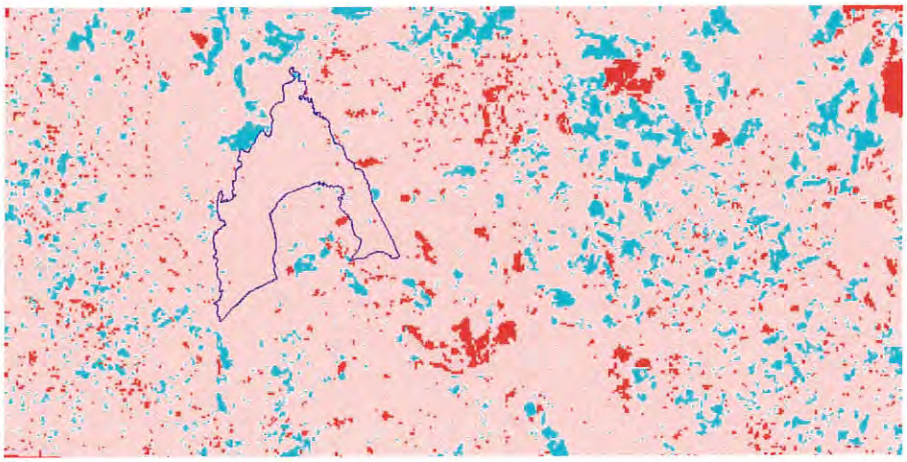
where:

$\Delta p^{m}_{(i,j)}$ = difference pixel value; $p^{m}_{(i,j)}$ = pixel value at line i, column j, and band m; t = time and c = constant

The technique was applied to the modified images as obtained in **section 7.3.1.4**. Basic statistics for the new image Dif.imageN can be seen in **Table 7.1.**The resultant images are presented in **Figure 7.9 a** and **b**. The same constants (scaling factor of 127) were included in the formulae.



(a)



(b)

Figure 7.9 (a) Normalised image differencing image (NID) (Dif.imageN); and **(b)** synthetic image (SDif.imageN) presenting the three categories which represent 'negative changes' (red), 'positive changes' (light blue) and 'no-change' (light pink). The contour in navy blue represents the INF. The image **(a)** has been subjected to a histogram equalisation contrast stretch for printing purposes

The same threshold (mean, plus or minus 1 standard deviation) was used again. The procedure produces complex results for analysis. One of the reasons is that the formulae include a ratio operation of images subtraction and images sum. In change detection, ratioing is expected to reduce the effects of variations in sun illumination angle, topography and sensor calibration, the so called environmental and system multiplicative factors (Thiam 1997). The author perceived a possibility of using the method to emphasising changes at low level. The technique was capable of detecting changes in the same way as the simple differencing and the regression method with differences that will be discussed in the next section.

7.3.1.6 Agreement analysis

Three different approaches were designed to assess vegetation changes in the study area. Comparison between the three methods (simple threshold (1), linear regression (2) and normalisation (3)) is complicated due to the basic differences between them. Therefore, the criterion used to define their adequacy was visual analysis of changes, specifically any changes in the area of the National Forest (INF) and areas 6 and 7 (in **Figure 7.1**). In those areas, very little change was expected as a result of anthropogenic interference, in the ten years period. The natural forest was not managed. Thinning only was applied to few stands of plantations. Thus, the best methods were expected to reflect that condition of non-alteration in the INF area.

Even if the three methods are based on some different assumptions making comparisons between them difficult, some analysis can be performed. The difficulties can be summarised as: a) methods 2 and 3 which shared the same basic images, presented results visually different – the areas under urbanisation pressure were not detected as negative change in the same percentage. Actually, the resultant images were quite different. When images were only differenced, rivers and floodplain areas were more detectable and urban versus non-urban areas more emphasised than when the images were normalised. Method 2 included that expansion of the urban area (mainly Irati town, but also Teixeira Soares and Fernandes Pinheiro) in the 'negative change' class. As a consequence, it presented the highest percentage of negative changes. Both methods identified positive and negative changes, but it seems, based on visual analysis (of bands 3,4 and 5 colour composites) of both dates that the method 3 tended to underestimate a bit the alterations in urban and semi-urban region of the study area; b) both methods similarly classified the INF and areas 6 and 7, where only natural changes was expected mainly in the 'non-change' class. However, by visual analysis, method 2 tended to exaggerate the size of one area considered as 'negative change' inside the INF; c) in method 3, the standard deviation was noticeably smaller, meaning less heterogeneity in this resultant image; data were more compressed (dynamic range of 65-250). Consequently a higher percentage of 'no-change' was presented; d) method 1 classified 'gain and losses' of vegetation in urban and semi-urban more sensitively than method 3. It was also the method that included more percentage of negative-changes in the area labelled as 10 (**Figure 7.1**), where the major deforestation occurred in the study area. It was similar to method 2 when classifying areas inside the INF and areas 6 and 7.

Table 7.2 presents information about the percentage of the area considered as unaltered, indicating 'positive changes' and 'negative changes' for each approach. As mentioned before, it must be remembered that positive changes usually reflected some agricultural management or plantations of exotic forest species and the negative changes included all the clearcuttings in pine plantations, a common occurrence in the region. It was built by vectoring the layers and cross-tabulating each two

thematic resultant images. The results for each class of changes (negative, positive and no-change) are presented in hectares. The generated images were obtained by different processes and philosophies and comparison between them is difficult. However, a measure of the differences among the three methods and the agreement area between them is presented. Both SDif.imageR (obtained by regression) and SDif.imageN (normalised image differencing) were analysed against SDif.image (simple image differencing).

Table 7.2 Cross tabulation between each two thematic layers

(1) SDif.image (rows) SDif.imageR (columns)			
Values in ha	**Negative change**	**No-change**	**Positive change**
Negative change	520.65	168.48	0.09
No-change	702.72	5585.49	451.80
Positive change	1.89	451.80	645.57
(1) SDif.image (rows) SDif.imageN (columns)			
Values in ha	**Negative change**	**No-change**	**Positive change**
Negative change	350.73	335.79	2.70
No-change	14.40	6435.63	4.77
Positive change	0.27	361.62	737.37

(1) SDif.image x SDif.imageR - Overall agreement: 81.91%		
	SDif.image (%)	SDif.imageR (%)
Negative change	42.49	75.54
No-change	89.99	86.53
Positive change	79.48	58.73
(2) SDif.image x SDif.imageN - Overall agreement: 91.27%		
	SDif.image (%)	SDif.imageN (%)
Negative change	95.99	50.89
No-change	90.22	99.70
Positive change	99.00	67.08

The first segment of the table reflects how each approach classified each category and how they agreed (grey cells), classifying areas in the same class. One image is read in the rows and the others in the columns. The grey cells show the area of agreement, in hectares. Both SDif.imageN and SDif.imageR, when compared with the basic simple image differencing (SDif.image) presented high percentage of overall agreement, which in this case means the sum of agreements (in hectares) for each class against the total area in hectares. The class 'no-change' (89.99% and 86.53% - in the SDif.image versus SDif.imageR (1); 90.22% and 99.70% - in the SDif.image versus SDif.imageN (2) presented the highest agreement percentage between the methods. However, the discrepancy between methods in the classes 'negative change' and 'positive change' was considerable. As a conclusion, it would be commented that the simple image differencing is: a) easier to be implemented; b) more straightforward in terms of analysis; and c) presented high percentage of overall agreement with the other two methods.

7.4 Categorising losses of vegetation

The analysed landscape involves a protected area (INF). No changes (at least no considerable changes) were expected in the natural forest areas over a long period of time, excepting natural disturbances (any discrete natural event in time that disrupts the natural communities) or the natural vegetation growth process. Besides INF, another area (labelled in **Figure 7.1** as (10)) with almost the same area of the National Forest and in the past covered by *Araucaria* forest, was severely disturbed in the 10 years period, as detected through the field work. Analysis of the information obtained in the field campaign and analysis of the images of both dates and the resultant difference images indicated that the disturbance actually presented different levels of deforestation. The previous methodology ought to be adapted to incorporate one more level of disturbance and be able to categorise these alterations. However, the threshold limits are very difficult to determine. A decision was taken to test methodologies that incorporate the analyst's perception of the phenomenon and of the impact of some changes over the environment. In the next **section**, an analysis of other change detection methods is presented regarding their capability for detection of those different levels of disturbances and change in the forest environment.

7.4.1 Image differencing classification

Some change detection approaches adopt image transformations to merge two temporal sets (USDA 1995) or use direct classification of a multspectral or multidate data set (Jensen *et al.* 1993; Soares *et al.* 1996; Sunar 1998). Change detection has been also accomplished by comparison of two independent image classification results (Häme *et al.* 1998). A major problem with this approach is that the final results incorporate errors that may have been committed in the individual classifications; some minor changes may not be even considered. According to Häme *et al.* (1998), it is difficult to differentiate if, for example, change in the value of biomass is due to an actual change or to an error in the interpretation model.

There are two classical approaches for image classification: the supervised approach and the unsupervised approach. According to Novo (1988), a third category can be included, the hybrid classification. The topic has been extensively detailed in classical textbooks. Lillesand and Kiefer (1994) defined well the differences between the two first approaches mentioning that "in the supervised approach we define useful information categories and then examine their spectral separability; in the unsupervised approach we determine spectrally separable classes and then define their informational utility". Thus, the fundamental difference is that supervised classification involves a training step followed by the classification procedure (Lillesand and Kiefer 1994). USDA (1995) mentioned that supervised classification is a method of image classification that depends upon the direct involvement

of the analyst in the pattern recognition process. Usually, supervised classification requires the analyst to select training samples from the data, which represent patterns to be classified.

In the present work, Dif.image5 (the image resultant from the band 5 image differencing process) was used to test its detection sensitivity for different levels of vegetation loss. The image presented some clear differences in degradation levels. These were used in the training process. The use of one single band in the process was incorporated because it is the simplest way of detecting changes (Alves *et al.* 1998b).

7.4.1.1 Data and methods

With this approach to analysis, a supervised classification was performed using. The objective of using supervised classification was to involve the analyst in the process of defining classes that could be later used as indicators of relative degradation in *Araucaria* forest managed areas. In the fieldwork, it was possible to define three levels of forest degradation, defined as 'severe', 'intermediate (moderate)' and 'slight (described below)' changes. The training areas were then searched for each of the three levels, plus two more categories: 'no-change' and 'biomass gain'. Local foresters were consulted to indicate sites that represented typical vegetation-degradation conditions. Samples for the 'no-change' category were sought only inside the INF and in areas 6 and 7 (from **Table 7.1**). 'Biomass gain' samples were sought in two areas close to the INF (labelled as 8, in **Figure 7.1**) which were reforested with species of the genus pine in 1992. These areas were visited in the field campaign. The area was covered by pasture and crops in 1987. 'Severe biomass loss' class was investigated in three areas - two stands located near to the INF, the third located in the top right position of the image window – labelled as 9, in **Figure 7.1**). All were forested with pine in 1987 and suffered clearcutting during the 10 years period of analysis. The areas were in process of conversion to agriculture. Two additional small areas, where deforestation had occurred, were located in the *Araucaria* forest (area labelled as 10, in **Figure 7.1**). The most difficult categories to define were 'intermediate' or 'moderate loss' and 'slight' or 'minimal' loss. During the field campaign, while field plots positions were determined by GPS, this procedure was also adopted to locate some natural areas around the INF. These were visited with the objective of locating different intensities of degradation in areas of *Araucaria* forest. The area labelled as 10 in **Figure 7.1** had more historical information available. Areas that were forested in 1987 (according to local foresters) were located. A rapid assessment was made in 1998 in order to classify secondary patches of *Araucaria* forest in two categories: a) 'intermediate loss' was classified by the absence of *A.angustifolia* and other important species (mentioned in **Chapter 2**) and by an evaluation of number of trees and volume by reconnaissance; b) 'slight loss' was defined as: areas where the original *Araucaria* forest was partially altered specially by the selective exploitation for wood purposes; with absence of *A.angustifolia* (or presence in small percentage); presenting higher density and volume than the 'intermediate loss' category.

7.4.1.2 Sample size and classification scheme

The training groups define the classes dispersion diagram and its distribution of probability, considering the distribution of normal probability for each class of the training. According to NETGIS (1998), 10 to 100 pixels per class is the minimum class size for any sampling design. Campbell (1996) considered that, as a general guideline, the operator should ensure that several individual training areas for each category provide a total of at least 100 pixels for each category. According to Lillesand and Kiefer (1994) the theoretical lower limit of the number of pixels that must be contained in one training set is n+1, where n is the number of spectral bands. He added that in practice a minimum of 10 to 100 pixels is used since the estimates of the mean vectors covariance matrices improve as the number of pixels in the training set increases. Campbell (1996) also mentioned that each training site must be large enough to provide accurate estimates for the properties of each informational class. Nevertheless, he also recommended that these areas should not be too big, as large areas include undesirable variation. Campbell (1996) suggested that the optimum number of training areas depends on the number of categories to be mapped, but he considered 5 to 10 to be reasonable numbers. In the present work, sampling was designed to be between 45 and 48 sample units for each category. Each sample units comprised 9 to 20 pixels. The method's effectiveness assessment separated the sample units in two groups: one group was used for the training set and the second one for the accuracy assessment.

Classification was based on the maximum likelihood decision rule (MAXVER), a powerful classification technique (Campbell 1996). According to this author, MAXVER maximises the probability of a correct classification given the information in the training data. It is based on normal distribution assumptions and it considers not only the mean, or average, values in assigning classification but also the variability of the brightness values in each class. MAXVER is an adequate procedure to be performed in imagery with the resolution of TM data (Ringrose *et al.* 1997). A pos-processing procedure with the objective of eliminating isolated points was performed. Depending on weight and threshold entered in the model the point will or not have its class substituted by the most frequent class in the neighbourhood. The smaller the weight and threshold, the bigger will be the number of substitutions. Maximum weight and threshold are scaled to 7. The best results were defined by visual analysis using weight and threshold equal 3. The results of the classification process (Dif.image5c) are shown in **Figure 7.10**.

When analysing the resultant image, it was noted that: a) inside the INF – in general there were no changes detected except for areas in process of regeneration - detectable mainly when close to internal earthen roads. A concentration of positive changes was found close to the number (1) in **Figure 7.10**. The area was a field trial of species of the genus Acacia. Because of its high mortality rate, it was abandoned as an experiment years ago. The area now presents aspects of second-growth

forest; and b) outside the INF: clearcuttings of pine plantations were captured as severe losses of biomass. In addition, some areas concentrated in the region labelled as (10) were also classified as 'severe loss'. These areas were located following the streams and in the low valleys. Intermediate losses of biomass were detected in pine plantations that suffered severe thinnings and in the area labelled as (10), around the areas classified as severe. Slight losses were distributed in the whole area.

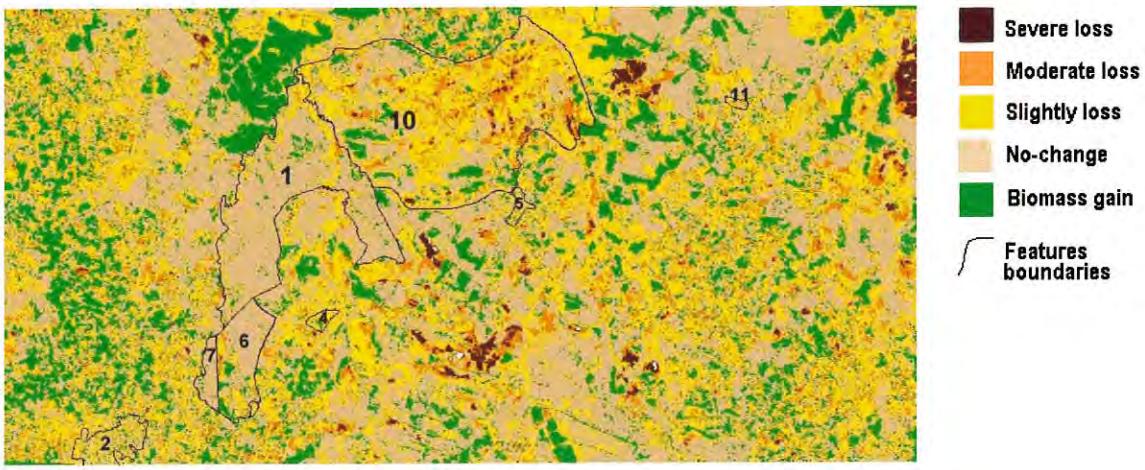


Figure 7.10 Classified change image (SDif.image5s)

Gain of biomass was concentrated in the areas where agriculture was substituted for forestry. Because the emphasis of the study was on identifying forest changes, minimal attention was given to optimising the classification results within the urban areas and other non-forest categories.

By the results, it was confirmed that TM band 5 and supervised classification are promising as tools to contribute to the monitoring of *Araucaria* forest. The differentiation between levels of degradation was achieved as a result of the training effort. The classification effectiveness involved the comparison of the categorised data to the reference set by the error matrix, presented in **Appendix 7.1**. When analysing the producer's accuracy it was noted that practically no errors were found in mapping 'biomass gain'. Some misclassification had occurred between the class 'no-change' and its immediate classes, 'slight loss' and 'biomass gain'. The class severe loss presented very little confusion, except occasionally with the class 'intermediate (moderate) loss'. As expected, however, 'intermediate loss' was the class that was less accurately classified (PA = producer's accuracy = 75.76%. PA is a measure of the proportion of total sampling units that were correctly classified). The user's accuracy (measure of commission errors), indicates the probability that a unit within an individual category is correctly classified. It presented values superior to 84% for all classes. A global accuracy of 90.66% was found. Considering the difficulties in defining some of the classes, it can be considered adequate.

Another traditional form for estimating overall accuracy has been the Kappa (or KHAT) coefficient of agreement. A description of the use of the Kappa statistics can be found in Congalton *et al.* (1983). The K with ^ (or KHAT) statistic is a measure of the difference between the actual agreement, between reference data and an automated classifier and the chance agreement between the reference data and a random classifier (Lillesand et all 1994). The formula is:

KHAT (Kappa) = (observed accuracy – chance agreement) / 1 – chance agreement

Lillesand et all (1994) considered that this statistic serves as an indicator of the extent to which the percentage correct values of an error matrix are due to "true" agreement versus "chance" agreement. As true agreement (observed) approaches 1 and chance agreement approaches 0, KHAT approaches 1. The statistic attempts to remove chance agreement from estimates of classification accuracy. A Kappa value close to 100% would mean that there is a perfect agreement between the classification and the field data. In the present work, the Kappa result was 87.80%. Foody (1992) has revealed that Kappa tends to underestimate the overall classification accuracy. Another statistic that was developed after Foody's (1992) is TAU (Ma *et al.* 1995), a Kappa-like technique that is suppose to be a more precise measure of accuracy that removes chance agreement from the estimate of classification accuracy. The TAU statistic (88.32%, in the present study), error matrix and the Kappa coefficient are presented in the **Appendix 7.1**.

Analysing the final results, it was found that even when good results are obtained in discriminating the different levels of change in the study area, this traditional approach suffers from limitations related to the assumption that each accuracy site can be unambiguously assigned to only one class. The defined classes in the present work might be better analysed if considered as continuous instead of discrete classes. In reality, the categories are not easily defined as cover types identified by species like 'pine' or 'eucalyptus' should be. This reasoning leads to the possible solution for this kind of problem by using mixing models (as discussed before) or fuzzy logic. Fuzzy classification analyses land cover at a scale below that of the spatial resolution of the sensor (Foody 1999a). It is a relatively recent innovation in accuracy assessment (USDA 1995), but some techniques for the analysis has been already developed, making possible the identification of changes, including land cover conversion and modifications (Foody 1999a). The approach was not implemented in this work; however, in retrospect it would seen to be a valuable tool to accurately assess changing in *Araucaria* forest and it is strongly recommended for further studies. As an alternative approach, fraction images derived from the linear mixing model (presented in the **Chapter 6**) were included in the next section.

7.5 Alternative process of change detection for monitoring purposes

The purpose of this **section** was to verify the possibility of developing an alternative change detection procedure for *Araucaria* forest. The study area is a small window of the whole biome which is now extremely fragmented and, in the past, covered a great part (almost 200,000 km^2) of Southern Brazil (**Figure 2.1**). There are no long-term programmes for continuous monitoring of the biome degradation and the last Forest Inventory was conducted in the area in 1984. The greatest part of the forest remnants is in private hands where forest management is allowed, meaning frequently depletion of the "commercial" species, some of them endemic of this forest type. Shifting and extensive cultivation, uncontrolled logging and unplanned land use are the reality of the region. As a consequence, important vegetable and animal species, might disappear in the near future due to extension of the heavily managed areas and lack of sufficient protected remnants.

Monitoring *Araucaria* forest is important in two directions: a) detecting changes in area (extension or defraction); and b) detecting degrees of degradation of the forest remnants. There is no experience of the use of remote sensing at regional scale in the region, for the purpose of forest monitoring. Even the traditional methods, based on visual analysis are focused in the Amazon Region, where a well established programme (PRODES) is being conducted. Recently, limited and local efforts are being reported on the use of digital classification based on pixel-by-pixel assumptions, again focusing on the Amazon Region and/or Atlantic Coastal Region. No effort has been concentrated on developing methodology for monitoring the biome *Araucaria* forest and, consequently, no early information about successful projects can be incorporated into the new initiatives. Visual interpretation and conventional digital analysis will always be necessary, and frequently the only way to perform some diagnoses. However, some techniques are laborious, time-consuming and demand so much expertise that the final costs are prohibitive even when they present reasonable accuracy (Hill 1999). Digital analysis offers a huge opportunity of having an updated panorama of large areas, as *Araucaria* forest, by monitoring progress toward the vegetation conditions, using traditional methodology or more recently developed procedures, involving image segmentation and unsupervised classification, for example. There are advantages and disadvantages in using unsupervised classification, and they are related to the limited control of the analyst over the menu of classes and their specific identities (Campbell 1996). For example, some spectrally homogeneous classes do not necessarily correspond to the classes under interest. However, some of these limitations maybe can be softened by the concomitant use of other techniques, as image segmentation (which is detailed in **section 7.5.2**).

In the present study, the tested method was based on Shimabukuro *et al.* (1998). The authors suggested the use of shade image for deforestation evaluation in the Amazon Region. The main reason for using shade image in deforestation assessment is that it is has been found to be well

correlated with vegetation canopy structure; in other words, dense vegetation has a medium proportion of shade within the canopy while deforested areas, defined by Shimabukuro *et al.* (1998) as bare soil, pasture and/or regrowth, have a comparatively small proportion (Adams *et al.* 1995). Shimabukuro *et al.* (1998) also mentioned limitations of conventional digital analysis of TM data based on pixel-by-pixel classification because it takes into account only the spectral variation of the scene and misses the contextual information of the objects, which is potentially important. The fact was also commented by Eastman (1997). He mentioned that, although some procedures are able to incorporate information about such image characteristics as texture and context, the majority of image classification is based solely on the detection of the spectral signatures (i.e., spectral response patterns) of land cover classes. The new approaches involving image segmentation and fraction images obtained by mixing models tend to improve the deforestation estimation.

7.5.1 Step 1 - Preparing the differencing image

The Linear Mixing Model (LMM) analysed in **Chapter 6** for the 1987 data set has the potential of being included in change detection studies because of its inherent characteristics of dimensionality reduction and the fact that one of its synthetic products, the shade image, is related to the forest canopy structure (Shimabukuro *et al.* 1998). The shade image for 1987 (**Figure 6.7d**) described in **Chapter 6** was used in the process. The same procedure described previously for the 1987 data set was also adopted for the 1997 data set. Only the procedure using three end-members was considered for 1997 and three fraction images were generated: component *A.angustifolia*, component *Araucaria* forest and component shade. The image differencing procedure was then performed and the shade image of 1987 was subtracted from the shade image of 1997. **Figure 7.11** presents the results of this differencing procedure (Dif.image-sh).

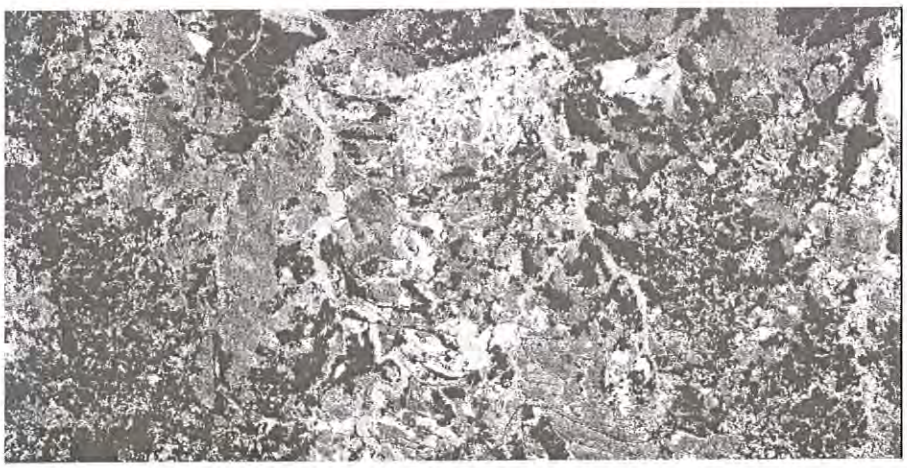


Figure 7.11 Image differencing obtained from the shade image of 1987 and the shade image of 1997 (Dif.image-sh). The image has been subjected to a histogram equalisation contrast stretch for printing purposes

The use of the fraction image obtained by the component *A.angustifolia* (Vt1) instead of the shade image was also investigated as well the combination of both (Vt1 and shade) to be involved in the segmentation process. As mentioned in the **section 6.10.1.3**, Vt1 presented interesting results regarding discrimination between the forest types of the study area. However, a more detailed analysis of the image led to the conclusion that the Vt1 image captured the presence of *P.elliottii* and the amount of *A.angustifolia* (in plantations and in the natural areas), but did not differentiate *P.taeda* from *A.angustifolia*. The resultant difference image (Dif.image aa) highlighted reforestation (biomass gain), but it was not capable of indicating other forms of change in vegetation that occurred in the area clearly. Actually, it reflected the tendency of Vt1 from 1997 in highlighting young plantations, in this case, pine plantations. Some preliminary tests classifying the image were also not well succeeded and, as a consequence, the image was not included in the next steps of the process. The image resultant of the image differencing process (Vt1 from 1997 – Vt1 from 1987) is presented in the **Figure 7.12**.



Figure 7.12 Simple differencing image Dif.image-aa

7.5.2 Step 2 – Segmenting the differencing image

The next step was to proceed to the thresholdings. This time, instead of testing the statistical approach (mean plus or minus 1 standard deviation for example), it was decided to test the segmentation procedure, which fragments the studied area into regions. Actually, image segmentation consists of partitioning an image into units that are homogeneous with the constraint that the union of no two neighbouring units be homogeneous (Beauchemin *et al.* 1997). According to Abeyta *et al.* (1998), once the image is broken up into regions, each region can be assigned to a theme (land-cover or vegetation class) using classification based on average characteristics or on the plurality from a per-pixel classification overlaid on the segmentation. The authors mentioned that region-growing image

segmentation algorithm was developed by C.E.Woodcock and J.Harward in 1992 and has been applied to forest stand delineation. Regions are a group of "contiguous pixels", which spread bidirectionally. A region growing scheme uses spectral and spatial (texture) properties of pixels to merge the pixels into homogeneous regions (Abeyta *et al.* (1998). One of the algorithms implemented in the software SPRING is based on the method of region growing. It is an iterative technique for grouping data, in which the spatially adjacent areas can be grouped. Initially, this segmentation process labels each pixel as a different area and starts grouping pixels until the whole area is processed. The process of image segmentation implies in the definition of the two categories of thresholds: a) similarity threshold (ST); and b) area threshold (AT) (Nascimento *et al.*, 1998). The similarity threshold defines the spectral similarity between two regions and it is the Euclidean distance between the mean digital number (in the present work, shade proportion) of two regions under which two regions are grouped together (Shimabukuro *et al.* 1998). The area threshold is the value of the minimum area (in number of pixels) which individualises a region. There is no ideal value defined *a priori* and the combinations between thresholds are normally defined by trial-and-error method.

The use of image segmentation in vegetation studies is increasing since algorithms are being implemented in various image-processing systems. In Brazil, it has been tested mainly in the Amazon Region by the INPE team (Alves *et al.* 1996; Alves *et al.* 1998a; Pereira *et al.* 1998; Shimabukuro *et al.* 1998) and has been implemented in the software SPRING. Few other studies were published using segmentation to analyse other Brazilian regions, but no one focusing the *Araucaria* forest region.

In this step of the present study, various combinations of ST and AT were tested and the best similarity threshold was 30 and the area threshold was again 30. The smaller the ST the bigger will be the number of borders included by the algorithm. The value 30 was the one that visually matched better with the five classes of forest changes previously defined. The decision is made based by displaying the shade image and the obtained segmented image superimposed to each other. An AT of 30 indicates regions with minimum size of 27,000.00 m^2, (2.7 hectares) which is adequate as a minimum size for forest stands or patches analysis. Smaller AT should create too many polygons and confuse the analysis. The selected similarity and area thresholds are not comparable with other studies performed using similar procedures (Nascimento *et al.* 1998; Alves 1998b). This happens not just because of the peculiarities in the areas but also because in this study the effort was dedicated to classify forest changes instead of classifying forest types.

In this work, the criteria for thresholds acceptance were the same as mentioned in the other adopted approaches: a) visual analysis; and b) the use of the INF as a guidance for no-change areas. The superposition of the segmented image over the images used to generate them is a qualitative but efficient approach to analyse the adequacy of the segmentation process (Nascimento *et al.* 1998). **Figure 7.13** presents Dif.image-sh and the labelled image obtained by the selected segmenting

thresholds (lines in red). The small picture is a zoomed window of an area that presented the five classes of vegetation losses identified previously, where: 'severe loss' (1), 'moderate loss' (2), 'slight loss' (3), 'no-change' (4) and 'biomass gain' (5).

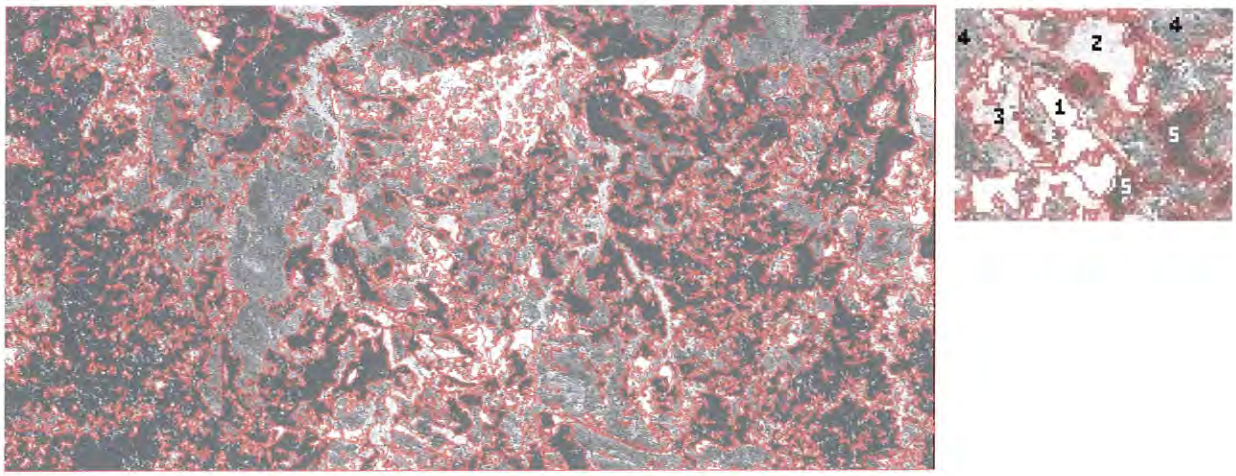


Figure 7.13 Segmentation of the Dif.image-sh and a zoomed window presenting the classes: 'severe loss' (1), 'moderate loss' (2), 'slight losss' (3), 'no-change' (4) and 'biomass gain' (5).

7.5.3 Step 3 - Classifying the segmented image

At this stage of the work, it was decided to test some other techniques that should be included as elements of a monitoring system of *Araucaria* forest. Besides supervised classification, analysed in the section above and in again in the next section, it was considered important to try some techniques that should diminish the impact of the training process in the classification step; an unsupervised approach was considered useful to be tested, after the success of other projects (Alves *et al.* 1996; Shimabukuro *et al.* 1998). Howard (1991) asserted that fully automated thematic mapping is unlikely to be achieved, since the human element is essential and adaptable in visually interpreting variations in the local landscape (e.g. tree shadows). However, one of the advantages of the unsupervised classification is that it requires less knowledge of the area than the techniques that are based on the training process and, for projects that involve huge areas it might have impacts over the expertise costs. Because the main purpose of this section was to generate an alternative procedure to detect changes in *Araucaria* forest, an unsupervised procedure was chosen. Unsupervised classifiers normally involve aggregating pixels from the image into spectral groupings. Those classifiers simulate the decision-making of a photo-interpreter, when recognising homogeneous areas of images, based on the spectral and spatial properties of images. The border information is used initially to separate areas and the spatial and spectral properties will group areas with same texture. For this study, the unsupervised classification was performed using the per-field clustering classifier - ISOSEG algorithm in SPRING for Windows 3.2. ISOSEG applies clustering and uses the vector mean and the covariance matrix of the segmented

regions (calculated separately) to define the centre of the classes (Nascimento *et al.* 1998). The classifier requires the specification of an acceptance threshold to aggregate or separate classes. This acceptance threshold is the maximum Mahalanobis distance over which shade proportions can be moved from the centre of a class (Shimabukuro *et al.* 1998). According to (Nascimento *et al.* 1998) the threshold can be considered as a hyperellipsoid in the attribute space; all regions that are inside of this hyperellipsoid will belong to the same class. The available options in SPRING are: 99.9%, 99%, 95%, 90% and 75%. All the options were tested in the present study and the best results were found when 75% was adopted because it provided the smaller number of classes, compatible with the purposes of this work. **Figure 7.14** presents the resultant classified image.

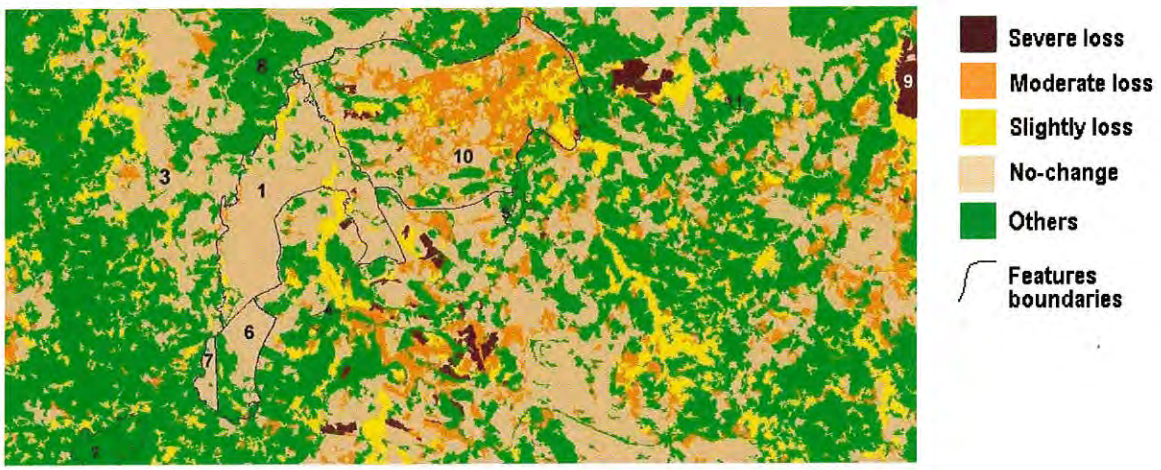


Figure 7.14 SDif.image-sh-u (ISOSEG classification of the segmented image Dif.image-sh)

The analysis of the **Figure 7.14** shows that: a) inside the INF – the only detectable positive changes were verified in the areas of former *Acacia* spp. stands. Some 'slight biomass loss' was also detected in the riparian areas which might be a consequence of the flooding process that normally occurs in the area; b) outside the INF: no changes were detected in the areas labelled as (6) or (7). The area (6) is protected (Estadual research institution) but the area (7) is in private hands. Wood exploitation is regulated by law (management plans); however, exploitations against the law as identified in the area labelled as (10) are possible. The mentioned area was almost totally classified as 'moderate biomass' loss and some few small patches were included in the 'severe loss' class. In the period of 10 years, the region, which is equivalent in area to the INF, was deeply altered. **Appendix 7.2** presents some facts and concerns related to the area. The class 'severe biomass loss' was only identified when a highly dense area were clearcut (commercial forest). Only a few natural areas were included in this same class.

The classifier incorporated agriculture, pasture and urban areas as 'biomass gain', in the same class as reforestation (young pine, mainly). The fact maybe can be partially explained by the characteristics of shade image. According to Adams *et al.* (1995) deforested or sparsely forested areas present low shade contents. Actually, the class should not be called biomass gain but "others". Therefore, gain in biomass was not well modelled; the shade image did not provide information enough to the discrimination between forestation (and/or reforestation) and other cover types present in the study area. It is a typical tool to analyse deforestation classes. The image, therefore was composed by the three levels of deforestation, the class 'no-change and the class 'others', where others features were included.

7.5.4 Supervised classification

With the objective of analysing performances of both approaches (supervised and unsupervised classifications) using the same shade differencing image, Dif.image-sh was used to perform a supervised classification, using MAXVER. The same sampling approach and same reference data mentioned in the **section 7.4.1.2** was used; in the training stage, only the categories 'no-change', 'severe loss' and 'biomass gain' were easily identified in the shade difference image. The categories 'moderate' and 'slight' loss presented more difficulties in their discrimination. **Figure 7.15** presents the classified image.

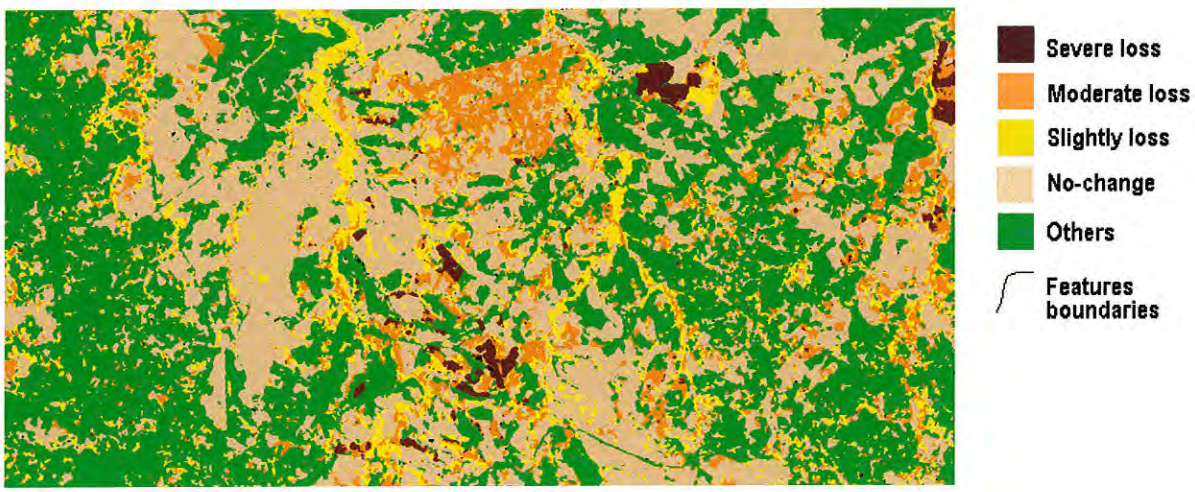


Figure 7.15 SDif.image-sh-s (supervised classification using MAXVER and Dif.image-sh)

The ability of the procedure to discriminate the change categories in the study area was measured by the producer accuracy (percentage of pixels within the training set known to have been classified correctly) and the user accuracy (percentage of pixels correctly assigned to a category for each change class). The Kappa and TAU statistics were also calculated, from the marginal totals of columns and rows of the classification matrix. The highest percentage of accuracy was achieved by the category 'biomass gain'. However, similar result was obtained as the one obtained using unsupervised classifier. The shadow image was not capable of differentiating real gain of biomass from urban areas or other cover types; thus, this result is highly influenced by the misclassification. The categories 'severe loss' and 'no-change' presented almost no confusions. The area of INF was almost totally isolated in the class 'no-change', as expected. The category 'severe loss' was composed by clearcutting of commercial pine plantations and few and small patches of *Araucaria* forest. The area labelled as (10) in **Figure 7.1** was almost entirely classified in the category 'moderate loss' and the class 'slight loss' basically concentrated in the floodplain areas and riparian forests. The most remarkable confusions were observed between the spectrally adjoining classes 'slight loss' and 'moderate loss'. As mentioned before, it was difficult to differentiate them when the training stage was being performed. The same number of classes defined in the previous approaches was maintained with the objective of comparisons. The overall accuracy was high i.e., 92.38%. (Kappa value = 89.06% and TAU statistic = 90.48%). However, these results have to be analysed by accepting the fact that the shade image recognises only classes related to deforestation or 'no-change', joining other cover types under the same class. The classification error matrix is presented in the **Appendix 7.1b**.

7.6 Agreement analysis

Three different approaches were presented with the objective of categorising losses of forest vegetation: (1) image differencing classification using a supervised classifier and TM band 5 (SDif.image5-s); (2) an alternative process of change detection using unsupervised classification and the shade fraction image (SDif.image-sh-u); and (3) a supervised procedure involving the same shade image (SDif.image-sh-s).

Table 7.3 presents a cross tabulation matrix comparing area of agreement in hectares among the three approaches, two by two. The table reflects how each approach classified each category and how they agreed (grey cells) classifying areas in the same class; a class called background was included to accommodate the non-classified pixels. One image is read in the rows and the others in the columns. The cross area shows the area of agreement, in hectares.

Table 7.3 Cross tabulation between each two image differencing classification approaches

(1) SDif.image-sh-u (rows)			SDif.image5-s (columns)		(Values in ha)	
	(a)	(b)	(c)	(d)	(e)	(f)
(a)	0	0	0	0	0	0
(b)	40.32	567.27	406.71	178.47	21.96	1.71
(c)	123.30	497.73	2464.65	3431.88	1630.62	75.78
(d)	79.74	89.19	907.92	4045.05	4627.35	433.26
(e)	70.11	37.35	673.92	5487.21	22641.39	1547.73
(f)	277.29	114.93	2816.64	13075.29	10745.64	14527.35
(2) SDif.image-sh-u (rows)			**SDif.image-sh-s (columns)**			
	(a)	(b)	(c)	(d)	(e)	(f)
(a)	0	0	0.45	1.44	1.17	45.45
(b)	2.79	664.20	162.81	8.28	5.04	6.39
(c)	2.52	264.42	6399.18	1.98	1.80	5.85
(d)	0.99	287.28	14.13	2431.62	463.32	404.37
(e)	9.27	94.41	2067.21	7044.48	27596.34	2152.62
(f)	37.44	11.34	199.62	466.02	2430.54	38349.99
(3) SDif.image5s (rows)			**Sdif.image-sh-s (columns)**			
	(a)	(b)	(c)	(d)	(e)	(f)
(a)	0.27	1.26	0.72	14.31	13.50	22.95
(b)	42.48	501.93	458.01	261.72	49.05	8.46
(c)	127.53	565.29	2460.60	3717.54	1844.01	128.43
(d)	80.19	85.32	951.39	4265.91	4196.70	374.31
(e)	36.00	12.33	285.75	4463.73	23679.90	2020.50
(f)	304.29	137.34	3113.37	13494.69	9883.80	14031.18

(a) background; (b) 'severe loss'; (c) 'moderate loss'; (d) 'slight loss'; (e) 'no-change'; and (f) 'biomass gain'

Table 7.4 presents a numeric visualisation table of how the classification systems agreed when labelling each class.

Table 7.4 Agreement among the three image differencing classification approaches

(1) Sdif.image-sh-u x SDif.image5-s Overall agreement: 48.28%		
	SDif.image-sh-u (%)	SDif.image5-s (%)
Severe loss	43.52	46.63
Moderate loss	33.90	29.99
Slight loss	15.43	39.73
No-change	57.08	74.34
Gain of biomass	87.59	34.96
(2) SDif.image-sh-u x SDif.image-sh-s Overall agreement: 75.36%		
	SDif.image-sh-u (%)	SDif.image-sh-s (%)
Severe loss	78.19	50.26
Moderate loss	90.96	90.96
Slight loss	24.35	24.43
No-change	70.82	90.49
Gain of biomass	92.42	93.62
(3) SDif.image5s x Sdif.image-sh-s Overall agreement: 49.04%		
	SDif.image5-s (%)	SDif.image-sh-s (%)
Severe loss	37.98	38.51
Moderate loss	27.82	33.85
Slight loss	42.86	16.27
No-change	77.64	59.70
Gain of biomass	34.35	84.60

Analysing **Table 7.4**, for the first group: (SDif.image-sh-u x SDif.image5-s) the agreement percentage between the unsupervised approach using the shade image and the supervised approach using band 5 were compared. The overall agreement occurred in less than 50% of the pixels. The table is interpreted by checking how close both approaches classified each class. Using the first line as an example: SDif.image-sh-u classified 43.52% of the pixels as 'severe loss' and SDif.image5-s classified 46.63% of the pixels in the same class. Even when it can be considered a good agreement, it does not means that the same pixels were considered in both cases. In this group, the highest disagreement occurred in the class gain of biomass. This might be explained by the fact that in the shade image features other than deforestation were included in this class. In TM band 5, however, they were sometimes considered biomass loss, sometimes biomass gain, according to the dynamics of the region urban and rural areas. The second group was composed by the two approaches (supervised and unsupervised) over the shade image. The class slight loss presented similar percentages in both cases, and the class gain of biomass as well. This class (biomass gain) reflected the high level of misclassification occurred when the shade image was used independent of the classifier used. The third group compared both supervised approaches, but over different background images (SDif.image5s x Sdif.image-sh-s). Here the main difference was the already mentioned discrepancies found when shade image and band 5 were involved.

Figure 7.16 presents a view of an internal dirt road with natural patches of *Araucaria* forest on both sides at the INF. In general, no changes were expected unless the changes related to the process of forest succession dynamics. The internal roads are kept without vegetation to facilitate movement inside the area as a precaution against forest fires.



Figure 7.16 View of *Araucaria* forest patch inside the INF

7. 7 CONCLUSIONS

One of the current preoccupations about public areas is defining how the changes in the natural environment are monitored over periods of time (Ferguson, 1996). Irati National Forest is located in an area where the urban pressure is increasing and it maintains one of the last large patches of *Araucaria* forest. Monitoring changes in the protected area (INF) and adjacent area is becoming increasingly important for the success of the regional forest conservation programme.

In this context, the objectives of this chapter were: a) to determine the magnitude of any changes in the forest types representing *Araucaria* forest region in the period of 10 years from 1987 to 1997 in the area of INF and adjacencies; and b) to develop a method by which future changes in *Araucaria* forest as a biome and in *A.angustifolia* as a species can be monitored.

This work used two sets of TM imagery (from 20/06/1987 and 18/08/1998), bands 1 to 5, and 7 (details in **section 3.4.1.2**). Images were rectified and resampled using a bilinear interpolator. The images were co-registered to each other and DNs were converted to reflectance (and multiplied by 255 to keep the same amplitude of raw images). Atmospheric effects were reduced as part of the process of radiometric corrections.

The literature describes a large number of change detection methods. Each of the methods exhibits different degrees of complexity in their procedures and results. The central piece of this work was image differentiation, used to detect changes occurred in the forest types during the period analysed. Image differencing is defined as a direct comparison method and considered as a very intuitive procedure for change detection. The methodology involving image differencing is easy to be implemented and has an inherent meaning in terms of cover changes (Cohen *et al.* 1999). According to Singh *et al.* (1989), image differencing is the method most widely used by different projects and exhibits superior performance, when compared with other methods. It can be applied using two single bands (as applied in **sections 7.3.1.1** and **7.4.1**) or applied to previously prepared sets and eventually incorporating more sophisticated procedures, as mixture models and image segmentation (**section 7.5**).

The first procedures tested in this work, using image differencing were: a) the single subtraction of images, using band 5 for each year, as the basic procedure and b) subtraction of NDVI images (NDVI97 from NDVI87). The basic overlay procedure, using band 5 provided the necessary picture of the environmental complexity and the decisions inputs for the next steps. Deforestation was highlighted (light grey), reforestation appeared in dark grey and no-changes, in general presented a medium tone of grey, giving a preliminary idea of three groups that should be characterised as: a) positive changes

in biomass; b) negative changes in biomass; and c) no-changes in biomass. Six bands were available for each time and the inherent multispectral ability would contribute to the analysis; actually, it is expected that the use of transforms should improve the change detection analysis. NDVI was chosen to perform this step because of its strong relationship with green biomass. To increase the NDVI capability of distinguishing between true change and natural variability a method using thresholds was applied. The process creates an "automatic decision process" to define what is 'change' and what is 'no-change', based on some pre-defined criteria, involving limits for each class. The central idea is that the values of a difference image follow a normal distribution (confirmed for the NDVI difference image) around the mean and the standard deviation (std.dv.) would indicate the lower and upper threshold values. The number of standard deviations to be used with the mean depends on the degree of suspected change. In the present work, by a "trial-and-error" procedure, it was found that mean plus or minus one std.dv. isolated 'negative' and 'positive' changes. The decision was made by verifying how the INF area was affected by the slicing process, considering its natural status of belonging to 'no-change' category. The problems encountered by using the method were: a) the results might be site specific; b) the definition of the threshold limits was somehow intuitive; c) the inclusion of the INF in the category 'no-change' was strategic but not necessarily true.

Two other variations of the image differencing method were tested, using NDVI. The first technique is called linear regression differencing and it is performed in two steps. The first stage of the regression method basically is intended to "produce" a new image of time 2 as an "adjusted" time 1 image, by regressing one against the other. It is considered one of the strongest environmental change detection approaches because it can also be effective as a pre-processing technique in change analysis, supposedly dealing with differences caused by environmental factors (offset) and calibration differences between measuring instruments (gain). When analysing the results of this first step, it was considered that: a) the procedure may be a good option to address problems of non-comparability due to sensor changes in imagery that did not suffer any calibration but not necessarily will be advisable when radiometric restored data is available; b) for future analysis regarding the usefulness of image regression, it is strongly suggested that, for this first step, samples of data not corrected radiometrically should be compared with pre-processed data; and c) samples involving only areas within the INF should be tested (this should reduce the effects of autocorrelation) and then compared with the procedure that included the whole area. In the second step of the method, the dependent image was adjusted with the regression coefficients obtained previously. A new image was thus obtained; the subtraction process was then repeated. The same threshold limits use in the **section 7.3.1.3** were applied to the generated difference image. It was visually detectable that the Dif.imageR and SDif.imageR presented larger amounts of area in the 'negative-change' class (red, in the SDif.imageR image) than the previous simple image differencing. . It was also apparent that the 'negative changes' class, in the one specific region with natural forest, was less remarkable in number and size of patches, when using the second approach. However, by the means available in the present work, it

was not possible to detect if the image adjustment collaborated to determine the best pair of images for forest change detection. It was also not possible to clarify if the differences in those areas occurred due to the impact of the coefficients in the adjusted image, to the impact of applying 1 std.dv. as threshold or any other reason and further analysis is recommended.

The Normalised Image Differencing (NID) technique is a combination of image differencing and image ratioing and it was applied to the adjusted data set. Based on visual analysis (of bands 3,4 and 5 colour composites) of both dates, it was considered that the method tended to underestimate the alterations in urban and semi-urban region of the study area and was less sensitive to few negative changes in the natural patches. As a conclusion, it would be commented that the simple image differencing is: a) easier to be implemented; b) more straightforward in terms of analysis; and c) presented high percentage of overall agreement with the other two methods.

Mather (1999) mentioned the recognition of categories as the first of the two steps in the image classification and stressed that the first requirement in the process should be to determine "*a priori*" the number and nature of the categories being recognised. It is always difficult to distinguish spectral reflectance characteristics on areas where the vegetation diversity is high. Mixed vegetation is a major source of confusion and the discrimination between disturbed and undisturbed forest is not easily achieved. However, by field observation, it was possible to determine that in some areas the high grading presented actually different levels of intensity. Some areas were over cut and some suffered intensive and selective thinning with the removal of the commercial species, with preference to the *A.angustifolia*. The assumption was that these different levels of deforestation should be captured by the adoption of more threshold levels. However, the process is not straightforward and defining the limits of the new levels only by statistics should be too artificial.

The solution appeared to incorporate the perception of the analyst in the definition of the different levels of biomass loss. For this, a classification scheme using supervised classification was tested and the difference image of band 5 was selected to be included in the process. In the fieldwork, it was possible to define three levels of forest degradation, defined as 'severe', 'intermediate (moderate)' and 'slight (described in the **section 7.4.1.1**)' changes. The training areas were then searched for each of the three levels, plus two more categories: 'no-change' and 'biomass gain'. Local foresters were consulted to indicate sites that represented typical vegetation-degradation conditions. Classification was based on the maximum likelihood decision rule (MAXVER). By the results, it was considered that TM band 5 and supervised classification are promising as tools to contribute to the monitoring of *Araucaria* forest. The differentiation between levels of degradation was achieved as a result of the training effort. The classification effectiveness involved the comparison of the categorised data to the reference set by the error matrix, with a global accuracy of 90.66%. The Kappa result was 87.80% and The TAU statistic was 88.32%.

Even with the good results obtained previously, doubts remained about the limitations of traditional image classification related to the assumption that each accuracy site can be unambiguously assigned to only one class. The defined classes might be better analysed if considered as continuous instead of discrete classes and the possible solution would be mixing models (as discussed before) or fuzzy logic. Fuzzy analysis was not implemented in this work; however, in retrospect it would seen to be a valuable tool to accurately assess changing in *Araucaria* forest and it is strongly recommended for further studies. As an alternative approach, fraction images derived from the linear mixing model were included, to test the assumption.

Considering the potential size of the area in Southern Brazil, a monitoring system to detect changes in *Araucaria* forest, has to be: a) repeatable on whatever time interval is appropriate; b) simple to carry out and not time-consuming; and c) effective in detecting changes.

The developed process involved the: a) calculation of the fraction images, derived from the Linear Mixing Model (LMM) for both analysed periods; b) calculation of the difference image, by subtracting one date from another; c) segmentation of the difference image; and d) adoption of an unsupervised classification, creating a map of the forest changes for the period. LMM is potentially useful for change detection studies because of its inherent characteristics of dimensionality reduction and the fact that one of its synthetic products, the shade image, is related to the forest canopy structure (Shimabukuro *et al.* 1998). The use of the fraction image obtained by the component *A.angustifolia* (Vt1) instead of the shade image was also investigated as well the combination of both (Vt1 and shade) to be involved in the segmentation process. The difference image of Vt1, however, did not discriminate any level of vegetation loss (it is presented in **Figure 7.12**) and was not included in the process. The segmentation process was performed over the shade difference image using the region growing algorithm and the contextual and spatial attributes of the regions were acquired by the following thresholds: similarity = 30 and area = 30. Both threshold were decided after experimental iterations by visual analysis. The algorithm used in the per-field unsupervised classification was the ISOSEG, a cluster technique. The best acceptance threshold was set to 75%, which guaranteed the classes of deforestation as 'severe loss', 'moderate loss' and 'slight loss' and the classes 'no-change' and 'biomass gain'. The resultant image highlighted the classes which are related to deforestation: 'severe loss', 'moderate loss' and 'slight loss' were located were they were expected, with an huge area near to the INF being classified as 'moderate loss'. The class 'no-change' highlighted the INF, and other two areas, expected to belong to that class. The great difference of these results, when compared to the image obtained by the supervised classification of band 5 was found in the class 'biomass gain'. The differencing shade image, due its characteristics related to vegetation canopy, did not discriminate other features, which have low shade contents. It is a typical tool to analyse deforestation classes. The image, therefore was composed by the three levels of deforestation, the class 'no-change and the class 'others', in which the others cover types were included.

A supervised classification, using MAXVER and the shade image were also performed. The classes 'moderate loss' and 'slight loss' were difficult to be labelled in the training process. The resultant image, however, presented similarities to the one obtained by unsupervised classification, with differences mainly in the class 'slight loss'; the classifier incorporated practically all riparian forests and floodplain areas in this class. In relation to the other cover types which are not related to deforestation, the same result obtained in the unsupervised approach was also obtained in the supervised one; due to the characteristics of the shade image, it does not differentiate urban, semi-urban or agriculture areas from young forest plantations. This is not a fatal restriction, however. The main focus of the study is related to forest degradation and both supervised and unsupervised approaches, as performed in this study delineated successfully deforestation classes.

It was concluded that the use of shade differencing image, image segmentation and a per-field unsupervised classification resulted in an alternative methodology for detecting deforestation in *Araucaria* forest. It should be a good option for situations as rapid assessments as well when powerful computational resources are not available, important aspects in the monitoring of this large and complex area. The methodology must be tested in other areas of the same biome and minor alterations might be necessary. Further research in the same direction is strongly recommended.

CHAPTER 8

CONCLUSIONS AND RECOMMENDATIONS

8.1 Introduction

Southern Brazil is one of the most developed regions of the country. All over the western world, the price to be paid for this model of development is usually environmental degradation. The region is not an exception, as it has experienced a rapid deterioration of its forest resources, mainly due to the impacts of explorative land use practices. Mixed Ombrophilous Forest with *Araucaria angustifolia* (Bertoloni) O. Kuntze (Veloso and Goes-Filho), abbreviated to *Araucaria* forest, were so imposing and extensive that both local and foreign visitors considered them 'inexhaustible' (Carneiro 1979). However, the forest inventory research project that was carried out in southern Brazil in 1978 (FUPEF 1978) estimated that only 3% of the original area remained, with just 0.7% being considered primary areas. Now, 22 years later, possibly less than 1% of the original *Araucaria* forest remains in Paraná State, around 6% in Santa Catarina State and approximately 1.5% in Rio Grande do Sul State, the three states that are called "*Araucaria* forest region".

Considering this scenario of rapid disappearance of forest resources, it should be desirable to establish as many reserves as possible. Despite the efforts of conservationists, however, the Paraná State Conservation Units currently totals only 2.41% of the State total surface and the Irati National Forest (INF) is the only federal area with remnants of *Araucaria* forest in this State (Jacobs 1997).

The present study was undertaken to evaluate the potential of digital image processing methods to characterise, over space and time, forest status in the Irati National Forest (INF) and neighbouring areas. The area might be considered under protection. Actually, there are no areas of *Araucaria* forest that are effectively protected in Paraná State because in reality the INF is a Federal Productive Unit, where sustainable management is allowed, and eventually expected. Officially, however, no selective or clear cutting was performed in the natural forest in the last 50 years. The field plots used for the Forest Inventory of the INF were used in this study to analyse the relationship between structural information and spectral response of four main forest types: a) *Pinus elliottii* Engelm. – slash pine; b) *Pinus taeda* L. – loblolly pine; c) *Araucaria angustifolia* plantation; and d) *Araucaria* natural growth forest..

8.2 Conclusions

The possibility of using multi-temporal data sets for assessing and monitoring forested areas has transformed RS, its products (satellite imagery or aerial photographs) and techniques in powerful allied of managers and conservationists. Archives of such imagery are available in many public institutions and some of them are free of charge, easily captured over the Internet. The first impression of ecologists, foresters and people involved with managing landscape might be that finally this is the solution for analysing past and present concomitantly and even predict future based on past patterns. However, comparison of images taken at different times is not straightforward. The signal received by the sensors is determined by a range of factors, not only by the ground surface reflectance, distorting somehow the captured image. The main factors that perturb the signal can be described as: a) atmospheric conditions; and b) differences in view and illumination geometry. As a third factor, it is also worthy to mention variations in sensor characteristics through time (degradation of TM visible bands, mainly band 1, for example) as potential sources of inaccuracies.

The first preoccupation of this study was to recognise the importance of data sets corrections in order to remove or at least to diminish those effects. The best arsenal of techniques should be applied to the imagery to guarantee or to improve the probability of multi-temporal sets appear as if they were acquired under the same conditions. In this work, **Chapter 3** highlighted the available approaches for making the data sets comparable for quantitative assessment of forest changes within the INF and adjacent areas. In the pre-processing phase, some methods are fundamental and necessary to any data set; others techniques, however, are specific to some objectives. The processing methods that were chosen to build a robust and reliable data set for change detection purposes were:

1) Geometric restoration (resampling, georeferencing and geometric registration of image to image): in the present study, two dates were analysed, but the procedure of image co-registration should be valid for time series as long they are registered to one of them, called "reference image". One of the basic recognition was that the accuracy of change detection process depends on the accuracy with which the images were registered to each other. In the present work, the resultant RMS was smaller than one pixel (0.486 or 14.58 m approximately). It was considered adequate for the purposes of the study, following the precaution suggested by Lunetta (1999), for whom the RMS error between two dates should not exceed 0.5 pixels, for environmental monitoring projects.

2) Radiometric restoration (conversion of digital numbers to radiance and conversion of radiance to reflectance) were performed in order to convert DNs to physical values and

consequently, allow the reduction of errors in the data analysis. Reflectance values were multiplied by 255 to present the same comprehensiveness of the original digital values.

3) Variations in atmospheric conditions (which affect the transmittance of surface reflected radiance) are a major concern in multi-temporal analysis and a large number of approaches are proposed in the literature for correcting their effects. The analysis of the conditions of the available satellite imagery (minimum cloud cover, high visibility, small seasonal land cover differences and topographical aspects of the area) contributed to the decision of calibrating the images using radiometric rectification, as proposed by Hall *et al.* (1991). The major difference between radiometric rectification and other atmospheric corrections methods is that it does not require sensor calibration or atmospheric turgidity data, and thus corrects images from a common scene in a relative, rather than an absolute sense. It is a laborious procedure but visual analysis corroborated the adequate correction of the images. However, the success of the method depends on the availability of dark and bright pseudo-invariant objects in the image data, what sometimes is a restriction. Considering the importance of the atmosphere in change detection studies, for future research is recommended the application of an absolute model for calibration data or the comparison of both procedures.

The studied region is the only region in the country where natural conifers (but also planted) grow up contiguously to deciduous species. The representative conifers in the INF are: *Araucaria angustifolia, Pinus elliottii* and *Pinus taeda* plantations. Representing indigenous forest is *Araucaria* forest, composed of almost a hundred deciduous woody species, but also being the natural environment to the conifer *A.angustifolia*. With the objective of characterising differences between these main forest types, three basic aspects were studied in **Chapter 4**: a) spectral responses and structural variables relationships; b) spectral responses and species composition; and c) spectral characteristics of each feature and their separateness. To accomplish the objectives, three experiments were set up and the results are:

1) *P.elliottii* presented brighter values than *P.taeda* and the differences in texture are identifiable in the mid-infrared region of the spectrum. Interestingly, *P.taeda* presented spectral responses that, under some aspects, can be more easily confounded with the responses of *A.angustifolia* than with the ones from *P.elliottii*. For pines, 49 plots were measured in the field and their co-ordinates recorded by GPS. TM band 5 (mid infrared region 1.55-1.75 μm) was the single band that showed the largest correlation against ground variables (basal area and volume). Its superiority was also verified against band transformations (vegetation indices). Forward stepwise linear regression procedure, including the TM single bands (bands 1 to 5) and six statistically significant transformed images was performed and the only significant variable was again band 5. This

contribution should be very useful for the Management Plan of the INF and might be adopted by the forest farmers of the region. The results are valid for mature stands. Future research should involve young stands and a sampling design that incorporates stratification by age. Data, however, must be collected outside the INF because the number of young stands available in the protected area is insufficient for statistical purposes.

2) The available field data could not be used to determine correlations between structural and spectral variables of *A.angustifolia* plantation as done for pine. However, samples acquired at stand level presented the tendency of confirming the correlation between basal area or volume and mainly band 5. Band 7 (mid infrared region 2.08-2.35 μm) is a key band in the differentiation of conifers and deciduous in the region. *A.angustifolia* plantation represents the coniferous group, but some stands presented natural regeneration of deciduous species in the understorey, as captured by band 7. Future experiments would involve some stratification in the design, by grouping stands which still presents characteristics of monoculture separated from those where the regeneration process is so established that the area is actually almost so mixed as a natural patch.

3) *Araucaria* forest in the INF is not continuous. In this work, it was divided into four patches and 93 plots were analysed on their percentage of *A.angustifolia* (basal area, volume or number of trees) against the same variables for the other 82 inventoried species together. The field design was systematic sampling and the location of the plots was recorded by GPS. There was a special interest in the analysis of the possible correlation between DBH (and consequently, basal area) and spectral responses, because of the straight correlation between DBH and crown parameters. The highest correlation coefficients between DBH and its derived variables (basal area, crown diameter and crown area) and spectral responses were against bands 5 and 7 (correlation coefficient around 0.5). The analysis of band 7 pattern for the natural stands highlighted its sensitivity to the presence of conifers (*A.angustifolia*) in the natural patches of the INF. Other analysis involving the equality of the four patch means showed that when the variable crown diameter was analysed, differences between patches were found, showing a lower concentration of *A.angustifolia* in patch 1. It was concluded that there is a difference in basal area (and its related variable crown diameter) of paraná-pine between the sampled areas and that, this difference can be captured by TM band 5. The sampling methodology designed as systematic is adequate for the purpose of analysing natural forests. However, future research is strongly recommended establishing plots as transects. The transects should have width equal two or three pixels, to be capable of capturing eventual concentrations of *A.angustifolia*. All trees of *A.angustifolia* should be measured and their concentration recorded by GPS. The transects should follow one direction (for example, north-south)

and cut the whole patch. This type of strategy certainly will contribute to confirm and extend some of the results obtained by this work.

Finally, **Chapter 4** presents the typical reflectance curves for the main forest types and for the non-vegetated areas in the INF. In the case of the forest cover types, the bands of the infrared region of the spectrum were fundamental in distinguishing conifers and deciduous species.

Chapter 5 continued the analysis started in **Chapter 4** to evaluate the fundamental aspects of spectral feature and dimensionality of the data, in a multispectral environment (Landsat TM 5). The preceding chapter identified the importance of the bands of the infrared region for the discrimination between coniferous and deciduous forests. **Chapter 5** analysed the dimensionality reduction of data (from six bands to two or three) without losing desirable information. The main results can by summarised in:

1) The RATIO, in its traditional form of dividing spectral information from the visible red (0.63-0.69 μm) by the near infrared (0.76-0.90 μm) was firstly tested and did not contribute to the objective of the work. However, intriguing results related to the discrimination of *A.angustifolia* in natural forests were obtained by the division of the near infrared band 4 (0.76-0.90 μm) by mid infrared band 7 (2.08-2.35 μm), labelled as ratio47. Band 4 is also known as "band of vegetation" and band 7 was already detected as important in the discrimination between conifers and deciduous. The results however, must be treated with reserve because they can be site-specific; additionally, the field methodology must be refined in order to confirm those results. NDVI was calculated and vegetated areas and non-vegetated areas were recognised. However, the expected ability on minimising topographic effects was not achieved. In addition, the recognised ability of the index in the estimation of woody biomass was not accomplished and it was concluded that NDVI should be better explored when the comparisons were made over time, which was studied in **Chapter 7**.

2) The adoption of NDVI or RATIO might be eventually successful, but in the case of TM, four bands are still excluded from the process. PCA is one of the recommended technique for reducing or removing duplication or redundancy in multispectral images and for compressing all of the information that is contained in the original channels to their principal components. Three different approaches were designed: a) standard PCA (all six bands as inputs) using the non-standardised procedure (covariances matrix); b) selective PCA – approach when high correlation exists between bands; and c) selective PCA – approach when low correlation exists between bands. Standard PCA revealed that the spectral data could be reduced to two or three eigenvectors. The infrared bands accumulated approximately 88% of total variance and consequently they dominated the

highest-orders principal components. Coniferous stands were clearly darker than the other type of vegetation in the first principal component, PC1. PC3 represented only around 2% of the variability but the image was strongly dominated by vegetation. Clear discrimination between the two species of Pine (slash and loblolly pine), and between *Araucaria angustifolia* (plantation) and *Araucaria* forest were achieved. A diminishing influence of topographical conditions was noted. PC3 appeared much more flat than PC1 and PC2. The selective PCA with highly correlated bands did not bring advantages when compared with the standard PCA. However, the most remarkable results were achieved with TM4 and TM7 together. The second PC of this combination presented the same tendency noted when ratio47 was analysed, a possible discrimination between conifers and non-conifers identified by tone and texture. Differences were noted also in tone and texture between *Araucaria* forest and other natural forested areas. However, besides these results, some points of the *Araucaria* forest appeared darker than the others did. This should indicate any kind of concentration of conifers, which was checked in the analysis of aerial photographs at a scale of 1:25,000 of the same area and by the analysis of field plots. An analysis of the phenomenon was performed and presented in **Appendix 7.2**. Finally, the wetness image from tasselled cap (TC) was considered a reasonable approach to be explored in future studies using data from *Araucaria* forest.

The approaches performed in the **Chapters 4** and **5** analysed information at pixel-level. This might be appropriate for large areas of crops or extensive areas with deep clear water but eventually it may be inappropriate for more complex environments. **Chapter 6** presents a review of the computer-assisted methods for handling mixed pixels (assuming linear mixing) and presents an assessment of the usefulness or effectiveness of the Linear Mixing Model (LMM) for detecting and identifying targets at the sub-pixel level. In the present work, LMM was applied with the objectives of: a) digital enhancement of phyto-physionomic characteristics from the INF; b) analysis of its capability for detecting the species *Araucaria angustifolia*, in plantation and natural stands; and c) analysis of its usefulness in the change detection process, which is discussed in **Chapter 7**. It deals with the multitemporal aspect of the problem. The philosophy of the method is that an object can be detected by its spectral properties (or by a combination of spectral and spatial properties), even if it cannot be detected by its spatial properties, i.e., when the object is smaller than the pixel dimension. However, there are two strangleholds in the process, which have to be discussed and considered before the application of any of the numerous approaches available to unmix the mixture. The first problem is related to the adequate number of end-members to be included in the model. The second constraint is related to questioning about "where and how" to capture the purest pixel to be included in the model. The next paragraphs present the strategy adopted in this work in relation to the two mentioned strangleholds:

1) Answering the first question, the potential number of end-members could be equal six (considering that in this approach the maximum number of end-members is equal to the number of bands involved in the study). The decision about of the number of end-members was made by finding the PCA eigenvectors or directions accounting for most of the variance in the data (**Chapter 5**). The three first PCs concentrated most of the variability in the studied data set. Their cumulative variance was 98.51%. Thus, three end-members could be considered as ideal in this study and the remaining variance included in the category "instrumental errors". However, most vegetated ecosystems are comprised of a mixture of canopy components, architecturally derived shadow and potentially exposed soil (Roberts *et al.* 1999). Because this feature was significant in many studies, a fourth component capturing soil as an end-member was alternatively also tested. The result was that the inclusion of the end-member soil diminished the error bands, but did not improve the quality of the images at least for classification purposes. Analysing the resultant images visually, it was considered that the variable soil maybe is a crucial variable when sparse vegetation, as arid zones and deserts are analysed. The forest cover of the study area is considered highly concentrated, mainly in the areas where mature forests are located. However, further analysis and a more detailed fieldwork is recommended to analyse its contribution when specific studies involving forest degradation (reduction of biomass by trees suppression) are performed.

2) The second preoccupation, related to the definition of end-members can be addressed by the following reasoning: *Araucaria angustifolia* is the only conifer growing up naturally that was inventoried in the study area and the work actually is focused on the discrimination between coniferous and deciduous woody species. Consequently, the first two end-members had to be: a) *A.angustifolia* in its pure form (pixel captured in the densest stands of plantations with caution to avoid the ones which are being invaded by deciduous species due to forest recuperation); and b) the signature of the natural *Araucaria* forest. The third element was shade. It was considered because of the architecture of the features (conical canopy for pine and rounded canopy for *A.angustifolia*) occurring in the area. An appreciation about the philosophy behind the adoption of shade as an end-member is presented in **Chapter 6**.

The number of different approaches to the unmixing problem is extensive. All algorithms proposed in this study assume that the spectral reflectance of the area can be modelled by a Linear Mixing Model (LMM). The Unconstrained Weighted Least-Square (UWLS) method was preferred to the Constrained Least-Square (CLS) approach for three reasons: a) smaller mean error per band and smaller total error; b) better visual performance; and c) because it could allow the derivation of overflow images to evaluate the fitness of the model.

According to the LMM assumptions, reflectance of a mixed area is given by the linear superposition of the reflectances of the constituent components. On the other hand, theoretically, non-linear mixing occur when individual photons interact with more than one type of material. Bosdogianni (1996) exemplified the idea mentioning that if the material is salt, the light might penetrate through the component of interest and interact with the component beneath. Borel *et al.* (1994) mentioned previous studies describing cases of non-linear mixing occurring when materials with different spectral reflectance compose a rough surface and the example was plowed field with exposed rocks or well-mixed particles of different minerals. The matter is complex. Williams (1991) concluded that the near infrared reflectance at the branch level of sugar maple is higher than the reflectance at the leaf level. Borel *et al.* (1994) proved that that effect is actually due to multiple reflections between leaves. Summarising the subject: if non-linear spectral mixing occurs always that two or more substances are mixed and the reflectance is measured, than Borel *et al.* (1994) is in the right direction when he assured that no linear combination of spectra can fully represent the measured composite scene spectra and, depending on the model adopted, the residuals can be large. This is a challenging subject and some methodology is still under investigation, theoretically and experimentally, as the radiosity model. There is no previous studies using unmixing modelling (linear or non-linear) for the *Araucaria* forest environment and the linear mixing model was successfully applied to the study area. As the results might be site specific and as non-linearity is to be considered, further research is recommended, in order to investigate these two questions: a) will the LMM be effective in others areas of *Araucaria* forest ?; and b) will non-linear models be more adequate than the LMM ?

The last contribution of this work to the spectral analysis of *Araucaria* forest in the INF and neighbouring areas (**Chapter 7**) studied the environment on a temporal basis. For this, Landsat TM 5 of two different dates (from 20/06/1987 and 18/08/1998), bands 1 to 5, and 7 were obtained. Landsat is one of the most important families of satellite imagery currently available. Together, MSS and TM can contribute enormously for monitoring programmes due their capability of retrieving retrospective data since early seventies. With the possibility of updated information each 16 days and the success of the Landsat 7 launching, it will continue being predominant, even when linked to new technologies.

The common sense "the accuracy of the change detection process will depend of the accuracy of pre-processing procedures - was accept integrally". The stages of performed image pre-processing and the precautions adopted during that phase are described in **Chapter 3** and summarised above. The objectives of **Chapter 7** were: a) to determine the magnitude of any changes in the forest types representing *Araucaria* forest in the period of 10 years, from 1987 to 1997, in the area of INF and adjacencies; and b) to develop a method by which future changes in *Araucaria* forest as a biome and in *A.angustifolia* as a species could be monitored.

Image differencing was the central piece in the strategy for forest change detection in this study. It is relatively easy to be implemented and strongly recommended in the literature. The strategy adopted was:

1) NDVI was selected due its potential contribution to the process. To increase the NDVI capability of distinguishing between true change and natural variability a method using thresholds was applied. By a "trial-and-error" procedure, it was found that mean plus or minus one std.dv. isolated 'negative' and 'positive' changes. The problems encountered by using the method were: a) the results might be site specific; b) the definition of the best threshold limits was somehow intuitive; c) the inclusion of the INF as the category 'no-change' limits was strategic.

2) Yet using NDVI, other two variations of the image differencing method were tested. The first technique, called linear regression differencing, was analysed in two stages: a) in the first stage, the regression is performed to produce a new adjusted image. It was concluded that the procedure might be a good option to address problems of non-comparability due to sensor changes in imagery that did not suffer any calibration. However, it will not necessarily be advisable when radiometric restored data is available, as the case of the present study. As a consequence, for future analysis regarding the usefulness of image regression, it is strongly suggested that, for this first stage, samples of data not corrected radiometrically should be compared with pre-processed data. It is also advisable to test sampling strategy involving only areas within the INF, with the objective of reducing effects of autocorrelation, inherent to the process; b) in the second stage, image differentiation was applied to the "new" image. However, by the means available in the present work, it was not possible to detect if the image adjustment collaborated to determine the best pair of images for forest change detection. It was also not possible to clarify if the differences in those areas occurred due to: a) the impact of the coefficients in the adjusted image; b) to the impact of applying 1 std.dv. as threshold; or c) any other reason and further analysis is recommended.

3) The second technique which was performed using NDVI was the Normalised Image Differencing (NID) which is a combination of image differencing and image ratioing; it was applied to the adjusted data set obtained as described in the previous section. Based on visual analysis (of bands 3,4 and 5 colour composites) of both dates, it was considered that the method tended to underestimate the alterations in urban and semi-urban region of the study area and was less sensitive to few negative changes in the natural patches.

4) Image differencing was firstly applied to the single band 5, with the objective of producing a picture of the environmental changing complexity. However, after the application of the

two methods described above, it was concluded that the simple image differencing is: a) easier to be implemented; b) more straightforward in terms of analysis; and c) presented high percentage of overall agreement with the other two methods.

In the field campaign, it was determined that in some areas the high grading presented, actually, different levels of intensity. Some areas could be considered overcut and others presented less intense selective thinning. By the procedures described above, however, deforestation was included in just one category, called 'negative changes'. One of the approaches to capture these different levels of disturbance was to incorporate the perception of the analyst, through supervised classification over the difference image of band 5. Five different categories were defined, three levels of forest degradation, 'no-change' and 'biomass gain'. Classification was based on the maximum likelihood decision rule (MAXVER). By the results, it was considered that TM band 5 and supervised classification are promising as tools to contribute to the monitoring of *Araucaria* forest. The differentiation between levels of degradation was achieved as a result of the training effort. The classification effectiveness involved the comparison of the categorised data to the reference set by the error matrix, with a global accuracy of 90.66%. The Kappa result was 87.80% and The TAU statistic was 88.32%.

The sequence of the study was based in the premiss that: a) traditionally, collecting data related to land cover alterations has been based on and developed for the investigation of a relatively small areas. Given the increasing need for continuous planning and management of forest reserves and extended regions, traditional methods might become inadequate and insufficient. The need for a low-cost source of information covering entire regions rather than merely sample areas is enormous; and b) even with the good results obtained previously, doubts remained about the limitations of traditional image classification related to the assumption that each accuracy site can be unambiguously assigned to only one class.

Considering the potential size of the area in Southern Brazil, a monitoring system to detect changes in *Araucaria* forest, need to be: a) repeatable on whatever time interval is appropriate; b) simple to carry out and not time-consuming; and c) effective in detecting changes. The strategy adopted two different approaches:

1) In the first approach the developed methodology consisted of the following steps: a) the calculation of the shade image, derived from the Linear Mixing Model (LMM) for both analysed periods; b) the calculation of the difference image, by subtracting one date from another; c) the segmentation of the difference image; and d) application of an unsupervised classification and creation of the forest changes map for the period. The unsupervised classification was based on a clustering algorithm called ISOSEG, which uses the covariance matrix and mean vector of the regions to estimate the classes. Even

being unsupervised, the procedure allows the analyst to define an acceptance threshold, which ultimately defines the regions size. It was found very helpful in the definition of forest degradation levels. The advantages and constraints of the methodology are detailed in **Chapter 7** and the analysis about the achievements and limitations of the method is presented specifically in is conclusion **section 7.8**. Summarising what was presented there, the major difference of the results, when compared to the image obtained by the supervised classification of band 5 was found in the class 'biomass gain'. The differencing shade image, due its characteristics related to vegetation canopy, did not differentiate other features, which have low shade contents. Reforestation, alterations in urban and semi-urban areas and agricultural crop and pasture rotation, for example, were included in the class 'biomass gain'. The image therefore was composed by the three levels of deforestation and by the class 'no-change'. Additionally, the class 'biomass gain' was renamed to 'others', in which the others cover types were included. It was concluded that the image differencing shade, when used for purposes as the ones of the present work is a useful tool to analyse differences in shadow caused, for example, by forest removal and it is consequently, sensitive to different levels of deforestation.

2) The adoption of a second approach was decided to diminish the uncertainties of the adoption of an unsupervised classification. However, before the discussion of the results obtained by that procedure, it could be important to discuss the role of image segmentation when combined with an unsupervised classification. Actually, when both techniques are combined, the results become more under control, at least in relation to the number of final classes. Image segmentation consists of partioning the image into units that are homogeneous. Generally, the segmentation algorithms are based in one of the basic properties of the grey level values: discontinuity and similarity. In the first category, the image partition is based on abrupt changes in the grey levels. Region growing is in the second category and groups homogeneous pixels or sub-regions in larger regions. The process starts with a series of "seed points" and pixels are annexed based on properties as similar grey levels and texture. The proportions, spatial and contextual attributes of the segments (regions) are somehow in the analyst control, as parameters such as similarity and minimum area have to be set. This is opposite to traditional unsupervised classification when it is not possible to control de number of categories obtained.

3) A supervised classification, using MAXVER and the shade image was designed. The classes 'moderate loss' and 'slight loss' were difficult to be labelled in the training process. The resultant image, however, presented similarities to the image obtained by unsupervised classification, with differences mainly in the class 'slight loss'; the classifier incorporated practically all riparian forests and floodplain areas in this class. Similar results were also obtained in relation to the class 'biomass gain'; due to the

characteristics of the shade image, it did not differentiate urban, semi-urban or agriculture changing from young forest plantations. This was not considered as a fatal restriction, however. The main focus of the study is related to forest degradation and both supervised and unsupervised approaches, as performed in this study delineated successfully deforestation classes. The overall accuracy was high i.e., 92.38%. (Kappa value = 89.06% and TAU statistic = 90.48%). However, these results have to be analysed by accepting the fact that the shade image recognises only classes related to deforestation or 'no-change', joining other cover types under one same class.

In this final part of the work, three different approaches were presented with the objective of categorising losses of forest vegetation: a) image differencing classification using a supervised classifier and TM band 5; b) an alternative process of change detection using unsupervised classification and the shade fraction image; and c) a supervised procedure involving the same shade image. As an additional effort in the analysis of the involved processes, a cross tabulation matrix comparing area of agreement in hectares among the three approaches, two by two, and a numeric visualisation table of how the classification systems agreed when labelling each class was presented.

The results obtained with the INF data set may probably not be sufficient for a generalisation of the procedures developed for the assessment of the spectral responses of the different forest types of the *Araucaria* forest region. They could be site specific. However and hopefully, the methodology will be applied successfully to other areas in the same biome. If results that are similar or identical to those from the INF are obtained, then the generalisation of the procedures would be suggested.

This study has analysed expert knowledge in the area of Remote Sensing of Environment. However, it does not intend to be exhaustive in any of the discussed aspects. After all, it intends to open a debate on the process of analysis of the biome *Araucaria* forest, with emphasis on the species *A.angustifolia*. Another purpose is to draw the government and public opinion to the process of devastation and degradation of *Araucaria* forest, presenting contribution on how this can be evaluated by RS and GIS procedures.

8.3 Future research directions and recommendations

Using Landsat TM5, this thesis highlighted some procedures that hopefully, will contribute to the monitoring of the *Araucaria* forest region. However, it is recognised that some of the objectives should be more easily reached if the spatial, spectral and radiometric resolutions in Landsat 5 were more adequate. The gap produced by the failure of Landsat 6 is now filled by Landsat 7 and soon SPOT-5 will be available. Brazilians will have their own remote sensing satellite system, the CBERS, a joined China-Brazil Project that is planned to be available this

year 2000. Actually, the first images are already being produced. Resolution will range from 5 to 260m (10m in the case of multispectral imagery). Including the technology like RADARSAT-1, it seems that the possibilities are being increased and a new paradigm is going to be reached. In this case, more facilities to species identification and more details in the spectral versus structural relationship of forest areas can be expected. However, new technology will always need to be linked to the available MSS and TM programmes, when change detection of long periods is the purpose.

The quantitative relationship between Landsat TM reflectance and structural variables for pine plantations is very promising as a tool for the management of the INF and eventually other forest farms. If field data are available (by re-measuring permanent sample plots or by the use of growth and yield models), it is recommended to confirm the predictability of structural variables such as volume using remote sensing. As it was showed in this research, the results are satisfactory, at least for stands where the crowns closure is less than complete.

Investigations are recommended on the subject of correlations between crown dimensions and spectral response for *A.angustifolia*, as a refinement to the procedures for assessing its presence in natural stands of *Araucaria* forest. This work used crown variables derived by regression and the confirmation of that correlation (DBH and crown diameter) in different stand densities would be valuable.

More investigations on the use of the ratioing of TM bands 4 and band 7 or other transformations based in their relationship is suggested, to confirm their grouped ability to discriminate *A.angustifolia* in the *Araucaria* forest, perhaps by grading patches with different percentage of *A.angustifolia*.

Linear Mixing Modelling is a powerful technique, for mixed pixels. However, sometimes the available image processing systems do not report the presence of under- and over-shoots. The methodology developed in this thesis to derive the overflow images was considered very useful in the analysis and it is recommended the procedure to be included in the software SPRING. In addition, some procedures, which were beyond the scope of this study, are expected to be very useful in the future studies in the same area. They should be listed as: a) the manual end-member selection method (MESM), as presented by Bateson *et al.* (1996); b) the applied analysis spectral analytical process (AASAP), as presented by Huguenin *et al.* (1997); c) the modified spectral mixture analysis (MSMA), presented by Tompkins *et al.* (1996); and d) the singular-value decomposition analysis (SVA) developed by J. Boardman in 1989 and suggested by Mather (1999).

Change detection has numerous approaches, some of them very promising when dense vegetation such as *Araucaria* forest is analysed. A suggested method could be procedures

involving PCA with one single band set or on two or more band sets together (time series analysis). By the results obtained **Chapter 5**, it is presumed that PCA has the ability to provide useful environmental change information from multi-date satellite imagery (Thiam 1997). Another method that has been tested successfully is CVA (Change Vector Analysis), which is based on the three spectral dimensions known as brightness, greenness and wetness (the Tasseled Cap - TC of Kauth *et al.* 1976; Cris *et al.* 1984). The tasselled cap wetness feature presented promising results in the analysis presented in **Chapter 5** of this work and was reported as important for separating classes of conifer forests and in conifer forest change detection (Cohen *et al.* 1999).

8.3.1 Change detection and pos-classification comparison methods

A final discussion is important to be included in this section: Is the image differencing change detection procedure a reasonable solution to the change detection of degradation in natural forests? According to Castelli *et al.* (1999), it introduces two major sources of imprecision, when it is compared to the pos-classification procedure. They can be divided in: reduction of accuracy of the classification process; and b) limited ability of providing information regarding the land cover class involved in the change. This reasoning is polemic; any review of the available literature will find successful and unsuccessful experiences when the procedure was adopted.

Therefore, would other approach using conventional classification of land cover be a better approach for change detection in the present study ? The question is difficult to be answered. The alternative approach could be the adoption of pos-classification comparison, as mentioned by Castelli *et al.* (1999). Pos-classification change detection were among the first change detection techniques applied to digital data (Yuan *et al.* 1999). It consists of two types of change detection methods, which are based on: a) two independent spectral or spectral-spatial classifications; and b) two independent true land cover class classifications. Summarising, it requires: a) two or more independent classifications, using comparable classification procedures; b) those classifications are compared, pixel by pixel; c) an image is generated when pixels are placed in different classes on the two scenes (Campbell 1996). According to Foody *et al.* (1999a) this type of approach is widely used for its simplicity. The most important aspect to be discussed when the adoption of the approach is related to the fact that the change detection accuracy is dependent of the individual accuracies. This means that some differences might not be true changes in land use, but differences in classification accuracy. Pre-processing approaches, as co-registration accuracy can have major influences on the process of change detection. One of its strengths is that it permits the compilation of a matrix of from-to changes, making possible the quantification of classes (Campbell 1996).

Two aspects were considered in the strategy adopted in the present work. The approach assumed that the traditional procedures would be inadequate to the detection of changes because: a) the objective was to detect forest degradation rates, not only clear cuttings. Forest degradation in *Araucaria* forest happens as a selective cut of the commercial species largest trees. As a consequence, the nature of this kind of change is more related to continuous classes than discrete boundaries. Different reasoning should be applied if the intention was to map different forest cover types, as pines, eucalyptus, etc. These should be clearly a discrete approach; b) conventional pos-classification comparison methods of change detection will only indicate complete changes in class membership. This means that if land cover conversion is operating at a scale finer than the spatial resolution of the sensor (in the case, Landsat TM) a mixed pixel will represent the land cover. Pos-classification approaches might be insensitive to these more gradual or subtle variations in the class properties (Foody *et al.* 1999a).

Two approaches are known as appropriate to deal with mixed pixels: linear mixing model and fuzzy classification. Foody (1999b) compared both approaches using Airbone thematic mapper (ATM) over an area of the University College of Swansea and concluded that both methodologies were able to unmix pixel composition accurately. This work adopted the first procedure, linear mixing model. However, as mentioned in **Chapter 7 section 7.8** fuzzy classification is suppose to be a valuable philosophy in the approaching of the problem, considering its potential for classification at any point along a continuum of classification fuzziness (Foody 1999b) and it is strongly recommended for further investigation.

References

Abeyta, A.M. and Franklin, J., 1998, The accuracy of vegetation stand boundaries derived from image segmentation in a desert environment. **Photogrammetric Engineering and Remote Sensing**, (64(1):59-66.

Adams, J.B., Smith, M.O., and Gillespie, A.R., 1989, Simple models for complex natural surfaces: a strategy for the hyperspectral era of remote sensing. In: **Proceedings of International Geoscience and Remote Sensing Symposium (IGARSS'89)** Twelfth Canadian Symposium on Remote Sensing, 1, pp.6-21.

Adams, J.B., Sabol, D.E., Kapos, V., Almeida F.,R., Roberts, D.A., Smith, M.O., and Gillespie, A.R., 1995, Classification of multispectral images based on fractions of endmembers: application to land-cover change in the Brazilian Amazon. **Remote Sensing of Environment**, 52:137-154.

Aguiar, A.P.D., Shimabukuro, Y.E., and Mascarenhas, N.D.A., 1999, Use of synthetic bands derived from mixing models in the multispectral classification of remote sensing images. **International Journal of Remote Sensing**, 20(4):647-657.

Ahern, F.J., Erdle, T., MacLean, D.A., and Kneppeck, I.D., 1991, A quantitative relationship between forest growth rates and Thematic Mapper reflectance measurements. **International Journal of Remote Sensing**, 12(3):387-400.

Aldred, A.H., and Sayn-Wittgenstein, L., 1972, Tree diameters and volume from large-scale aerial photographs. **Information Report**, Forest Management Institute, Ottawa. (FMR-X-40).

Aldred, A.H., 1976, Measurement of tropical trees on large-scale aerial photographs. **Information Report**, Forest Management Institute, Ottawa. (FMR - X-86).

Alves, D.S., Moreira, J.C., Kalil, E.M., Soares,J.V., Fernandez,O., Almeida, S., Ortiz, J.D., and Amaral. S., 1996, Mapeamento do uso da terra em Rondônia utilizando técnicas de segmentação e classificação de imagens TM. In: **Proceedings of VIII Simpósio Brasileiro de Sensoriamento Remoto**. Salvador, Brazil, April 14-19 1996 - CD-Rom.

Alves, D.S., Pereira, J.L.G., Sousa, C.L., Soares, J.V., Moreira, J.C., Ortiz, J.O., Shimabukuro, Y.E., and Yamaguchi, F., 1998(a), Análise comparativa de técnicas de classificação de imagens do sensor Landsat/TM para caracterização de áreas desflorestadas. In: **Proceedings of IX Simpósio Brasileiro de Sensoriamento Remoto**. Santos, Brazil, 11-18 Sep.1998 - CD-Rom.

Alves, D.S., Pereira, J.L.G., Souza, C.L., Soares, J.V., and Yamaguchi, F., 1998(b), Classification of the deforested area in central Rondônia using TM imagery. In: **Proceedings of IX Simpósio Brasileiro de Sensoriamento Remoto**. Santos, Brazil, 11-18 Sep.1998 - CD-Rom.

Alves, D.S., Pereira, J.L.G., Sousa, C.L.de, Soares, J.V., and Yamaguchi, F., 1999, Characterizing landscape changes in central Rondônia using Landsat TM imagery. **International Journal of Remote Sensing**, 20(14):2877-2882.

Anderson, G.L., Hanson, J.D., and Haas, R.H., 1993, Evaluating Landsat Thematic Mapper derived vegetation indices for estimating above-ground biomass on semiarid rangelands. **Remote Sensing of Environment**, 45:165-175.

Ardö, J., 1992, Volume quantification of coniferous forest compartments using spectral radiance recorded by Landsat Thematic Mapper. **International Journal of Remote Sensing**, 13(9):1779-1786.

Avery, G., 1958, Composite aerial volume table for southern pines and hardwoods. **Journal of Forestry,**10:741-45.

Barrett, E.C., and Curtis, L.F., 1999, ***Introduction to Environmental Remote Sensing***, Stanley Thornes (Publishers) Ltd, Cheltenham, 4th ed.

Bateson, A., and Curtiss, B., 1996, A method for manual endmember selection and spectral unmixing. **Remote Sensing of Environment**, 55:229-243.

Beauchemin, M., and Thomson, K.P.B., 1997, The evaluation of segmentation results and the overlapping area matrix. **International Journal of Remote Sensing**, 18(18):3895-3899.

Bernardes, S., 1996, **Índices de vegetação e valores de proporção na caracterização de floresta primária e estádios sucessionais na área de influência da Floresta de Tapajós-Estado do Pará**. INPE, S.José dos Campos. M.Sc. Dissertation.

Boardman, J., 1989, Inversion of imaging spectrometer data using singular value decomposition. In: **Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS'89)**, Vancouver, British Columbia, Canada, 10-14 July 1989.

Borel, C.C., and Gerstl, S.A.W., 1994, Nonlinear spectral mixing models for vegetative and soil surfaces. **Remote Sensing of Environment**, 47:403-426.

Bosdogianni, P., 1996, **Mixed pixel classification in Remote Sensing**. Department of Electronic and Electrical Engineering, University of Surrey, UK. Ph.D. Thesis.

Brown Jr., and Brown, G.G., 1994, Habitat alteration and species loss in Brazilian forests. In: **Tropical Deforestation and Species Extinction**. The IUCN Forest Conservation Programme. Chapman & Hall, London, UK. 2nd ed.Edited by T.C.Whitmore and J.A.Sayer.

Campbell, J.B., 1996, ***Introduction to Remote Sensing***. Taylor & Francis, London, 2nd ed.

Campos, J.B., 1996, Unidades de Conservação no Estado do Paraná – Ações e Contradições. **IF Série Registros**, 17:1-11.

Carneiro, N., 1979, Foreword In: ***Bibliography of the genus Araucaria***, Bibliografia do gênero Araucaria, V.1. Curitiba, Fundação de Pesquisas Florestais do Paraná-FUPEF.

Carr, J.R., and Matanawi. K., 1999, Correspondent analysis for principal components transformation of multispectral and hyperspectral digital images. **Photogrammetric Engineering and Remote Sensing**, 65(8):909-914.

Casteli, V., Elvidge, C.D., Li, C., and Turek, J.J, 1999, Classification-based change detection: theeory and applications to the NALC data set. In: ***Remote Sensing change detection – Environment monitoring methods and applications***, Chapter 4. Taylor & Francis, London.

CCRS - Canada Centre for Remote Sensing, 1999, **Fundamentals of Remote Sensing – Tutorial**. Ottawa, Ontario, Canada.

CSES - Center for the Study of Earth from Space, 1996, **Atmosphere Removal Program (ATREM) User's Guide**, Version 2.0.

Chang, C. and Du, Qian, 1999, Interference and noise-adjusted principal component analysis. **IEEE Transactions on Geoscience and Remote Sensing**, 37(5):2387-2396.

Chavez, P.S., Jr., 1975, Atmospheric, Solar, and M.T.F. Corrections for ERTS Digital Imagery. **Proceedings, American Society of Photogrammetry**, p.69-69a.

Chavez, P.S., Jr., and Kwarteng, A.Y., 1989, Extracting spectral contrast in Landsat Thematic Mapper Image Data using selective Principal Component Analysis. **Photogrammetric Engineering and Remote Sensing**, 55(3):339-348.

Chavez, P.S., 1992, Comparison of spatial variability in visible and near-infrared spectral images. **Photogrammetric Engineering and Remote Sensing**, 58:957-964.

Chavez, P.S., Jr., and MacKinnon, D.J., 1994, Automatic detection of vegetation changes in the southwestern United States using remotely sensed images. **Photogrammetric Engineering and Remote Sensing**, 60(5):571-583.

Chen, S.C. and Herz, R., 1996, Estudos quantitativos e calibração radiométrica de dados digitais do Landsat-5. In: **Proceedings of VIII Simpósio Brasileiro de Sensoriamento Remoto**. Salvador, Brazil, April 14-19 1996 - CD-Rom.

Cohen, W.B., and Fiorella, M., 1999, **Comparison of methods for detecting conifer forest change with Thematic Mapper imagery**. In: Remote Sensing change detection – Environment monitoring methods and applications. Edited by R.S.Lunetta and C.D.Elvidge. Taylor & Francis, London.

Congalton, R.G., and Mead, R.A., 1983, A quantitative method to test for consistency and correctness in photointerpretation. **Photogrammetric Engineering and Remote Sensing**, 4(1):69-74.

Crist, E.P. and Cicone, R.C., 1984, A physically-based transformation of Thematic Mapper data – the TM Tasseled Cap. **IEEE Transactions on Geoscience and Remote Sensing**, 22, 256-263.

Crosta, A.P., and McM.Moore, J., 1989. **Enhancement of Landsat Thematic Mapper Imagery for Residual Soil Mapping in SW Minas Gerais State, Brazil: A Prospecting Case History in Greenstone Belt Terrain.** *Proceedings of the 7th* (ERIM) *Thematic Conference: Remote Sensing for Exploration Geology. Calgary, 2-6 Oct, pp.1173-1187.*

Crosta, Á.P., 1992, ***Processamento digital de imagens de sensoriamento remoto***. Ed.rev.Campinas, Ed.UNICAMP.

Dawkins, H.C., 1963, Crown diameters: their relation to bole diameter in tropical forest trees. **Commonwealth Forest Review**, 42(4):318-333.

De Wulf R.R., Goossens, R.E., De Roover, B.P., and Borry, F.C.,1990, Extraction of forest stand parameters from panchromatic and multispectral SPOT-1 data. **International Journal of Remote Sensing**, 11:1571-1588.

Dillenwijn, F.S., 1966, ***Inventário do Pinheiro no Paraná***. Curitiba, Paraná. BADEP/CERENA.

Disperati, A.A., 1981, **The mapping of stands of Paraná pine (*Araucaria angustifolia* (Bert.) O. Ktze.) in the forest of south-west Paraná State (Brazil) using computer-aided analysis of Landsat MSS data.** Geography Department, Bedford College, University of London. D.Phil. Thesis.

Eastman, J.R., and McKendry, J.E., 1991, ***Change and Time Series Analysis***. Series: Explorations in Geographic Information Systems Technology – Volume 1. United Nations Institute for Training and Research - UNITAR European Office, Geneva, Switzerland.

Eastman, J.R., 1997, **Idrisi for Windows – User's Guide Version 2.0**. Clark Labs for Cartographic Technology and Geographic Analysis, Clark University, Worcester, MA.

Ehrlich, R., and Full, W.E., 1987, Sorting out geology-unmixing mixtures. In: **Use and Abuse of Statistical Methods in Earth Sciences**. Edited by W.B.Size. Oxford University Press, N.York.

Epiphanio, J.C.N., Almeida Jr., A.C.de, and Formaggio, A.R., 1997, Wheat development evaluated by remote sensing using two vegetation indices. **Anais da Academia Brasileira de Ciências**, 69(4):471-478.

Ferguson, I.S., 1996, ***Sustainable Forest Management***. Oxford University Press, Melbourne, Australia.

Foody, G.M., 1992, On the compensation for chance agreement in image classification accuracy assessment. **Photogrammetric Engineering and Remote Sensing**, 58(10):1459-1460.

Foody, G.M., and Cox, D.P., 1994, Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions. **International Journal of Remote Sensing**, 15(3):619-631.

Foody, G.M., and Boyd, D.S., 1999a, Detection of partial land cover change associated with the migration of inter-class transitional zones. **International Journal of Remote Sensing**, 20(14):2723-2740.

Foody, G.M., 1999b, The continuum of classification fuzziness in thematic mapping. **Photogrammetric Engineering and Remote Sensing**, 65(4):443-451.

Foschi, P.G., and Smith, D.K., 1997, Detecting sub-pixel woody vegetation in digital imagery using two artificial intelligence approaches. **Photogrammetric Engineering and Remote Sensing**, 63(5):493-500.

Frohn, R.C., 1998, ***Remote Sensing for Landscape Ecology: new metrics indicators for monitoring, modelling, and assessment of ecosystems***. Lewis Publishers, Boca Raton, Florida, USA.

FUPEF – Fundação de Pesquisas Florestais, 1978, ***Inventário do Pinheiro no Sul do Brasil***. Relatório Final. Curitiba, Paraná.

Fung, T., and LeDrew, E., 1987, Application of principal component analysis to change detection. **Photogrammetric Engineering and Remote Sensing**, 53(12): 1649-1658.

Galvão, F., Kuniyoshi, Y.S., and Roderjan, C.V. 1989. Levantamento fitossociológico das principais associações arbóreas da Floresta Nacional de Irati-PR. **Floresta**,Vol.XIX (1 and 2):30-49.

Gantzel, O.L., 1979, **Avaliação das florestas de *Araucaria angustifolia* (Bert.) O. Ktze. do sul do Brasil através de imagens do satélite Lansat 2**. Curitiba, Paraná, Brazil, Setor de Ciências Agrárias, Universidade Federal do Paraná. M.Sc.Disseration.

Gillespie, A.R., Smith, M.O., and Sabol, D.E., 1990, A field measure of the "shade" fraction. In: **Proceedings of Airbone Science Workshop. Pasadena, California**

Golfari, L. 1971, Coníferas aptas para reflorestamento nos Estados do Paraná, Santa Catarina e Rio Grande do Sul. **Boletim Técnico IBDF** n.1.

Gougeon, F.A., 1999, Automatic individual tree crown delineation using a valley-following algorithm and a rule-based system. In: **Proceedings of International Forum – Automated interpretation of high spatial resolution digital imagery for forestry**. Victoria, British Columbia, Canada, February 10-12 1998.

Govaertz, Y.M., Verstraete, M.M, Pinty, B., and Gobron, N., 1999, Designing optimal spectral indices: a feasibility and proof of concept study. **International Journal of Remote Sensing**, 20(9):1853-1873.

Gubert Filho, F.A., and Oliveira, J.D.de, 1991, Proposta do Sistema Estadual de Unidades de Conservação e áreas de Uso Regulamentado. **In: Anais do III Congresso Florestal e do Meio Ambiente do Paraná**. Curitiba, APEF/APRE/SEAB/SEMA/SEIC.

Guyot, G., and Gu, X.F., 1994, Effect of radiometric corrections on NDVI-Determined from SPOT-HRV and Landsat-TM data. **Remote Sensing of Environment**, 49: 169-180.

Hadjimitsis, D.G., Clayton, C.R.I., and Hope, V.S., 1999(a), The application of covariance matrix atmospheric correction method in the remote sensing of reservoirs. In: **Proceedings of the RSS99 – Earth Observation – From Data to Information**. Cardiff University and University of Wales, Swansea, September 8-10 1999.

Hadjimitsis, D.G., Clayton, C.R.I.,Hope, V.S., and Retalis, A., 1999(b), A new method of removing atmospheric effects from Landsat TM images using pseuso-invariant targets. In: **Proceedings of the RSS99 – Earth Observation – From Data to Information**. Cardiff University and University of Wales, Swansea, September 8-10 1999.

Hall, F.G., Shimabukuro, Y.E., and Huemmrich, K.F., 1995, Remote sensing of forest biophysical structure using mixture decomposition and geometric reflectance models. **Ecological applications**, 5(4):993-1013.

Hall, F.G., Strebel, D.E., Nickeson, J.E., and Goetz, S.J. 1991, Radiometric rectification: toward a common radiometric response among multidate, multisensor images. **Remote Sensing of Environment**, 35(1):11-27.

Häme, T., and Heiler, I., 1998, An unsupervised change detection and recognition system for forestry. **International Journal of Remote Sensing**, 19(6):1079-1099.

Hill, J., and Sturm, B., 1991, Radiometric correction of multitemporal Thematic Mapper data for use in agricultural land-cover classification monitoring. **International Journal of Remote Sensing**, 12(7):1471-91.

Hill, R.A., 1999, Image segmentation for humid tropical forest classification in Landsat TM data. **International Journal of Remote Sensing**, 20(5):1039-1044.

Holben, B.N., and Justice, C.O., 1980, The topographic effect on spectral response from nadir-pointing sensors. **Photogrammetric Engineering and Remote Sensing**, 46:1191-1200.

Holben, B.N., and Shimabukuro, Y.E., 1993(a), Linear Mixing model applied to coarse spatial resolution data from multispectral satellite sensors. **International Journal of Remote Sensing**, 14(11):2231-2240.

Holben, B.N., and Shimabukuro, Y.E., 1993(b), Linear mixing model applied to AVHRR LAC data. In: **Proceedings of VII Simpósio Brasileiro de Sensoriamento Remoto**. Curitiba, Brazil.

Horler, D.N.H., and Ahern, F.J., 1986, Forestry information content of Thematic Mapper data. **International Journal of Remote Sensing**, 7(3):405-428.

Horwitz, H.M., Nalepka, R.F., Hyde, P.D., and Morgenstern, J.P., 1971, Estimating the proportions of objects within a single resolution element of a multispectral scanner. In: **Proceedings of the 7th International Symposim of Remote Sensing of the Environment** – ERIM, nº 10259-1-x pp.1307-1320.

Howard, J.A., 1991, ***Remote Sensing of Forest Resources – Theory and Application***, Ed.Chapman & Hall.

Hueck, K.,1953, Distribuição e habitat natural do Pinheiro-do Paraná (*Araucaria angustifolia*) - Contribuições para a fitossologia paulista. **Boletim da Universidade de São Paulo**, *156*:5-24, Faculdade de Filosofia, Ciências e Letras.

Huete, A., Justice, C., and Liu, H., 1994. Development of Vegetation and Soil Indices for MODIS-EOS. . **Remote Sensing of Environment**, 49:224-234.

Huguenin, R.L., Karaska, M.A., Van Blaricom, D., and Jensen, R., 1997, Subpixel classification of Bald Cypress and Tupelo Gum trees in Thematic Mapper Imagery. **Photogrammetric Engineering and Remote Sensing**, 63(6):717-725.

Huseby, R.B., Halck, O.M., and Solberg, R., 1999, A model-based approach for geometrical correction of optical satellite images. In **Proceedings of IGARSS' 99 – IEEE 1999 International Geoscience and Remote Sensing Symposium**, Hamburg, Germany, June 28 July 02 1999.

IBAMA, Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis, 1989, **Inventário Florestal da Flona de Irati: Florestas Plantadas**. Executed by FUPEF (Fundação de Pesquisas Florestais do Paraná).

IBAMA, Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis, 1990, **Inventário Florestal da Flona de Irati: Florestas Nativas (Volume 1)**. Executed by FUPEF (Fundação de Pesquisas Florestais do Paraná).

Igbokwe, J.I., 1999, Geometrical processing of multi-temporal satellite images for change detection studies. **International Journal of Remote Sensing**, 20(6):1141-1148.

Ingebritsen, S.E., and Lyon, R.J.P., 1985, Principal component analysis of multitemporal image pairs. **International Journal of Remote Sensing**, 6(5):687-696.

Itte, K.I., and Meyer, P., 1993, Geometric and radiometric correction of TM data of mountainous forested areas. **IEEE Transactions on Geoscience and Remote Sensing**, 31:764-770.

Jacobs, G.A., 1997, Unidades de Conservação no Estado do Paraná. Reflexões sob um contexto histórico-ambiental. In: **Proceedings of Congresso Brasileiro de unidades de Conservação**. Curitiba, Vol.II, Trabalhos Técnicos, 68-80.

Jacobs, G.A., 1999, Evolução dos remanescentes florestais e áreas protegidas no Estado do Paraná. In: **Proceedings of Forest 99 – Fifth International Congress and Exposition on Forests**. Bio 5023. Curitiba, Paraná Brazil, 15 to 18 June 1999. CD-Rom.

Jakubauskas, M.E., and Price, K., 1997, Empirical relationships between structural and spectral factors of Yellowstone lodgepole pine forests. **Photogrammetric Engineering and Remote Sensing**, 63(12):1375-1381.

Jensen, J.R., Cowen,D.J., Althausen,J.D., and Narumalani, S., 1993, An evaluation of the Coastwatch Change Detection Protocol in South Carolina. **Photogrammetric Engineering and Remote Sensing**, 59(6):1039-1046.

Kaufman, Y.J., and Sendra, C., 1988, Algorithm for automatic atmospheric corrections to visible and near-IR satellite imagery. **International Journal of Remote Sensing**, 9:1357-1381.

Kauth, R.J., and Thomas, G.S., 1976, The tasseled cap – a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. In **Proceedings of the Symposium on Machine Processing of Remotely Sensed Data,** Lars, Purdue University, West Lafayette, IN.

Kazmierczak, M.L., 1992, **Estimativa de variáveis dendrométricas em florestas homogêneas a partir de dados Landsat TM**. INPE, S.José dos Campos - INPE-5455-DI/497. M.Sc. Dissertation.

Kitchens, L.J., 1998, ***Exploring statistics – A modern introduction to data analysis and inference***. Brooks/Cole Publishing Company (Thomson Publishing Inc.) Pacific Grove, CA, USA.

Klein, R.M. 1963. Observacões e considerações sobre a vegetação do planalto nordeste catarinense. **Sellowia**, 15:39-54.

Krajicek, J. E., Brinkman, K. A. and Gringrick, S.F., 1961, Crown competition: a measure of density. **Forest Science**, **7**(1):35-42.

Kressler, F.P., and Steinocher, K.T., 1999, Detecting land cover changes from NOAA-AVHRR data using spectral mixture analysis. **International Journal of Applied Earth Observation and Geoinformation**, 1(1) 21-26.

Labovitz, M.L., and Marvin, J.W., 1986, Precision in geodetic correction of TM data as a function of the number, spatial distribution and success in matching of control points: A simulation. **Remote Sensing of Environment**, 20:237-52.

Lange Jr., F.L., and Maximiano, G.A., 1994, Utilization of Spot image for the purposes of forest stratification in Mangueirinha Indian Reserve, Paraná. In: **Proceedings of ISPRS Symposium: Resource and Environmental Monitoring**. Rio de Janeiro, Brazil, September 26-30 1994.

Lillesand, T.M., and Kiefer, R.W., 1994, ***Remote Sensing and Image Interpretation***, Wiley & Sons, N.York, 3rd ed.

Liu, H.Q., and Huete, A., 1995, A feedback based modification of the NDVI to minimise canopy background and atmospheric noise. **IEEE Transactions on Geoscience And Remote Sensing**, 33(2):457-465.

Longhi, S.J., 1980, **A estrutura de uma floresta natural de *Araucaria angustifolia* (Bert.) O.Ktze no sul do Brasil**. Curitiba, Paraná Brazil, Setor de Ciências Agrárias, Universidade Federal do Paraná. M.Sc. Dissertation.

Loughlin, W.P., 1991, Principal component analysis for alteration mapping. **Photogrammetric Engineering and Remote Sensing**, 57(9):1163-1169.

Lunetta, R., 1999, Applications, project formulation, and analytical approach. In: ***Remote Sensing change detection – Environment monitoring methods and applications***, Chapter 1. Taylor & Francis, London.

Ma, Z., and Redmond, R.L., 1995, Tau coefficients for accuracy assessment of classification of remote sensing data, **Photogrammetric Engineering and Remote Sensing**, 61(4):435-439.

Maack, R. 1950, Notas complementares à apresentação preliminar do mapa fitogeográfico do Estado do Paraná (Brasil). **Arquivos do Museu Paranaense**, 7: 351-361.

Maack, R. 1968, ***Geografia Física do Estado do Paraná*** Curitiba, Codepar.350 p. 1st. ed.

Machado, A.A., and Siqueira, J.D.P., 1979, Distribuição natural da *Araucaria angustifolia* (Bert.) O. Ktze. In: **Proceedings of IUFRO meeting: a problemática das araucárias**, Curitiba, Paraná, Brazil.

Magnussen, S., and Boudewyn, P., 1999, In: **Proceedings of International Forum – Automated interpretation of high spatial resolution digital imagery for forestry**. Victoria, British Columbia, Canada, February 10-12 1998.

Markham, B.L., and Barker, J.L., 1985, Spectral characterisation of the Landsat Thematic Mapper sensors. **International Journal of Remote Sensing**, 6:697-716.

Markham, B.L., and Barker, J.L., 1986, Landsat MSS and TM post-calibration dynamic ranges, exoatmospheric reflectances and At-Satellite temperatures. **EOSAT Landsat Technical Notes**.

Markham, B.L., and Barker, J.L., 1987, Thematic Mapper bandpass solar exoatmospheric irradiances. **International Journal of Remote Sensing**, 8:517-523.

Mas, J.F., and Ramirez, I., 1996, Comparison of land use classifications obtained by visual interpretation and digital processing. **ITC Journal**, 3/4:278-283.

Mather, P.M., 1999, ***Computer processing of remotely-sensed images – an introduction***. John Wiley and Sons, Chichester, 2nd ed.

Michener, W.K., and Houhoulis, P.F., 1997, Detection of vegetation changes associated with extensive flooding in a forested ecosystem, **Photogrammetric Engineering and Remote Sensing**, 63(12):1363-1374.

Miller, K.R., 1996, Conserving biodiversity in managed landscapes. In: **Biodiversity in managed landscapes**, edited by R.C. Szaro and D.W. Johnston. Oxford University Press, N.York, Usa.

Milne, A.K., 1988, Change detection analysis using Landsat imagery: A review of methodology. In: **Proceedings of IGARSS'88 Symposim**, Edinburgh, Scotland.

Minor, C.O., 1951, Stem-crown diameter relations in southern pine. **Journal of Forestry**, 49(7):490-93.

Mispan, M.R., 1997, **Multi-sensor remote sensing data for change detection analysis: a case study from Peninsular Malaysia**. Faculty of Science, University of Nottingham, UK. D.Phil. Thesis.

Moran, M.S., Jackson, R.D., Clarke, T.R., Qi, J., Cabot, F., Thome, K.J., and Markham, B.L, 1995, Reflectance factor retrieval from Landsat TM and Spot HRV data for bright and dark targets. **Remote Sensing of Environment**, 52:218-230.

Morisette, J.T., Khorram, S., and Mace, T., 1999, Land-cover change detection enhanced with generalised linear models. **International Journal of Remote Sensing**, 20(14):2703-2721.

Mouat, D.A., Mahin, G,G., and Lancaster, J., 1993, Remote sensing techniques in the analysis of change detection. **Geocarto International**, (2):39-50.

Murphy. R.J., and Wadge, G., 1994, The effects of vegetation on the ability to map soils using imaging spectrometer data. **International Journal of Remote Sensing**,15:63-86.

Muchoney, D.M., and Haack, B., 1994, Change detection for monitoring forest defoliation. **Photogrammetric Engineering and Remote Sensing**, 60(10):1243-1251.

Myneni R.B., and Asrar, G., 1994(a), Atmospheric effects and spectral vegetation indices. **Remote Sensing of Environment**, 47:390-402.

Myneni R.B., and Williams, D.L., 1994(b), On the relationship between FAPAR and NDVI. **Remote Sensing of Environment**, 49:200-211.

Nascimento, P.S.de R., Batista, G.T., and Almeida Filho, R., 1998. Efeito de pré-processamento (ajuste) no desempenho da segmentação e classificação de imagens Landsat-TM. In: **Proceedings of IX Simpósio Brasileiro de Sensoriamento Remoto**. Santos, Brazil, September 11-18 1998 - CD-Rom.

NETGIS, 1998, **SPRING – Sistema de processamento de informações georeferenciadas**. 4 vol. NETGIS – Geoprocessamento e Informática Ltd. São José dos Campos, SP, Brazil.

Ni, W., 1999 Modeling the bidirectional reflectance for a semi-arid landscape. In: **Proceedings of IGARSS'99 - IEEE Transactions on Geoscience And Remote Sensing Symposium**, Hamburg, Germany, June 28 – July 02 1999.

Novo, E.M.L.M., 1989, ***Sensoriamento Remoto: princípios e aplicações***. Edgar Blücher Ed., São Paulo, Brazil.

Oliveira, Y.M.M.de, 1980, **Correlações entre parâmetros dendrométricos em *Araucaria angustifolia*, utilizando fotografias aéreas**. Curso de Pós-Graduação em Engenharia Florestal, Setor de Ciências Agrárias, Universidade Federal do Paraná. M.Sc. Dissertation.

Oliveira, Y.M.M.de, and Rotta, E. 1982. Levantamento da estrutura horizontal de uma mata de araucaria no primeiro planalto paranaense. **Boletim de Pesquisa Florestal**. Curitiba,(4). Unidade Regional de Pesquisa Florestal Centro-Sul - EMBRAPA.

Olsson, H., 1994, **Monitoring of local reflectance changes in boreal forests using satellite data**. Swedish University of Agricultural Sciences. Remote Sensing Laboratory. Report 7, Umea. Pos-grad. Dissertation.

Osako, L.S., and Amaral, G., 1998, Extração de informações espectrais do Landsat-5/TM na Província Uranífera de Lagoa Real (BA). In: **Proceedings of IX Simpósio Brasileiro de Sensoriamento Remoto**. Santos, Brazil, September 11-18 1998 - CD-Rom.

Ouaidrari, H., Bégué, A., and Imbernon, J., 1996, Extraction of the pure spectral response of the landscape components in NOAA-AVHRR mixed pixels-application to the HAPEX-Sahel degree square. **International Journal of Remote Sensing**, 17(12):2259-2280.

Pereira, J.L.G., Batista, G.T., Thalês, M.C. and Roberts, D.A., Proporção de components derivadas de imagens Landsat TM para caracterização da cobertura da terra na Amazônia. In: **Proceedings of IX Simpósio Brasileiro de Sensoriamento Remoto**. Santos, Brazil, September 11-18 1998 - CD-Rom.

Pereira, J.L.G., Batista, G.T., and Roberts, D.A., Classificação da cobertura da terra na Região Amazônica através de classificação de regiões em imagens de proporção de componentes. In: **Proceedings of IX Simpósio Brasileiro de Sensoriamento Remoto**. Santos, Brazil, September 11-18 1998 - CD-Rom.

Perry, C.R., and Lautenschlager, L.F., 1984, Functional equivalence of Spectral Vegetation Indices. **Remote Sensing of Environment**, 14:169-182.

Philip, M.S., 1994, ***Measuring trees and forests***. CAB International, Wallingford. 2nd ed.

Picchiotti, A., Casacchia, R., and Salvatori, R., 1997, Multitemporal principal component analysis of spectral and spatial features of the Venice Lagoon. **International Journal of Remote Sensing**, 18(1):183-196.

Pollock, R., 1999, Individual tree recognition based on a synthetic tree crown image model. In: **Proceedings of International Forum – Automated interpretation of high spatial resolution digital imagery for forestry**. Victoria, British Columbia, Canada, February 10-12 1998.

Ponzoni, F.J., and Shimabukuro, Y.E., 1998, Relação entre índices de vegetação (NDVI e SAVI) e parâmetros biométricos de plantios de *Eucalyptus urophylla* e *Eucalyptus camaldulensis* em Cachoeira do Manteiga (MG). **Revista Árvore**, 22(3):357-366.

Prefeitura Municipal de Irati, 1999, **Irati, Loved City**, General Data.
Internet address: http://www.irati.pr.gov.br/english/index_loc.htm

Ricotta, C., and Avena, G.C., The influence of principal component analysis on the spatial structure of a multispectral dataset. **International Journal of Remote Sensing**, 20(17):3367-3376.

Ringrose,S., Vanderpost,C., and Matheson, W., 1997, Use of image processing and GIS techniques to determine the extent and possible causes of land management/fenceline induced degradation problems in the Okavango area, northern Botswana. **International Journal of Remote Sensing**, 18(11):2337-2364.

Ripple, W.J., 1985, Asymptotic reflectance characteristics of grass vegetation. **Photogrammetric Engineering and Remote Sensing**, 51:1915-1921.

Ripple, W.J., Wang, S., Isaacson, D.L., and Paine, D.P., 1991, A preliminary comparison of Landsat Thematic Mapper and SPOT-1 HRV multispectral data for estimating coniferous forest volume. **International Journal of Remote Sensing**, 12(9):1971-1977.

Roberts, D.A., Batista, G.T., Pereira, J.L.G., Waller, E.K., and Nelson, B.W., 1999, Change identification using multitemporal spectral mixture analysis: applications in eastern Amazonia. In:

Remote Sensing change detection – Environment monitoring methods and applications. Taylor & Francis, London.

Robinove, C.J., 1982, Computation with physical values from Landsat digital data. **Photogrammetric Engineering and Remote Sensing**, 48(5):781-84.

Robinson, A.H., Morrison, J.L., Muehrcke, P.C., Kimerling. A.J., and Guptill, S.C., 1995, **Elements of Cartography**. John Wiley & Sons,Inc. N.York, 6th ed.

Roger, R.E., 1996, Principal components transform with simple, automatic noise adjustment. **International Journal of Remote Sensing**, (17), 2719-2727.

Rondeaux, G., Steven,M., and Baret, F., 1996, Optimization of soil-adjusted vegetation indices. **Remote Sensing of Environment**, 55:95-107.

Roriz, A. (editor), 1998, **Dicionário Essencial Europa-Oxford,** compiled by J.Whitlam and L.C.Rait. Oxford University Press (1996) and Ed.Europa, São Paulo, Brazil.

Rosa, R., 1992, ***Introdução ao Sensoriamento Remoto***, Editora da Universidade Federal de Uberlandia, Minas Gerais, Brazil, 2nd ed.

Royer , A., Vincent, P., and Bonn, F., 1985, Evaluation and correction of viewing angle effects on satellite measurements of bi-directional reflectance. **Photogrammetric Engineering and Remote Sensing**, 51:1899-1914.

Sabol Jr., D.E., Adams, J.B., and Smith, M.O., 1992, Quantitative subpixel spectral detection of targets in multispectral images. **Journal of Geophysical Research**, 97(E2):2659-2672.

Sant'anna, L.,1998, **Landless' Movement – Mining a Rich Vein**. In: Web page of newspaper "O Estado de S.Paulo", English section.
Internet address: http://www.estado.com.br/english/semterra/msteng1.html

Santos, J.R., Lacruz, M.S.P., Keil, M., and Kramer, J.M.C., 1999, A linear spectral mixture model to estimate forest and savanna biomass at transition areas in Amazonia. In **Proceedings of IGARSS' 99 – IEEE 1999 International Geoscience and Remote Sensing Symposium**, Hamburg, Germany, June 28 – July 02 1999.

Schowengerdt, R.A., 1997, ***Remote Sensing: Models and Methods for Image Processing***. S.Diego: Academic Press, 2nd ed.

Shimabukuro, YE., and Smith, J.A., 1991, The least-squares mixing models to generate fraction images derived from remote sensing multispectral data. **IEEE Transactions on Geoscience and Remote Sensing**, 29(1):16-20.

Shimabukuro, YE., and Smith, J.A., 1995, Fraction images derived from Landsat TM and MSS data for monitoring reforested areas. **Canadian Journal of Remote Sensing**, 21(1):67-74.

Shimabukuro, Y.E., Batista, G.T., Mello, E.M.K., Moreira, J.C., and Duarte, V., 1998, Using shade fraction image segmentation to evaluate deforestation in Landsat Thematic Mapper images of the Amazon Region. **International Journal of Remote Sensing**, 19(3):535-541.

Shimizu, J.Y., and Oliveira, Y.M.M.de, 1980, Distribution, variation and uses of *Araucaria angustifolia* gene resources in Southern Brazil. In: **Proceedings of IUFRO Symposium – Melhoramento genético florestal e produtividade de espécies florestais de rápido crescimento**. Águas de São Pedro. São Paulo, Brazil.

Silva, A.de B., 1998, **Curso de Extensão – Técnicas de geoprocessamento em geociências** - 2º Módulo – Sistemas de Informações Geográficas (SIG). Instituto de Geociências, UNICAMP, SP,Brazil.

Silva, J.A., 1978, Comprimento do tronco, relação altura-diâmetro e altura-idade do pinheiro brasileiro (*Araucaria angustifolia* Bert. 0. Ktze) e folhosas associadas em ambiente natural. **Brasil Florestal,** 9 (34):43-50.

Silva, J.A., Salomão, A.N., Gripp, A., and Leite, E.J., 1997, Phytosociological survey in Brazilian forest genetic reserve of Caçador. **Plant Ecology**, 133:1-11.

Singh, A., 1989, Change detection techniques using remotely-sensed data. **International Journal of Remote Sensing**, 6:989-1003.

Slater, P.N., Biggar, S.F., Holm, R.G., Jackson, R.D., Mao, Y., Moran, M.S., Palmer, M., and Yuan, B., 1987, Reflectance- and radiance-based methods for the in flight absolute calibration of multi-temporal sensors. **Remote Sensing of Environment**, 22:11-37.

Smith, M.O., Ustin, S.L., Adams, J.B., and Gillespie, A.R., 1990, Vegetation in desert: I. A regional measure of abundance from multispectral image. **Remote Sensing of Environment**, 31:1-26.

Smith, C., Pyden, N., and Cole, P., 1994, ***ERDAS IMAGINE*** Version 8.1 – Field Guide – ISG 660, 3rd ed.

Soares, V.P., Hoffer, R.M., Brites, R.S., and Ribeiro, C.A.A.S.,1996, Monitoramento de Plantações de *Eucalyptus* spp no Vale do Rio Doce através de imagens TM/Landsat 5. In: **Proceedings of VIII Simpósio Brasileiro de Sensoriamento Remoto**. Salvador, Brazil, April 14-19 1996 - CD-Rom.

Soares, V.P., and Hoffer, R.M., 1996, Cover types change classification using multitemporal Landsat TM data. **Revista Árvore**, Viçosa-MG, 20(4):545-554.

Sohl, T.L., 1999, Change analysis in the United Arab Emirates: an investigation of techniques. **Photogrammetric Engineering and Remote Sensing**, 65(4):475-484.

Sohn, Y., and McCoy, R.M., 1997, Mapping desert shrub rangeland using spectral unmixing and modelling spectral mixtures with TM data. **Photogrammetric Engineering and Remote Sensing**, 63(6): 707-716.

SOS Mata Atlântica, INPE and ISA, 1998, ***Atlas da Evolução dos remanescentes florestais e ecossistemas associados no domínio da Mata Atlântica***, São José dos Campos, Brasil.

Sousa, C.J.S., 1998, Discriminação das classes de uso/cobertura vegetal em uma área do Médio Vale do Paraíba-SP, a partir das images fração solo e vegetação. In: **Proceedings of IX Simpósio Brasileiro de Sensoriamento Remoto**. Santos, Brazil, September 11-18 1998 - CD-Rom.

Sousa, C.L., and Ponzoni, F.J.,1998, Avaliação de índices de vegetação e de bandas TM/Landsat para estimativa de volume de madeira em floresta implantada de *Pinus* spp. In: **Proceedings of IX Simpósio Brasileiro de Sensoriamento Remoto**. Santos, Brazil, September 11-18 1998 - CD-Rom.

Sousa, C.L., 1997, **Avaliação de imagens-índice e de imagens-proporção para detecção de alteração no volume de madeira em florestas implantadas**. Instituto Nacional de Pesquisas Espaciais, S.José dos Campos, S.Paulo, Brazil, MSc Dissertation.

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, understory vegetation and background reflectance. **International Journal of Remote Sensing**, 11(1):95-111.

Steinnocher, K., and Kressler, F., 1999, Application of spectral mixture analysis for monitoring urban development. In: **Proceedings of the RSS99 – Earth Observation – From Data to Information**. Cardiff University and University of Wales, Swansea, September 8-10 1999.

Sunar, F., 1998, An analysis of changes in a multi-data set: a case study in the Ikitelli area, Istanbul, Turkey. **International Journal of Remote Sensing**, 19(2):225-235.

Switzer, P., Kowalick, W.S., and Lyon, R.J.P., 1981, Estimation of Atmospheric Path-Radiance by the Covariance Matrix Method. **Photogrammetric Engineering and Remote Sensing**, 47:1469-1476.

Sziekielda, K.H., 1986, ***Satellite Remote Sensing for Resources Development***, Graham and Trotman, London.

Tanré, D., Deroo, C., Duhaut, P., Herman, M., and Morcrette, J.J., 1990, Description of a computer code to simulate the satellite signals in the solar spectrum: the 5S code. **International Journal of Remote Sensing**, 11:659-668.

Teillet, P.M., 1986, Image correction for radiometric effects in remote sensing. **International Journal of Remote Sensing**, 7:1637-1651.

Thiam, A.K., 1997, **Geographic Information Systems and Remote Sensing Methods for Assessing and Monitoring Land Degradation in the Sahel Region: The case of South Mauritania**, Department of Geography, Clark University, Worcester, Massachussets, USA. Ph.D. Thesis.

Thomas, I.L., Benning, V.M., and Ching, N.P., 1987, ***Classification of Remotely Sensed Images***. Adam Hilger Ed, Bristol, UK.

Tompkins, S., Mustard, J.F., Pieters, C.M., and Forsyth, D.W., 1997, Optimization of endmembers for spectral mixture analysis. **Remote Sensing of Environment**, 59:472-489.

Townshend, J.R.G., Justice, C., Gurney, C., and McManus, J., 1992, The impact of misregistration on change detection. **IEEE Transactions on Geoscience and Remote Sensing**, 30:1054-1060.

Trotter, C.M., Dymond, J.R., and Goulding, C.J., 1997, Estimation of volume timber in a coniferous plantation forest using Landsat TM. **International Journal of Remote Sensing**, 18(10):2209-2223.

Tucker, C.J., Fung, I.Y., Keeling, C.D., and Gammon, R.H., 1986, Relationship between atmospheric CO_2 variations and a satellite-derived vegetation index. **Nature**, 319.

USDA, Forest Service, 1995, ***Guidelines for the use of imagery for vegetation mapping*** - EM-7140-25. Engineering Staff. Edited by Paul A.Maus.

Veblen, T.T., Burns, B., Kitzberger, T., Lara,A., and Villalba, R., 1995, The ecology of the conifers of Southern South America. In: **Ecology of the Southern Conifers**. Melbourne University Press. Edited by Neal J. Enright and R.S.Hill.

Veronese, V.F., Lins, H.B., and Ferreira, M.C., 1998, Acurácia da correção geométrica de imagens de satélite em função do método de aquisição dos pontos de controle. In: **Proceedings of GIS BRASIL 98**. Curitiba, PR, Brazil.

Viana, V.M., Tabanez, A.J.A. & Juan, L.A. 1992, Restauração e manejo de fragmentos florestais. IN: **Proceedings of Congresso Nacional sobre Essencias Nativas, II**. São Paulo., São Paulo, Instituto Florestal de São Paulo, pp.400-406.

Vogelman, J.E., 1988. Detection of forest change in the Green Mountains of Vermont using mutispectral scanner data. **International Journal of Remote Sensing**, 9:1187-1200.

Vogelman, J.E., and Rock, B.N., 1988, Assessing forest damage in high-elevation coniferous forests in Vermont and NewHampshire using Thematic Mapper data. **Remote Sensing of Environment**, 24:227-246.

Volkart, C.M., 1969, Determinación de la relación diametro copa/diametro do tronco en *Araucaria angustifolia y Pinus elliottii* en la Provincia de Misiones. In: **Proceedings of Congresso Forestal Argentino (Actas)**. 1., Buenos Aires, 1970. p. 231-7.

Watts, S. and Halliwell, L. (eds.), 1996, **Essential Environment Science - Methods & Techniques**, Routledge, London.

Whitmore, T.C., and Sayer, J.A., (eds.),1994, Deforestation and species extinction in tropical moist forests. In: **Tropical Deforestation and Species Extinction**. The IUCN Forest Conservation Programme. Chapman & Hall, London, UK. 2nd ed.

Williams, D.L., 1991, A comparison of spectral reflectance properties at the needle, branch and canopy level for selected conifer species. **Remote Sensing of Environment**, 35:79-93.

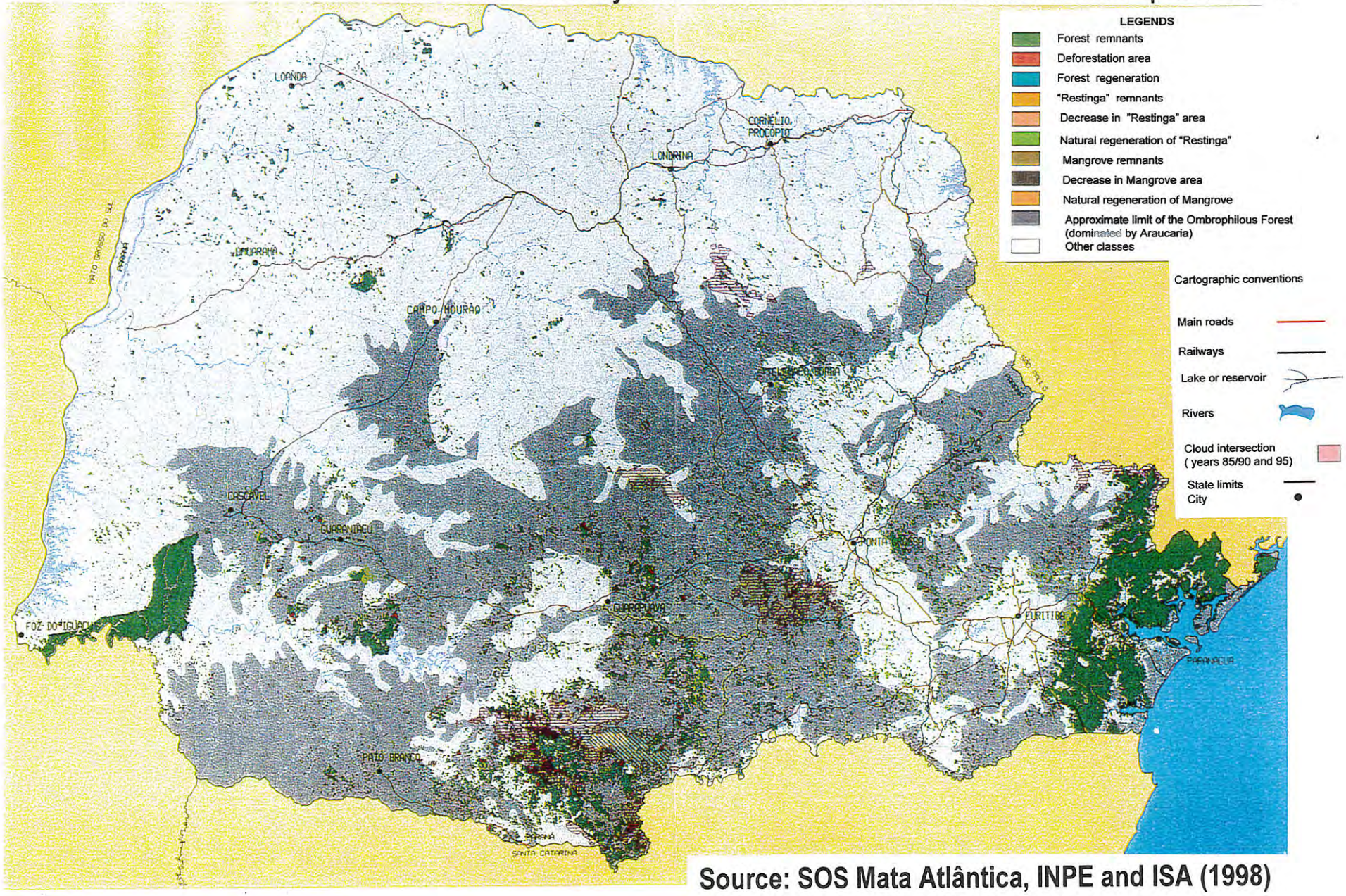
Wulder, M., 1999, Image spectral and spatial information in the assessment of forest structural and biophysical data. In: **Proceedings of International Forum – Automated interpretation of high spatial resolution digital imagery for forestry**. Victoria, British Columbia, Canada, February 10-12 1998

Yuan, D., Elvidge,C.D., and Lunetta, R.S., 1999, Survey of multispectral methods for land cover change analysis. In: ***Remote Sensing change detection – Environment monitoring methods and applications***, Chapter 2. Taylor & Francis, London.

Zachow, R., 1991, **Modelo de Uso Múltiplo dos Recursos Naturais da Floresta Nacional do Jamari**., Curso de Pós-Graduação em Engenharia Florestal, Setor de Ciências Agrárias da Universidade Federal do Paraná. M.Sc. Dissertation.

Appendices

APPENDIX 2.1 - PARANÁ – Forest Remnants and Ecosystems associated with the Mata Atlântica domain – period 1990-95



Source: SOS Mata Atlântica, INPE and ISA (1998)

APPENDIX 2.2

Global Positioning System

The NAVSTAR (**Nav**igation **S**ystem with **T**ime **a**nd **R**anging – Global Positioning System) is a full-time, all-weather world wide radio-navigation aid. According to Frohn (1998) GPS allow the technician to navigate to and specifically georeference a particular object or point on the landscape, as well as fix locations of objects found in the field.
The system is operated by the United States Department of Defense and it is composed basically by three segments:

a) The Space Segment: it consists of 24 satellites in earth orbit in six orbital planes inclined 55 degrees to the equator. The orbital period is approximately 12 hours and according to COPEL (1995), the satellites orbit at an altitude of approximately 20,200-km. Each satellite transmits a signal encoded on two frequencies, L1 (1575.42 MHz) and L2 (1227.60 MHz). These satellites transmit several different messages encoded on the carrier. There are two kinds of code, the coarse acquisition (C/A) code, and the precision (P) or encrypted precision (Y) code. The C/A code (1.023 MHz) is the code that civilian receivers use to determine position. The P code is a higher frequency code at 10.23 MHz, which allows better accuracy in measuring the transit time of the signal, and therefore better range measurements. The Y code is the encrypted version of the P code, which is in use now that the system is operational. To use Y code users must have an authorized receiver with the appropriate cryptographic key.

b) Control Segments: the master control station at Falcon Air Force (AFB) heads it in Colorado, USA. At this and other monitoring stations at Diego Garcia, Ascension Island, Kwajalein, Hawaii, and Onizuka AFB, accurate measurements of the satellite orbits are made.

c) User Segment: this segment consists of all the GPS receivers on the planet. They operate on the same basic principle of pseudorange measurements, which consists of the time shift need to correlate the actual signal transmitted by a satellite and the replica generated by the receiver multiplied by the speed of light (propagation speed of the radio signal). The distance measurement is called pseudorange because of the errors introduced into the system by relativistic effects, ionosphere delays, and selective availability introduced to degrade system performance by the United States government. When one satellite pseudorange is determined, the distance describes a sphere on which the receiver could lie. When two satellites are included in the solution, the intersection of the spheres is the solution, which is a circle. When a third satellite is included, it resolves to two points. The fourth satellite not only resolves the

ambiguity but also determines the system time, which is needed for accurate measurement of the signal travel time. The receiver will then compute a transformation of the position that is in an earth centered earth fixed frame of reference to a useful system such as latitude and longitude or UTM grids, for example.

In the case of this study, it was used the COPEL (the State Energy Company) base station. It is located at their office in Irati in the co-ordinates: latitude 25º 28' 27.131392" S and longitude 50º 40' 07.599639" W. The altitude is 835.10 m above the sea level.

The methodology to capture the signal is based on the calculation of the distance to each satellite by the receiver. Since the base station (BASE) knows both the position of the satellite and the receiver (FIELD), it can calculate the distance from the satellite precisely. The difference between the actual and measured signal is the differential correction for the pseudorange to the satellite. The information then saved for post-processing, made every evening because of memory management in the receivers. All the 93 plots installed in the Forest Inventory of the natural patches of the *Araucaria* forest were recovered by locating them in the field and taking their position. The same procedure was used to the 49 plots of pine species. The use GPS under canopy is not an easy task, for many reasons, as: a) sometimes is only possible to capture less than four satellites; b) sometimes the crown closure is close to 100% (which difficult the signal receiving). Therefore, the planning of field work must take into account: a) eventual necessity of climbing trees to improve the changes of receiving the signal; and b) concentrate work in the hours during the day when more satellites are aligned in the study region to increase the probability of success in the determination of each point co-ordinates.

APPENDIX 2.3

List of native forest species inventoried at the Irati National Forest - INF

Common name	Scientific name	Family
Açoita-cavalo	Luehea divaricata	Tiliaceae
Angico	Parapiptadenia rigida	Leguminosae
Araçá	Psidium catleianum	Myrtaceae
Araçatinga	Non-identified	Myrtaceae
Ariticum	Rollinia rugulosa	Annonaceae
Aroeira	Schinus terebinthifolius	Anacardiaceae
Bracatinga	Mimosa scabrella	Leguminoseae
Branquinho	Sebastiana klotzschiana	Euphorbiaceae
Bugreiro	Lithreaea brasiliensis	Anacardiaceae
Cambará	Gochnatia polymorpha	Compositae
Cambuí	Siphoneugenia sp	Myrtaceae
Canafístula	Cassia Leptophylla	Leguminoseae
Canela-lajeana	Ocotea pulchella	Lauraceae
Canela-amarela	Nectandra grandiflora	Lauraceae
Canela-branca	Nectandra lanceolata	Lauraceae
Canela-fogo	Cryptocaria aschersoniana	Lauraceae
Canela-guaicá	Ocotea puberula	Lauraceae
Canela-imbuia	Nectandra megapotamica	Lauraceae
Canela-pimenta	Ocotea sp	Lauraceae
Canela-porco	Ocotea sp	Lauraceae
Canela-sassafrás	Ocotea pretiosa	Lauraceae
Canelão	Ocotea sp	Lauraceae
Capororoca	Rapanea umbellata	Myrsinaceae
Capororoca-branca	Rapanea sp	Myrsinaceae
Carne-de-vaca	Styrax leprosus	Styracaceae
Caroba	Jacaranda puberula	Bignoniaceae
Cataia	Drimys brasiliensis	Winteraceae
Caujuja	Clethra scabra	Clethraceae
Caúna	Ilex theezans	Aquifoliaceae
Cedro	Cedrela fissilis	Meliaceae
Cedro-branco	Não identificada	Meliaceae
Cerejeira	Eugenia involucrata	Myrtaceae
Congonha	Ilex dumosa	Aquifoliaceae
Corticeira	Erytrina falcata	Leguminosea
Cuvatã	Cupania vernalis	Sapindaceae
Erva-mate	Ilex paraguariensis	Aquifoliaceae
Farinha-seca	Albizia polycephala	Leguminoseae
Figueira	Ficus sp	Moraceae
Guabirobeira	Campomanesia xanthocarpa	Myrtaceae
Guaçatunga	Casearia decandra	Flacourtiaceae
Guaçatunga-preta	Casearia obliqua	Flacourtiaceae

Continuation of the	Appendix 2.3	
Common name	**Scientific name**	**Family**
Guamirim	Myrcia sp	Myrtaceae
Guamirim-ferro	Calyptranthes sp	Myrtaceae
Guamirim-vermelho	Myrceugenia sp	Myrtaceae
Imbuia	Ocotea porosa	Lauraceae
Jaboticabeira	Myrciaria trunciflora	Myrtaceae
Jacaranda	Dalbergia brasiliensis	Leguminoseae
Juvevê	Zanthoxyllum kleinii	Rutaceae
Leiteiro	Sapium grandulatum	Euphorbiaceae
Limoeiro-do-mato	Randia armata	Rubiaceae
Louro-pardo	Cordia tricotoma	Boraginaceae
Mamica-de-cadela	Zanthoxyllum rhoifolia	Rutaceae
Maria-mole	Pisonia sp	Nyctaginaceae
Maria-preta	Diatenopterix sorbifolia	Sapindaceae
Miguel-pintado	Matayba elaegnoides	Sapindaceae
Monjoleiro	Acacia poliphylla	Leguminoseae
Murta	Blepharocalyx sp	Myrtaceae
Palmeira-jerivá	Cocus sp	Arecaceae
Pau-andrade	Persea major	Lauraceae
Pau-d'alho	Cinnamomum sp	Lauraceae
Pau-ferro	Non-identified	Non-identified
Pau-marfim	Balforodendron riedelianum	Rutaceae
Peloteira	Non-identified	Solanaceae
Pessegueiro-bravo	Prunus brasiliensis	Rosaceae
Pimenteira	Capsicodendron dinisii	Canelaceae
Pinheiro-do-paraná	Araucaria angustifolia	Araucariaceae
Saboneteira	Quillaja brasiliensis	Rosaceae
Santa-rita	Laplacea fruticosa	Theaceae
Sapopema	Sloaneae lasiocoma	Elaeocarpaceae
Sapuva	Machaerium stipitatum	Leguminosae
Tarumã	Vitex megapotamica	Verbenaceae
Tenente-josé	Aeschrion sp	Simaroubaceae
Timbó	Lonchocarpus sp	Leguminoseae
Tupixaba	Baccharis sp	Compositae
Uvarana	Cordyline dracaenoides	Liliaceae
Vassourão	Piptocarpha sp	Compositae
Vassourão-branco	Piptocarpha angustifolia	Compositae
Vassourão-preto	Vernonia discolor	Compositae
Vassourinha	Baccharis ligustrina	Compositae
Voadeira	Ilex microdonta	Aquifoliaceae

APPENDIX 3.1

Software SPRING

SPRING (System for Processing of Georeferenced Information) is a second generation geographical database. Available for UNIX and Windows environments, it developed by INPE (National Institute of Space Researches) with the following characteristics: a) it operates as a geographical database and supports great volume of data; b) it works with vector and raster data, integrating image processing data with SIG; c) it has a friendly and powerful environment through the combination of menus and windows with a spatial language user programmable (LEGAL - Space-geographical Language based on Algebra); and d) it is versatile being capable of operating with all its functionality in environments that vary from personal computer to RISC stations (NETGIS 1998).

Objectives of the software:

SPRING is an innovative system, projected initially for workstations nets based on RISC architecture and UNIX environment. Developed using advanced programming techniques, SPRING combines highly interactive interface, using "X Window System" (NETGIS 1998).

Pressure of the Brazilian RS and GIS community, and the perception that this type of tool had actually a huge demand motivated INPE team for the development of SPRING based on two premises: *integration of data and use easiness.* In the first case, it was verified that the complexity of Brazilian environmental problems requests a strong capacity of data integration among satellite images, thematic maps, and terrain numeric models. Additionally, many of the available systems nationally are complex on their use and time demanding in terms of training. SPRING is a highly attractive option in the geoprocessing area, because it became "software of public domain" and it can be acquired over the Internet.

The Internet INPE address is: http: // www.dpi.inpe.br/spring.html. SPRING can be acquired free of charge in its Portuguese, Spanish or English version.

The minimum software and hardware demands for Platform PC to execute the product correctly are:
Microcomputer IBM/PC Pentium 100 Mhz or superior
Memory RAM of 16 Mbytes (recommended 32)
Hard disk of 1 Gbytes
Monitor of coloured video SVGA, 14 " NI, dp 0.28 mm

Drive 31/2 ", 1.44 Mbytes.
Unit of CD-ROM.

Programming in LEGAL - Algebra of Maps

LEGAL is the short name for "Space Language for Algebraic Geoprocessing". The algebraic operations in LEGAL involve given georeferencied data represented under raster or vector form. It allows procedures modelling under descriptive form, where data and statements are declared and organised according to grammatical rules. Modelling occurs under the form of a program (NETGIS 1998).

LEGAL programs were written by the author of the research with supervision of INPE team, especially Flavio Ponzoni and Claudio Barbosa. Programs were prepared for: a) operations as data preparation: image (geometric and radiometric) restoration, geometric co-registration and image transformations. Programs in LEGAL were also written for the image processing display presentation, performed in **Chapters 6** and **7**.

APPENDIX 4.1

Descriptive statistics of Experiment 1 – spectral responses and structural variables relationships of Pine species

Table 4.1.1 Descriptive statistics of field data collected from stands within forest management units

Species: *P.elliottii* : Number of plots = 44				
Variable	Mean	St.Dev.	Maximum	Minimum
Age (years)	23.4	2.0	28.0	18.0
Density (trees ha^{-1})-TREE	30.4	10.7	51.0	7.0
DBH (cm)	31.3	4.5	44.2	24.6
Height (m)	26.5	1.8	30.9	23.2
Basal Area (m^2 ha^{-1})	27.4	11.4	59.0	16.0
Volume (m^3 ha^{-1})	471.6	127.4	829.7	211.5
Species: *P.taeda* : Number of plots = 06				
Age (years)	20.0	0	20.0	20.0
Density (trees ha^{-1})-TREE	32.7	20.9	71.0	18.0
DBH (cm)	32.0	7.0	41.0	25.5
Height (m)	24.0	3.4	29.5	19.7
Basal Area (m^2 ha^{-1})	38.7	7.9	49.0	25.0
Volume (m^3 ha^{-1})	421.4	64.2	514.6	352.8
Species: *P.elliottii* and *P.taeda* : Number of plots = 50				
Age (years)	22.9	2.2	28.0	18.0
Density (trees ha^{-1})- TREE	30.7	12.1	71.0	7.0
DBH (cm) - DBH	31.4	4.8	44.2	22.1
Height (m) - HEIGHT	26.2	2.2	30.9	22.1
Basal Area (m^2 ha^{-1}) - BA	37.6	9.3	59.0	16.0
Volume (m^3 ha^{-1}) - VOL	465.6	122.2	829.7	211.5

Table 4.1.2 Analysis of Variance for Volume - *P.elliottii* (1) and *P.taeda* (2)

```
Analysis of Variance for Volume
Source     DF        SS        MS        F        P
Spp         1     13323     13323     0.89    0.350
Error      48    718101     14960
Total      49    731425

                                  Individual 95% CIs For Mean
                                  Based on Pooled StDev
Level       N      Mean     StDev  -------+---------+---------+---------
1          44     471.6     127.4                      (------*-----)
2           6     421.4      64.2   (----------------*----------------)
                                  -------+---------+---------+---------
Pooled StDev =    122.3                360       420       480
```

Analysis of variance for Volume data (**Table 4.1.2**) for *P.elliottii* (Species 1) and *P.taeda* (Species 2) showed a p-value > than .10, failing to reject the null hypothesis. Therefore, the differences in volume can be considered insignificant. The confidence intervals of Volume also do overlap with species, confirming that they are not significantly different from each other. Considering that *P.taeda* data were included in *P.elliottii* limits, the data were therefore analysed as one unique set.

Table 4.1.3 Correlation between forest plot parameters

Correlations (Pearson) - P.elliottii					
	AGE	TREE	DBH	HEIGHT	BA
TREE	-0.437*				
DBH	0.525*	-0.681*			
HEIGHT	0.520*	-0.235	0.629*		
Basal Area	-0.256	0.856*	-0.320*	-0.057	
Volume	0.007	0.731*	-0.222	0.252	0.910*
Correlations (Pearson) – P.elliottii and P.taeda					
	AGE	TREE	DBH	HEIGHT	BA
TREE	-0.345*				
DBH	0.379*	-0.721*			
HEIGHT	0.544*	-0.350*	0.509*		
Basal Area	-0.235	0.792*	-0.307*	-0.083	
Volume	0.061	0.635*	-0.206	0.219	0.885*

(*) Marked correlations are significant at $p < 0.05$

As it can be observed in the **Table 4.1.3**, tree density (TREE) was negatively correlated (at the 95 per cent confidence level) with AGE (-0.437 and -0.345), certainly due to the management practice of systematic or selective thinning, with the consequent periodical removal of trees although this practice was not conducted as planned during long periods at the INF. TREE (number of trees) was also negatively correlated to DBH (-0.681 and -0.721) and HEIGHT (-0.235 and -0.350) as a natural consequence of initial density and thinning. DBH (0.525 and 0.379) and HEIGHT (0.520 and 0.544) were positively correlated with AGE, as expected, as they are related to tree growth. Both Basal Area and Volume were obtained from individual trees rather than from the mean values. Volume was very weakly correlated to any measured variable, as DBH (-0.222/-0.206) and HEIGHT (0.252 and 0.219).

Table 4.1.4 ANOVA applied to the eight groups of data sets

One-way Analysis of Variance

Analysis of Variance

Source	DF	SS	MS	F	P
Factor	7	11.25	1.61	0.25	0.972
Error	392	2526.55	6.45		
Total	399	2537.80			

Level	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev
a) v5	50	6.980	2.638	(----------*----------)
b) mean	50	7.491	2.434	(----------*----------)
c) v1245	50	7.317	2.573	(----------*----------)
d) v2356	50	7.547	2.454	(----------*----------)
e) v4578	50	7.343	2.702	(----------*----------)
f) v5689	50	7.497	2.557	(----------*----------)
g) v24568	50	7.449	2.421	(----------*----------)
h) v13579	50	7.416	2.518	(----------*----------)
				6.60 7.20 7.80 8.40

Pooled StDev = 2.539

The F-test p-value of 0.972 indicates that there is no quite significant evidence at α=0.05 of differences among the means. These results indicate that, in the case of present study the use of v5 (group "a" - the central point of the window and the pixel located by GPS) can be used as a reasonable approximation of the exact location of the plot.

Table 4.1.5 Correlations Coefficients of VOL (volume) and BA (Basal Area) and respective logarithms (L) against spectral reflectance and the logarithm of the spectral reflectance (Pearson) - *P.elliottii* and *P.taeda*

	TM1	TM2	TM3	TM4	TM5
LBAB	-0.045	-0.288*	-0.181	-0.138	-0.618*
LOGV	-0.074	-0.104	-0.183	-0.149	-0.749*
LTM1	1.000*	0.140	0.260	0.011	0.186
LTM2	0.126	0.997*	0.546*	0.536*	0.340*
LTM3	0.272*	0.575*	0.991*	0.396*	0.388*
LTM4	0.013	0.522*	0.365*	0.996*	0.364*
LTM5	0.158	0.374*	0.314*	0.406*	0.975*

(*) Marked correlations are significant at p < 0.05

where: LBA = Logarithm of Basal Area; LOGV = Logarithm of Volume; LTM1 to LTM5 = Logarithm of TM1 to TM5 respectively.

Table 4.1.6 Correlations (Pearson)- *P.elliottii* and *P.taeda*

Species: P.elliottii						
	Age	TM1	TM2	TM3	TM4	TM5
TM1	-0.009					
TM2	0.087	0.181				
TM3	0.051	0.238	0.633*			
TM4	-0.348*	0.003	0.526*	0.411*		
TM5	0.113	0.185	0.439*	0.305*	0.374*	
TREE	-0.437*	-0.103	-0.108	-0.139	0.193	-0.483*
DBH	0.525*	-0.051	0.071	0.083	-0.372*	-0.217
Height	0.520*	-0.165	-0.032	-0.146	-0.474*	-0.143
BA	-0.256	-0.124	-0.210	-0.165	-0.003	-0.642*
VOL	-0.007	-0.072	-0.178	-0.137	-0.135	-0.706*
Species: P.elliottii and P.taeda						
TM1	0.035					
TM2	0.161	0.219				
TM3	0.103	0.264	0.636*			
TM4	-0.325*	0.032	0.509*	0.367*		
TM5	0.021	0.180	0.398*	0.249	0.393*	
TREE	-0.345*	-0.159	-0.139	-0.103	0.023	-0.487*
DBH	0.379*	0.129	0.095	0.082	-0.214	0.240
Height	0.544*	-0.078	0.070	-0.078	-0.219	-0.108
BA	-0.235	-0.120	-0.213	-0.140	-0.065	-0.639*
VOL	0.061	-0.053	-0.147	-0.099	-0.154	-0.716*

(*) Marked correlations are significant at p < 0.05

Transforms

Transformations are pixel to pixel operations using different bands through a defined rule, generating a new band that combines the information of the original bands. Some of transforms are mentioned and discussed more deeply in **Chapters 5,6** and **7**. For the purpose of this experiment, transforms as Vegetation Indices, ratios, fraction images or synthetic images were created to compare their performance and to verify their capability on detecting variation on ground level. Therefore, besides the single bands 1 to 5 and 7, already mentioned, the following transformations were performed:

a) Vegetation Indices:

```
NDVI =  (TM4 - TM3) / (TM4 + TM3))
SAVI =  (((TM4 - TM3) / (TM4 + TM3 + 0.5))* 1.5)
```

Other indices based on the relation between the near-infrared band (band 4) and red (band 3) were tentatively performed, as Thiam's Transformed Vegetation Index-TTVI (proposed by Thiam 1997) and RATIO (TM3/TM4), or even Greenness Condition Index-GI (cited by Jakubauskas *et al.* 1997). However, for the conditions of this study, the results were not better than those obtained by each of these bands themselves and the relationship between volume (or basal area) and band 3 or band 4 was weak. The most traditional Vegetation Indexes (NDVI, SAVI, TTVI and RATIO) presented very low correlation with the ground variables, probably due to the weak relation between bands TM3 and TM4 and their low correlation with ground variables. Modified Indices, as Green NDVI and Moisture vegetation Index - MVI (Sousa *et al.* 1998) were also calculated, but they both were based on band 4 and, for the same reason, were not well succeed. An index with the same formula of the NDVI but using bands 2 and 5 (NDVI52) was also created based on the correlation of these two bands and ground variables as Basal area and Volume.

Mid-IR/red reflectance index `(MIRI = (TM7 - TM3) / (TM7 + TM3)` was tested as mentioned by Jakubauskas *et al.* (1997), but the presence of band 7, for the condition of this work did not improve any index in the relationship with volume. Better results were obtained using band 5 instead band 7 and the correlation results can be seen in **Appendix 4.1-Table 4.1.7** for
`MIRIm = (TM5 - TM3) / (TM5 + TM3).`

b1) Scene-specific Tasseled Cap Transformation, as suggested by Jakubauskas *et al.* (1997):

```
TC1B = 0.4395*TM1 + 0.5416*TM2 + 0.5406*TM3 + 0.4647*TM4 + 0.0704*TM5 + 0.0169*TM7
TC2G = -0.2254*TM1 - 0.2054*TM2 - 0.3599*TM3 + 0.8795*TM4 - 0.0437*TM5 - 0.0461*TM7
TC3W = 0.4154*TM1 + 0.1777*TM2 - 0.3618*TM3 - 0.0537*TM4 - 0.7526*TM5 - 0.3095*TM7
```

b2) Greenness(G) and brightness (B) images (derived by , generated by Crist and Cicone (1984):

```
G=((TM1*-0.2848) + (TM2*-0.2435) + (TM3*-0.5336) + (TM4*0.7243) + (TM5*0.0840) + (TM7*-0.1800) + 100.5);
B=((TM1*0.3037) + (TM2*0.2793) + (TM3*0.4743) + (TM4*0.5585) + (TM5*0.5082) + (TM7*0.1863) + 0.5);
```

c1) Synthetic bands: fraction images (Lsoil, Lshade and Lveg) derived from Linear Mixing Model (details about the technique can be found in **Chapter 6).** The generated fraction images for shade, soil and vegetation (in this case the pure pixel were captured in area of pine plantation) were included in the data set. They were created by extracting the spectral signatures of each component using the whole set of available bands (bands 1 to 5 and 7).

c2) Synthetic bands: PCA - Principal Components Analysis (details about the technique can be found in **Chapter 5**). The image generated by the first principal component of two different band combinations: a) TM5 bands 1 to 5 and 7 (PCA1-7) and b) TM5 bands 2 and 5 (PCA2-5) were included in the analysis.

A correlation matrix (**Table 4.1.7**) was generated based on the list above and the following synthetic images were significantly (at $p < 0.05$) correlated with volume and basal area:

Table 4.1.7 - Synthetic images correlated with volume and basal area

Mnemonic	Resulted R (for volume)	Resulted R (for basal area)
MIRIm	-0.56	-0.50
NDVI52	-0.62	-0.54
Lshade	0.33	0.38
Lsoil	-0.36	-0.41
TC3W	0.45	0.40
PCA2-5	-0.59	-0.53

Correlation values for the transformed spectral data were lower than for untransformed bands especially band 5. Forward stepwise linear regression procedure, including the TM original bands (5 variables) and the statistically significant transformed images (6 variables) was performed. Again, the result was that the only significant variable was band 5 itself.

APPENDIX 4.2

Descriptive statistics of Experiment 3 - spectral responses and species composition in the *Araucaria* forest

One-Way Analysis of Variance for Basal Area (whole plot (BA/plot) and only *Araucaria angustifolia* (BA_Aa))

a) Analysis of variance for BA/plot

```
Source      DF        SS        MS         F        P
Patch        3     2.398     0.799      0.94    0.425
Error       89    75.763     0.851
Total       92    78.161

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level       N      Mean     StDev  -+---------+---------+---------+-----
1          44    3.1327    0.8588                      (------*------)
2          32    2.9084    1.0214                 (-------*-------)
3           7    2.6843    0.6233   (----------------*----------------)
4          10    2.7360    1.0205        (-------------*--------------)
                                    -+---------+---------+---------+-----
Pooled StDev =    0.9226           2.00      2.40      2.80      3.20

Tukey's pairwise comparisons

Family error rate = 0.0500
Individual error rate = 0.0104
Critical value = 3.70

Intervals for (column level mean) - (row level mean)

                1           2           3

       2     -0.3365
              0.7851

       3     -0.5338     -0.7831
              1.4307      1.2314

       4     -0.4489     -0.7021     -1.2413
              1.2424      1.0470      1.1379
```

b) Analysis of Variance for BA_Aa

```
Source      DF        SS        MS         F        P
Patch        3     5.370     1.790      3.61    0.016
Error       89    44.185     0.496
Total       92    49.554

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level       N      Mean     StDev  -+---------+---------+---------+-----
1          44    0.6932    0.6048     (----*-----)
2          32    1.1419    0.8203                (------*-----)
3           7    0.9000    0.5825  (-------------*-------------)
4          10    1.3120    0.7862                (----------*----------)
                                    -+---------+---------+---------+-----
Pooled StDev =    0.7046           0.40      0.80      1.20      1.60

Tukey's pairwise comparisons

Family error rate = 0.0500
Individual error rate = 0.0104
Critical value = 3.70

Intervals for (column level mean) - (row level mean)

                1           2           3

       2     -0.8770
             -0.0204

       3     -0.9570     -0.5273
              0.5433      1.0111

       4     -1.2646     -0.8380     -1.3205
```

```
            0.0270      0.4977      0.4965
```

One-Way Analysis of Variance for Crown Diameter (CD_p)

a) Analysis of Variance for CD_p

```
Source      DF        SS        MS        F        P
Patch        3     23179      7726     5.56    0.002
Error       89    123655      1389
Total       92    146834

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev  --+---------+---------+---------+----
1           44     36.69     27.72    (----*---)
2           32     60.59     45.46              (----*----)
3            7     48.64     28.09  (----------*----------)
4           10     84.45     49.23                    (---------*--------)
                                    --+---------+---------+---------+----
Pooled StDev =    37.27              25        50        75       100
```

Tukey's pairwise comparisons

```
Family error rate = 0.0500
Individual error rate = 0.0104
Critical value = 3.70
Intervals for (column level mean) - (row level mean)

                1           2           3

       2      -46.6
               -1.2

       3      -51.6       -28.7
               27.7        52.6

       4      -81.9       -59.2       -83.9
              -13.6        11.5        12.3
```

One-Way Analysis of Variance for TM5 and TM7 spectral responses

a) Analysis of Variance for TM5

```
Source      DF        SS        MS        F        P
Patch        3     90.07     30.02     3.44    0.020
Error       89    776.97      8.73
Total       92    867.03

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev  -----+---------+---------+---------+-
1           44    15.228     3.332                            (----*-----)
2           32    13.823     2.747                   (-----*------)
3            7    13.517     1.259         (------------*-------------)
4           10    12.285     2.497  (-----------*----------)
                                    -----+---------+---------+---------+-
Pooled StDev =    2.955              11.2      12.8      14.4      16.0
```

Tukey's pairwise comparisons

```
    Family error rate = 0.0500
Individual error rate = 0.0104

Critical value = 3.70

Intervals for (column level mean) - (row level mean)

                1           2           3

       2      -0.391
               3.201

       3      -1.435      -2.920
               4.856       3.531

       4       0.235      -1.263      -2.577
               5.651       4.338       5.042
```

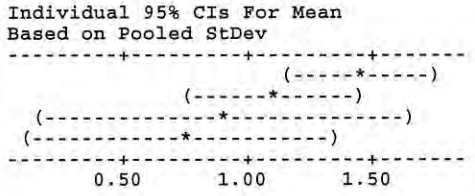
---

b) Analysis of Variance for TM7

Source	DF	SS	MS	F	P
Patch	3	5.94	1.98	1.97	0.125
Error	89	89.71	1.01		
Total	92	95.65			

Level	N	Mean	StDev
1	44	1.458	1.073
2	32	1.107	0.965
3	7	0.917	0.932
4	10	0.730	0.822

Pooled StDev = 1.004



Tukey's pairwise comparisons

Family error rate = 0.0500
Individual error rate = 0.0104

Critical value = 3.70

Intervals for (column level mean) - (row level mean)

	1	2	3
2	-0.259 0.961		
3	-0.528 1.610	-0.906 1.286	
4	-0.192 1.648	-0.575 1.328	-1.107 1.482

APPENDIX 5.1

List of vegetation indices calculated in preliminary stages of work

In **Chapter 5**, the Vegetation Indices (VIs) RATIO and NDVI were calculated and discussed. However, previous investigations were taken in this study, with the objective of selecting good performances, among the numerous suggestions available. Eventually, some modified indices using different bands were also tested. The indices were supposed to be capable of contribute at least with one of the following objectives: a) general cover-type discrimination, b) coniferous versus deciduous vegetation discrimination; c) discrimination between *A.angustifolia* plantation and the two species of the genus *Pinus* spp most frequently planted in the region, *P.taeda* – loblolly pine and *P.elliottii* – slash pine; d) discrimination between *Araucaria* forest and other type of forest cover composed only by deciduous species (sub-tropical forest or riparian areas); and, finally e) investigation about the discrimination of clusters of *A.angustifolia* in the *Araucaria* forest at pixel level. The VIs tested were:

NDVI = ((TM4 - TM3) / (TM4 + TM3))

GNDVI = ((TM4 - TM2) / (TM4 + TM2))

MVI = ((TM4 - TM5) / (TM4 + TM5))

SAVI = ((((TM4 - TM3) / (TM4 + TM3 + 0.5))* 1.5)

TTVI = (SQR((TM4 - TM3) / (TM4 + TM3)) + 0.5)

VCI = (TM7 / TM4)

GI = (TM4 / TM3)

RATIO = (TM3 / TM4)

RVI = (TM4 / TM3)

NRVI = ((RVI-1)/ (RVI + 1)))

MIRI = ((TM7 - TM3) / (TM7 + TM3)) * 127 + 128

MIRIm = ((TM5 - TM3) / (TM5 + TM3)) * 127 + 128

TC1B = 0.4395(TM1) + 0.5416(TM2) + 0.5406(TM3) + 0.4647(TM4) + 0.0704(TM5) + 0.0169(TM7)

TC2G = -0.2254(TM1) - 0.2054(TM2) - 0.3599(TM3) + 0.8795(TM4) - 0.0437(TM5) - 0.0461(TM7)

TC3W = 0.4154(TM1) + 0.1777(TM2) - 0.3618(TM3) - 0.0537(TM4) - 0.7526(TM5) - 0.3095(TM7)

Greenness = -0.2848(TM1) – 0.2435(TM2) – 0.5436(TM3) + 0.7243(TM4) + 0.0840(TM5) – 0.1800 (TM7)

Brightness = 0.3037 (TM1) + 0.2793(TM2) + 0.4743(TM3) +0.5585(TM4) +0.5082(TM5) + 0.1863(TM7)

Wetness = 0.1509 (TM1) + 0.1793(TM2) + 0.3279(TM3) +0.3406(TM4) -0.7112(TM5) - 0.4572(TM7)

APPENDIX 6.1

Fraction images obtained using the Unconstrained Least-Square method – model with three end-members



(a)



(b)



(c)

where:

(a) Fraction image of Vt1 – Component *Araucaria angustifolia* focusing the INF; (b) Fraction image of Vt2 – Component *Araucaria* forest; and (c) Fraction image of the Component Shade obtained using the Unconstrained Least-Square method (CLS).

APPENDIX 7.1

a) Classification Error Matrix (columns:reference data) (File:clas-supmaxver-dif9787-b5)

	severe_loss	moderate_loss	slight_loss	no-change	biomass_gain	omission	Totals
severe_loss	805 17.09%	91 1.93%	1 0.02%	0 0.00%	0 0.00%	77 1.64%	974
moderate_loss	51 1.08%	372 7.90%	15 0.32%	0 0.00%	0 0.00%	4 0.08%	442
slight_loss	0 0.00%	28 0.59%	409 8.69%	26 0.55%	0 0.00%	9 0.19%	472
no-change	0 0.00%	0 0.00%	74 1.57%	1126 23.91%	63 1.34%	0 0.00%	1263
biomass_gain	0 0.00%	0 0.00%	0 0.00%	1 0.02%	1557 33.06%	0 0.00%	1558
Totals	856	491	499	1153	1620	90	4709

	Producer's Accuracy	User's Accuracy
severe_loss	94.04%	89.74%
moderate_loss	75.76%	84.93%
slight_loss	81.96%	88.34%
no-change	97.66%	89.15%
biomass_gain	96.11%	99.94%

Classification Accuracy: 90.66 %
Comission errors : 7.43 %
Omission errors : 1.91 %
K(HAT) statistic : 87.80 %
TAU statistic : 88.32 %

b) Classification Error Matrix (columns:reference data) (file:cla-supmver-dif9787-sh.amo)

	severe_loss	moderate_loss	slight_loss	no-change	biomass_gain	Omission	Totals
severe_loss	681	0	9	0	0	4	694
	10.27%	0.00%	0.14%	0.00%	0.00%	0.06%	
moderate_loss	15	485	277	38	0	0	815
	0.23%	7.31%	4.18%	0.57%	0.00%	0.00%	
slight_loss	3	31	99	2	0	0	135
	0.05%	0.47%	1.49%	0.03%	0.00%	0.00%	
no-change	0	98	2	2179	10	13	2302
	0.00%	1.48%	0.03%	32.86%	0.15%	0.20%	
biomass_gain	0	0	0	3	2682	0	2685
	0.00%	0.00%	0.00%	0.05%	40.45%	0.00%	
Totals	699	614	387	2222	2692	17	6631

	Producer's Accuracy	User's Accuracy
severe_loss	97.42%	98.70%
moderate_loss	78.99%	59.51%
slight_loss	25.58%	73.33%
no-change	98.06%	95.19%
biomass_gain	99.63%	99.89%

Classification Accuracy: 92.38 %
Comission errors : 7.36 %
Omission errors : 0.26 %
K(HAT) statistic : 89.06 %
TAU statistic : 90.48 %

APPENDIX 7.2

Araucaria angustifolia and *Araucaria* forest reserves

The *Araucaria* forest inside the INF

One of the current preoccupations about public areas is defining how the changes in the natural environment are monitored over periods of time (Ferguson, 1996). In this context, the objective of this section is: a) to obtain reliable field information about the past distribution (presence and density) of the *A.angustifolia* as a species in the protected area of INF; b) to determine the magnitude of any changes in distribution and extent of *A.angustifolia* as a species and the *Araucaria* forest in the period of 10 years, from 1987 to 1997, in the area of INF and adjacencies.

As mentioned in **Chapter 2**, according to the Brazilian legislation for Conservation Units, the competence of legislating on the matter of water, forests, hunting, angling and their exploitation is federal. States can legislate only complementarity to account for local peculiarities or supply gaps or deficiencies in the federal legislation. The National Forests (NFs) are conservation units but, in contrast to National Parks (virtually untouchable) as they are supposed to be in constant production through sustainable development concepts. The NFs were designed to a rational utilisation of the natural renewable resources, production and research, respecting the mechanisms of ecosystem support. According to Barrett and Curtis (1999) these areas might aptly be termed "protected landscapes". The INF is a productive unit; as mentioned in **Chapter 2**, it commercialises wood originated from the planted forests and produces honey and mate-tea.

However, there are several competing uses of the land in such environments, including protecting forests whose age and structure are characteristic of a mature or ancient ecosystem. Some of the uses are conflicting with nature, as mentioned by Ferguson (1996). However, part of the area is composed of patches of *Araucaria* forest which are among the last remnants of the mature ecosystem in the region and monitoring changes in the area is an important task. The INF is more than just a productive forest farm; it provides areas of refuge for wildlife and recreation for the region urban population. Therefore, an inventory should be considered the baseline in the process of use control in the INF. With this objective, 93 plots were installed in the natural *Araucaria* forest areas. The INF has 1,508.09 ha of native forest, but they are not continuous. This has resulted in the installation of the plots in four patches, three of them in continuous areas and one in three small areas close of each other. The results of the forest inventory and the relationship between structural variables (basal area, volume), derived variables (crown diameter and crown area) and spectral responses were already discussed in

the **section 4.4.3**. It was concluded that there was insufficient evidence to say that the patches differ in total basal area. However, *A.angustifolia* is so imposing as a species that when analysing only its population there was evidence of differences in basal area in patches 1 and 2 and differences in crown diameter in patches 1 and 2 and 1 and 4. When the spectral responses were included in the analysis, it was found that patches 1 and 4 were different when band 5 was the analysed variable.

In this section, some investigations had to be made when analysing: a) the available information about the past distribution (presence and density) of the *A.angustifolia* as a species in the protected area of INF – objective (a); and b) the determination of the magnitude of changes in distribution and extent of *A.angustifolia* as a species and the *Araucaria* forest in the period of 10 years, from 1987 to 1997 in the area of INF and adjacencies – objective (b). The first step was to determine the existence of changing process inside the area of the INF. The previous analysis showed that the only activity detectable by the sensors was related to the natural forest successional dynamics and the regeneration of natural vegetation in the abandoned areas.

Testing differences in the plots between dates

Some statistical approaches were designed to capture information about changes using the spectral response in 1987 and in 1997 in the same 93 plots. Basic descriptive statistics of the spectral variables included in the analysis are presented in **Table 7.2.1**.

Table 7.2.1 Descriptive statistics of the spectral variables for 1987 and 1997; comparative analysis between the two dates using Matched pairs *t*-test

Variable	Mean	StDev	Min	Max	
TM1-87	17.937	0.423	17.000	20.000	
TM2-87	11.473	0.761	9.800	13.440	
TM3-87	7.244	0.720	6.000	10.100	
TM4-87	38.850	4.795	28.770	58.550	
TM5-87	14.299	3.070	6.660	27.880	
TM7-87	1.218	1.020	0.000	6.110	
TM1-97	18.774	0.554	18.000	20.000	
TM2-97	12.366	1.325	10.000	22.000	
TM3-97	11.129	1.329	10.000	22.000	
TM4-97	33.215	4.535	22.000	50.000	
TM5-97	17.323	3.692	7.000	30.000	
TM7-97	2.215	1.121	1.000	8.000	
T-Test of the Mean; N= 93					
Test of μu = 0.000 vs μu < 0.000					
Variable	Mean	StDev	SE Mean	T	P
Differb1	0.837	0.641	0.066	12.60	0.0
Differb2	0.892	1.362	0.141	06.32	0.0
Differb3	3.885	1.380	0.143	27.16	0.0
Differb4	5.635	2.235	0.232	-24.32	1.0 (*)
Differb5	3.023	2.281	0.237	12.78	0.0
Differb7	0.996	0.727	0.075	13.22	0.0

The matched pairs *t*-test is recommended in the *before-after experiments* in which the same subject serves in both groups, or when the same data set is used in different times. The matched pairs are reduced to a single sample by taking the difference between the two observations in each pair (Kitchens 1998). Theses differences form a single sample that can be analysed using the one-sample technique. Then, the variable Differ for each band is found by:

Differbx = (1997 spectral response in band *x*) – (1987 spectral response in band *x*).

There were increases in the absolute reflectance values for the bands 3, 5 and 7. In the case of band 4 there were decreases in the absolute values for reflectance, when measuring both dates. Only slight positive differences between the absolute values of reflectance (1987 and 1997) were found for bands 1 and 2. The results showed that the null hypothesis: that the difference between 1987 and 1997 is zero must be rejected (P value less than 0.05). The differences therefore can be explained by the natural differences in the overall scene in the period of 10 years time and by the natural forest succession dynamics. Values for band 4 differ b4 (*) in **Table 7.2.1** presented atypical numerical performance comparing to the others bands because actually it presents an inverse relationship (absolute values for 1997 were smaller than the values measured for 1987).

Considering the results above, a linear regression was performed for bands 4 and 5 to test how predicable the values of 1997 should be using the values of 1987 as independent variables. **Table 7.2.2** presents the results.

Table 7.2.2 Linear regression performed using information from 1987 as an independent variable and information from 1997 as a dependent variable

Regression: TM5-97 = 3.78 + 0.947 TM5-87				
Predictor	**Coef**	**StDev**	**T**	**P**
Constant	3.778	1.136	3.33	0.001
TM5-87	0.94720	0.07771	12.19	0.000
S = 2.288	R-Sq = 62.0%		R-Sq(adj) = 61.6%	
Analysis of Variance				
Source	DF	SS	MS	F	P
Regression	1	777.89	777.89	148.58	0.000
Error	91	476.43	5.24		
Total	92	1254.32			
Regression: TM4-97 = 0.64 + 0.839 TM4-87				
Predictor	**Coef**	**StDev**	**T**	**P**
Constant	0.637	1.794	0.36	0.723
TM4-87	0.83857	0.04583	18.30	0.000
S = 2.108	R-Sq = 78.6%		R-Sq(adj) = 78.4%	
Analysis of Variance				
Source	DF	SS	MS	F	P
Regression	1	1487.5	1487.5	334.86	0.000
Error	91	404.2	4.4		
Total	92	1891.7			

For both cases (band 5 and band 4) slope and intercept are highly significantly different from zero. The coefficient of determination R^2 = 62% for TM5 and R^2= 78.6% for TM4 represents reasonable proportions of 1997 data being explained by 1987 data. The analysis of variance compares the variation explained by the best-fit line with the residual variation and the null hypothesis is that none of the variation in 1997 data can be explained by the regression line. In both the cases, the relationship is highly significantly different from zero.

Figure 7.2.1 presents the INF map, with the 93 plots allocated in the four patches. To make a more focused analysis, two areas were selected: they were called patches 3 and 4. They were smaller than patches 1 and 2 and consequently easier to be analysed.



Figure 7.2..1 Location of the plots in the INF. Labels 1 to 4 represent the four patches of *Araucaria* forest.

In **Chapter 5** it was possible to determine that concentrations of coniferous species can be visualised in the natural stands of *Araucaria* forest using the second component of PCA from bands 4 and 7 or the ratio image from bands 4 and 7. In this chapter, the question is: a) is the same methodology adequate for the 1997 data set?; and b) are the results for 1987 image specific? Considering that changes inside the area due to external causes were not expected, it should be possible to perform the same methodology and obtain the same tendency of discrimination of *Araucaria angustifolia* (conifer) in the natural stands of *Araucaria* forest.

Araucaria angustifolia in patch 3

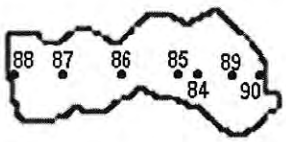
Figure 7.2.2 presents: Component 2 (PC2) of the PCA from band 4 and band 7 (details in **section 5.3.1.2.2**) for the image of 1987; (b) PC2 of the PCA obtained in analogue form for the image of 1997; (c) the visualisation of the plots number. The annexed **table** presents: CA_a (crown area - a variable which intends to represent crown closure) %NA (the percentage of the number of trees of the species *A.angustifolia* in the plot) and %BA (the percentage of basal area of the species *A.angustifolia*); (d) scanned picture of the aerial photograph, where the presence and concentration of *A.angustifolia* was photointerpreted. Areas in the photograph where the species is concentrated were circled by a white line.



(a)



(b)





(d)

Plot	CA_a	%NA	%BA
84	0	0	0
85	5.13	22	40
86	1.04	12	14
87	4.39	38	54
88	2.73	15	27
89	4.91	29	38
90	6.20	27	46

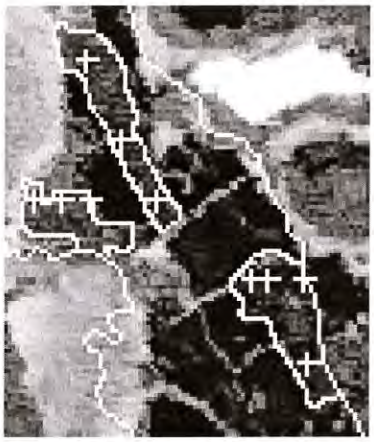
(c)

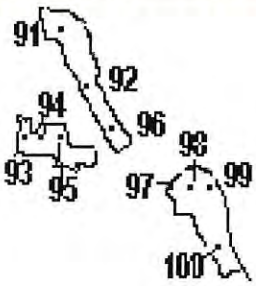
Figure 7.2.2 (a) location of the plots in patch3 over the PC2 of PCA from bands 4 and 7 for 1987; **(b)** location of the plots in patch 3 over the PC2 of PCA from bands 4 and 7 for 1997; **(c)** sketch of the plot and table with information about occupation of the plot with *A.angustifolia*; **(d)** aerial-photograph of the same patch and some the natural clusters of *A.angustifolia* in the *Araucaria* forest.

Araucaria angustifolia in patch 4

The same methodology was repeated in patch 4 (**Figure 7.2.3**). Actually, the patch is not a continuous area as is in patch 1, 2 and 3, but the small fragments are highly concentrated in terms of the presence of the species *Araucaria angustifolia*.









Plot	CA_a	%NA	%BA
91	7.38	28	42
92	4.01	19	33
93	4.35	19	38
94	0.67	03	08
95	4.58	13	34
96	2.56	22	29
97	8.15	93	98
98	11.63	60	79
99	3.86	61	85
100(*)	6.60	76	50

Figure 7.2.3 Allocation of the plots in patch 4. Images from 1997 and 1998 PC2 of PCA from bands 4 and 7, sketch of the plots, statistics about the plots and the aerial-photograph showing the areas with more concentration of *A.angustifolia*
(*) as mentioned before, 93 were available of the planned 100. Seven plots were not installed because they actually were positioned outside the INF. However, the original labelling was maintained.

The contiguous areas along the patches are formed by other category of vegetation; this fact helped its identification as highly dense and consequently contribute with the definition of the typical patterns (tone and texture) for the areas with and without *A.angustifolia*.

The tendency of the visual analysis is to confirm that PC2 of PCA from band 4 and band 7 discriminates coniferous species from deciduous species quite well. The conifers that normally grow in the region are *A.angustifolia, Podocarpus lambertii*, *P.elliottii* and *P.taeda*. The two species of pine are exotic species and will never appear in the same area as natural patches of *Araucaria* forest. *P.lambertii* belongs to the same biome as *A.angustifolia,* but normally grows up in clusters (many trees of the same species) and was not inventoried in the INF. Therefore, the conifer detectable in the PC2 image must be *A.angustifolia*. The superposition of the plots over the image made possible a visual analysis comparing tone, texture and the percentage of the number of trees of *A.angustifolia* in each plot. The variable percentages of basal area (%BA) and crown area of *A.angustifolia* were also used in the interpretation. All the concentrations of dark points in the image coincided with plots with high density of paraná-pine, in both data sets (1987 and 1997). Lighter areas in the images normally presented much less trees of paraná-pine (or less basal area) than the darker areas. The image ratio between band 4 and band 7 should also be used, but it seems to be little saturated. The methodology should be tested in other areas. However, it seems that it is possible to monitor presence of the species by this procedure.

Analysing the objective (b), the nature of changes might be divided in three categories: a) changes caused by the natural grow processes; b) changes caused by physical events like floods, fires, mortality, desease, earthquakes, etc; and c) changes caused by human interference. The INF have around 1,500 ha of *Araucaria* forest, and the conclusion of this section is that, in the 10 years period, there were no negative changes in the INF detectable by the procedures adopted in the present work. Some positive changes were verified as a consequence of the process of natural regeneration.

The *Araucaria* forest outside the INF

Forest fragments need management strategies and practices in order to prevent the continuation of current processes of forest degradation and loss of biodiversity (Viana *et al.*1992). In regions with intense agriculture, protection against anthropogenic disturbance of these fragments is unlikely to be enough. Monitoring and management techniques will be important in the process. According to Miller (1996), emerging scientific knowledge and local experience suggest that biodiversity will be retained to the extent that whole regions are managed co-operatively among protected areas and other uses. In Southern Brazil, a dominant problem in this decade is how to manage the few natural *Araucaria* forest fragments that remain. Some of these fragments are very small and are becoming degraded in terms of biodiversity because of intensive human interference. The challenge is how to reconcile economic

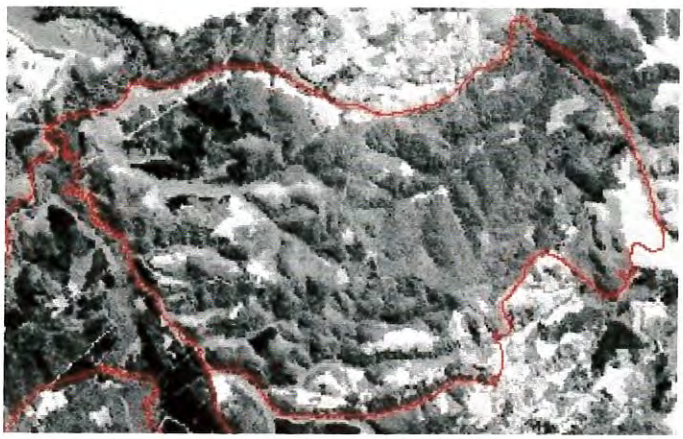
development and conservation of biological resources and how to use the natural resources without destroying the possible choices of future human generations. Even if not enough, important elements in the efforts to save natural biodiversity are the establishment and maintenance of protected areas and the sustainable management of the remaining areas.

The study area is not large enough to be representative of the whole biome, but it is recognised that just a few areas remain with *Araucaria* forest in its primitive or hardly disturbed state. Jacobs (1997) mentioned in his conclusions the urgency of "the establishment of the first Conservative Unit in the biome *Araucaria* forest in Estate of Paraná, which is recommended since 1966 by the Inventário do Pinheiro no Paraná (Dillewijn (1966) and still not implemented". Jacobs (1999) stressed the importance of creating the first Conservative Unity of the Estate in the *Araucaria* forest. He commented that "each day it is more difficult to find areas which can guarantee the conservation of at least a small percentage of the *Araucaria* forest ". Our perception is that there are so few remaining areas with *Araucaria* forest that the ones (**Chapter 7 - Figure 7.1**) labelled as (6) and (7) should be involved in a process of area protection by law. They ideally could be summed to the area of the INF and the whole area should be converted legally to any form of "Environmental Protection Area".

This should avoid problems that have occurred in the area labelled as (10). That area is a sad example of what might happen to the few remnants of *Araucaria* forest if immediate action is not taken. In the aerial photographs from 1980 (**Figure 7.2.4a**), the area was totally covered by *Araucaria* forest; in some places it already presented characteristics of a secondary forest and in others it still presented characteristics of old-growth forests. In the image of 1987 (**Figure 7.2.4b**), the differences in tone and texture between the patches of *Araucaria* forest inside the INF and the disturbed area is apparent. It was clear that most of the area was at that time covered by a "managed" forest. Finally, the image of 1997 (**Figure 7.2.4c**) showed that the area was converted to another kinds of land use. If the management allowed by law was followed, certainly the differences in forest density should not be so remarkable.



(a)



(b)



(c)

Figure 7.2.4 Evolution of deforestation of *Araucaria* forest in the neighbourhood of the INF: **(a)** aerial photograph (05.1980); **(b)** TM band 5 (06.1987); and **(c)** TM band 5 (08.1997). The images have been subjected to a histogram equalisation contrast stretch for printing purposes. The line in red encloses the devastated area.

Current status of the area

Personal information gathered at the INF demonstrated that the area was partially invaded by the so called "landless labourer" (as defined by Roriz (1998)). This is an extremely active movement in Brazil. The 'landless movement' (MST Movimento Sem-Terra) is one of the most radical and militant mass organisations in the world today. According to Sant'anna (1998), in 1984, Cascavel, Paraná, was the venue for the 1st National Meeting of the MST. These origins explain the predominance in the organisation of activists from the South of Brazil. The majority have already got their land, but some might become MST professionals (Sant'nana 1998). In the same way, the Catholic Church has been prominent since the beginning, through the Basic Christian Communities (Comunidades Eclesiais de Base) and the Pastoral Land Commission (Comissão Pastoral da Terra).

The movement is based on the indisputable fact that there is an excessive concentration of land and income in Brazil. According to Sant'anna (1998), mentioning the National Institute of Colonisation and Agrarian Reform (INCRA), 2% of rural owners hold 56% of the land and, of 452 million hectares of potentially agricultural land, 184 million are unproductive. For foresters, ecologists and conservationists, "unproductive" is a key word because forested areas are considered unproductive under certain definitions. Being unproductive, they can be clearcut. Some of the areas, which still have some kind of mature or ancient forest cover, belong to the Government and these lands are in danger as forest remnants if they are not transformed to protected areas by law. However, some contradictions arrive. The first is related to the Environmental Legislation, which regulates what is authorised to be "managed", and to what extent. The Brazilian Environmental Legislation is considered one of the most effective in the world. And the State of Paraná has also one of the best complementary legislations in the country. Then, why this scenario emerge? The answer is the eternal divide between theory and practice. Unfortunately, referring to conservation only on papers is not enough and the land use conflicts are increasing. As one of those involved in the conflict, the argument of the movement for using the 'invasion' strategy is that their aim is not to steal properties, but to force the Government to speed up the process of agrarian reform. In the meanwhile *Araucaria* forest is still a great preoccupation for all those involved with the conservation of nature.

In general, the scenarios are not too favourable to conservationists. The problem is serious and the movement is political. Probably the only possibility of interfering in favour of the *Araucaria* forest is also to act politically.

Recommendations

Final comments must address a problem that on the one hand is beyond the scope of this work but is certainly in the kernel of forest conservation philosophy, in Brazil. The issue is related to the destiny of the few remnants of *Araucaria* forest in Southern Brazil (and in the planet, since the biome is endemic to that region). In 1968, Reinhard Maack alerted the world to the consequences of the rapid deforestation of the *Araucaria* forest. Many efforts were dedicated in the 1970s and 1980s to characterising loss of the biome. Almost all relevant publications advocated the creation of: a) representative protected areas; b) mechanisms for the monitoring of the biodiversity in the existing reserves, which was prioritised in the International Workshop on Biodiversity - Monitoring in Federal Protected Areas: Defining the Methodology held in June of 1997; and c) mechanisms to decelerate and discourage the process of logging in the forest remnants, by augmenting the restrictions on selective felling. Actually, the legislation (Management Plans) in the Paraná State for logging the *Araucaria* forest is more restrictive than the other States' legislation. The main restriction is that a maximum logging of four species can be approved and the maximum exploitation cannot exceed 40% of the stock of trees with DBH equal to or larger than 40 cm. The summary of our approach is that a "minimum programme" to ensure the effective conservation of forest sites that support high levels of species diversity and endemism should be applied (Whitmore *et al.* 1994). The strategy of establishing reserves and parks is accepted, but they noted the fact that the viability of this strategy clearly depends upon the maintenance of reasonable environmental stability in the surrounding lands. The authors also stressed the fact that, if forests surrounding protected areas approach the structure and species composition of the natural forest, then more species will be maintained. The optimal form of land use outside parks and reserves is near-natural forests under sustainable management for timber or non-timber products (Whitmore *et al.* 1994). The natural forests of the Irati National Forest are the only "relatively" protected remnants of the biome *Araucaria* forest in Paraná State. Considering the preoccupations of authors such as Jacobs (1997) and the recommendations of the Workshop of Biodiversity in 1997 it might be a strong reason for declaring it untouchable, though it is a Federal Productive Unit.

Procedures for monitoring *Araucaria* forest using approaches highlighted in this thesis and the philosophy suggested by Whitmore (1994) should be incorporated in the State of Paraná Regional Planning of Development. Other areas should be incorporated to the web of "protected" or "semi-protected" areas in the *Araucaria* forest in the whole region of occurrence (Southern Brazil) by incentives to the creation of: a) new areas in the network of Paraná State Conservation Units; b) Private Reserves (whole area or part); c) Indigenous Reserves; and d) Municipal Protected Areas. This should increase the size of the protected units with areas where *A.angustifolia* and the threatened associated species (*Ocotea porosa* (Nees.) L.Barroso (imbuia) and *Ocotea pretiosa* (Nees.) Mez (canela-sassafrás), for example) are still certainly present.