School of Spatial Sciences

Multiscale Remote Sensing for Assessment of Environmental Change in the Rural-Urban Fringe

Graeme L Wright

"This thesis is presented as part of the requirements
for the award of the Degree of Doctor of Philosophy
of the Curtin University of Technology"

April 2000

ABSTRACT

The objective of this study was to investigate the application of multiscale satellite remote sensing data for assessment of land cover change in the rural-urban fringe. Inherent in this assessment process was the interpretation of multispectral data collected by several medium resolution satellite systems and evaluation of the quality of the resulting change information. Each dataset was acquired for a single date and classified at two levels of detail using standard classification algorithms. The optimum classification approach for each date was identified and the changes in land cover evaluated in several ways. The contribution of spatial and thematic errors and their propagation through the analysis process was investigated.

Data for this research were acquired over an area approximately 4.5 km square located in the southern metropolitan area of Perth, Western Australia. At the time of the initial data acquisition in 1972 the area was predominantly rural and comprised mostly dense pine plantations, however by the final stages of data acquisition in 1991, the area was almost completely given over to urban residential land use. Changes were interpreted from classified Landsat Multispectral Scanner (MSS), Landsat Thematic Mapper (TM) and SPOT (*System Pour l'Observation de la Terre)* High Resolution Visible (HRV) multispectral data, and were compared to reference maps compiled from medium scale aerial photographs. The geometric properties of high resolution panchromatic IRS1-D data were also evaluated to test the geometric potential of high resolution satellite data.

Supervised and unsupervised classification algorithms were used for derivation of land cover maps from each multispectral dataset at two levels of detail. Data were classified onto four general levels at the broadest (Level I) classification, and into nine levels at the finest (Level II) classification. The Kappa statistic and its variance were used to determine the optimum classification approach for each dataset and at each level of detail. No significant differences were observed between classification techniques at Level I, however at Level II the supervised classification approach produced significantly better results for the Landsat TM and SPOT HRV data. Classification at the more general Level I did not produce substantially higher classification rates compared to the same data at Level II. Additionally, higher spatial resolution data did not provide increased accuracy, however this was due mainly to a much greater

complexity of land covers present at the time the higher resolution Landsat TM and SPOT HRV data were recorded.

Land cover changes were assessed separately at Level I for all datasets, and also between Landsat TM and SPOT HRV data at Level II. Integrated multiscale assessment of land cover change was undertaken using classified Landsat MSS data at Level I and Landsat TM data at Level II. This enabled the continuity of change to be established across classification levels and sensor systems, even though there were variations in the level of detail extracted from each image.

The sources of spatial and thematic errors in the data were investigated and their effects on change assessment analysed. The evaluation of high resolution panchromatic satellite data emphasised the contribution to the analysis of spatial errors contained within the reference data. The multiscale data also indicated that combined propagation of spatial and thematic errors requires investigation using appropriate simulation modelling to establish the influence of data uncertainty on classification and change assessment results.

This research provides useful results for demonstrating a process for the integration of information derived from remotely sensed data at different measurement scales. Availability of data from an increasing range of remote sensing platforms and uncertainty of long term data availability emphasises the need to develop flexible interpretation and analysis approaches. This research adds value to the existing data archive by demonstrating how historical data may be integrated regardless of the spectral and spatial characteristics of the sensors.

ACKNOWLEDGEMENTS

I would like to take this opportunity to acknowledge the assistance provided by many people who have facilitated the completion of this dissertation. To those I have not mentioned by name, I say thank you, and acknowledge their valuable contribution of advice, support or material assistance.

Sincere gratitude is expressed to Professor Graham Lodwick for his continued support and encouragement during the period of this research. As supervisor and colleague he has alway encouraged self development and achievement, and has provided an academic environment which makes possible the fulfilment of these ambitions.

The assistance and support of my colleagues in the School of Spatial Sciences, Curtin University of Technology is appreciated. In particular, Mrs Nancy Antoine and Mrs Lori Patterson for their contribution to the finalisation of the manuscript, and Mr Jochen Franke in processing the aerial mosaics of the study area through the Helava digital photogrammetric workstation.

A range of data providers is acknowledged for the continued support of research and development at Curtin. Remote Sensing Services at the Department of Land Administration (DOLA) and the Australian Centre for Remote Sensing (ACRES) provided data from the SPOT and Landsat remote sensing satellites, and the Indian Space Research Organisation (ISRO) made available a sample image of IRS1-D data of the Perth area. Main Roads Western Australia (MRWA) supplied digital road centreline data of the Perth metropolitan area, which were used for the rectification of the IRS1-D satellite data. Data supplied by all of these organisations were crucial to the outcomes of this research and is gratefully acknowledged.

To my dear Irene, and sons Andrew and Nicholas, I express my deepest gratitude for your understanding during the preparation of this thesis. Without your constant support and encouragement, completion of this research program would not have been possible. Thank you.

TABLE OF CONTENTS

Page

Chapter **Page**

Chapter **Page**

Chapter **Page**

Chapter **Page**

LIST OF PLATES

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

Table **Page**

Table **Page**

Table **Page**

LIST OF ACRONYMS

ACRES	Australian Centre for Remote Sensing
AMG	Australian Map Grid
AVHRR	Advanced Very High Resolution Radiometer
C-CAP	CoastWatch Change Analysis Project, USA
DOLA	Department of Land Administration, Western Australia
ERS	Earth Resources Satellite
GCP	Ground Control Point
GIS	Geographic Information System
GPS	Global Positioning System
HRV	High Resolution Visible Sensor
IDRISI	Idrisi for Windows Version 2
IFOV	Instantaneous Field of View
IRS1-D	Indian Remote Sensing Satellite 1-D
ISODATA	Interactive Self Organising Data Analysis Technique
ISRO	Indian Space Research Organisation
LISS	Linear Imaging Self Scanning Sensor
MRWA	Main Roads Western Australia
MSS	Multispectral Scanner
NALC	North American Landscape Characterisation Program
NASA	National Aeronautics and Space Administration, USA
NDVI	Normalised Difference Vegetation Index
NOAA	National Oceanographic and Atmospheric Administration, USA
PAN	Panchromatic
PCA	Principal Components Analysis
RMSE	Root Mean Square Error
SPOT	System Pour l'Observation de la Terre
TM	Thematic Mapper
USGS	United States Geological Survey
XS	Multispectral sensor

Chapter 1

INTRODUCTION

This chapter discusses the concept of remote sensing techniques as an information source for temporal assessment of land-related development in the rural-urban fringe. Characteristics of near urban areas and their relationship to remote sensing techniques are outlined. The problem of the continuity of satellite remote sensing programs is discussed and the integration of data with a range of spectral and spatial resolutions is proposed. Image processing approaches for temporal assessment and analysis issues associated with multiscale assessment of remote sensor data are introduced. The objectives of the research are stated, as well as significance and benefits of the research. The chapter concludes with an outline of the methodology and thesis structure.

1.1 Background

1.1.1 Remote Sensing for Urban Studies

Remote sensing systems have the potential to detect, measure and assess scene variables based upon the response of the sensor to reflected or emitted radiation from the target. The major impetus for the application of remote sensing is to provide information that is not attainable by other means, or to reduce the cost of provision of the information. Trotter (1991) indicates that remote sensing can provide a very cost effective data acquisition and interpretation source for geographical information systems, while at the same time providing data that are *current, sufficiently accurate, comprehensive and available to a uniform standard over the entire area of interest.*

The modern era of digital remote sensing has developed from a sound basis of interpretation of aerial photographs using the classic parameters of interpretation - shape, size, pattern, tone, texture, shadow, site and association (Lillesand and Kiefer, 1994). Manual interpretation of analogue data relies upon the familiarity and adaptability of the human interpreter to consider all available information in the inductive process of interpreting remotely sensed data, and the deductive process of relating the interpretation to resource information requirements (Witmer, 1977).

Digital approaches to interpretation rely upon development of analysis techniques that utilise the unique quantitative characteristics of the digital data. Contrast stretching, classification and texture analysis emulate their analogue counterparts, while techniques including artificial intelligence approaches have been devised to facilitate incorporation of ancillary data and to assist with high level interpretation, and decorrelation stretches have been devised to reduce interference between multispectral bands. Regardless of the interpretation approach, features of interest are only detectable because they are correlated in a strong and predictable manner with their ground level spectral reflectance characteristics (Duggin and Robinove, 1990).

Correct interpretation may be achieved through assessment of direct or hybrid variables (Jensen, 1996). Direct interpretation of variables comes about when the remotely sensed data are able to provide fundamental biological or physical information directly. Measurement of position and height are common examples. Information regarding hybrid variables is obtained by measuring one or more biophysical variables, that when

modelled, provide the required information. Mapping of land use is one such hybrid variable where the final conclusion regarding specific land use classes is based upon assessment of the land cover classes present in the imagery (Green *et al*, 1994). It is only the spectral characteristics of land cover targets that are measurable via the remote sensor, and from these and other optional ancillary data, specific land use types may be identified.

Production of spatial interpretations and statistical summaries from remotely sensed data cannot be seen as an end in themselves. Trotter (1991) indicates that remote sensor data may form the basis of much input to geographic information systems (GIS) through the potential to not only lower the cost of the data acquisition and interpretation, but also to do so while meeting the other essential requirements for data in GIS. To fulfil this role the data must be current, sufficiently accurate, comprehensive and available to a uniform standard over the entire area of interest.

Digital interpretation approaches may be broadly categorised as either image enhancement or feature extraction techniques, depending whether they fulfil a role of improving the interpretability of the data, or serve to specifically identify features based upon spectral and/or spatial characteristics (Lillesand and Kiefer, 1994). Land cover changes were investigated by Byrne *et al* (1980) using Landsat MSS data analysed through the image enhancement process of Principal Components Analysis (PCA), with final target identification achieved through manual interpretation of the resultant images. Beaubien (1994) applied a similar analysis to Landsat TM data using a linear combination of spectral bands, then used the feature extraction process of multispectral classification to identify the targets. The specific analysis approach utilised depended upon an assessment of the target characteristics and sensor system parameters, and how these affect the clarity and content of the information contained within the data.

Remote sensing in urban and near-urban areas is influenced by a complex assemblage of disparate land covers which form a heterogeneous agglomeration of targets resulting from human activities, vegetation types, soils and water bodies (Forghani, 1994). As with many mapping problems, urban classes are represented by a continuum of cover classes rather than possessing well-defined boundaries (Forster, 1980b).

Forster (1985) indicates that the application of remote sensing approaches to urban studies are useful because they can provide regular and rapid updating of land use data, and are spatially more relevant for reporting urban development than census-type statistical data. Additional limitations to the application of remote sensing to landuse mapping of the rural-urban fringe do exist, and include the effects of atmospheric attenuation, inconsistent interpretation approaches, and temporal variations in the sensing parameters associated with the solar zenith angle, visibility and solar irradiance.

Assessment of the rural-urban fringe has been widely investigated by utilising a range of remotely sensed data and digital processing techniques (Barnsley and Barr, 1996; Forster (1980b); Haack *et al*, 1987; Jackson *et al*, 1980; Martin *et al*, 1988 and others). These studies individually encompass a range of data sources including a variety of spatial and spectral resolutions. Forster (1985) identified spatial resolution and temporal frequency as the major factors governing the level of information available from remote sensor data acquired from satellite platforms.

Digital analysis of remotely sensed data has broad application for interpretation of urban and near-urban areas. Well established processing techniques are available, and some sophisticated and innovative approaches have been developed for data integration and interpretation. With the launch of commercial remote sensing satellites, data with an extensive range of spectral and spatial resolutions are becoming available (Fritz, 1996), and multiscale data integration will become increasingly important. Spatial resolution has been considered with respect to target size and arrangement, with conflicting results. Spectral resolution has been investigated, but to date has been constrained by the limited range of sensing systems carried on remote sensing satellites. One of the challenges will be to develop techniques to combine the results of interpretation from these evolving sensor systems to provide coherent change mapping products.

1.1.2 Temporal Assessment

The ability to reliably detect and identify land covers is significantly affected by the temporal characteristics of the sensor. Foresam and Millette (1997) consider that although temporal considerations related to diurnal or seasonal timing may be complex, temporal analysis considerations are often more important than some spectral, spatial and radiometric factors.

Multitemporal remote sensor data may be analysed in two distinct ways for scene interpretation. Firstly, multidate images may be analysed in order to improve the quality of interpretation by incorporating knowledge of changes within the scene. For example, the deciduous nature of some trees may be used to separate deciduous from coniferous vegetation types. Alternatively, temporal variations in multidate images may be examined to determine actual changes that have taken place in the area of interest between imaging dates. This second approach is to be applied in this research and relies on the principle of *change detection*. Singh (1989) defines change detection as the *process of identifying differences in the state of an object or phenomenon by observing it at different times.*

The fundamental goals of change detection are to compare point by point spatial representations in time by controlling variances caused by external factors such as illumination, sensor calibration and variations in orbital parameters, and to measure change caused by differences in the variables of interest (Green *et al*, 1994). Control of the former will facilitate achievement of the latter. Careful assessment of all scene dependent and acquisition and processing variables is required to optimise the capabilities for change detection.

Jensen (1996) indicates that while high spectral and spatial resolutions are often required for object identification, detection of the occurrence of change but without identifying the specific change, is frequently less demanding. Consequently, there is potential that even with spatially complex urban and near-urban targets, areas of change may be identified if not the actual change that has occurred.

While visual interpretation of multitemporal aerial photographs and satellite images provides useful change information, there is scope for more automated approaches involving digital analysis. Ridd and Liu (1998) indicate there is a need for change detection procedures that will automatically correlate and compare multitemporal images for the purposes of highlighting change. Future availability of imagery acquired at short time intervals and with consistent image quality should provide adequate raw data for most short term change detection analyses. However, there is increasing uncertainty regarding the continuity of present remote sensing programs and sensing

systems, and therefore the potential to undertake long term monitoring using data sets from similar sensors, such as the Landsat TM, is threatened.

General guidelines for sensor systems utilised for temporal assessment include a systematic period between overflights, consistent time of overpass to minimise diurnal sun angle effects, similar scale and look angle geometry, and recording of radiant flux in a consistent manner and in relevant spectral regions (Jensen, 1996). The objective of these constraints is to minimise or control variations between epochs of data acquisition. However, even when using data from temporally adjacent epochs of a sensor such as SPOT HRV, variations still remain. Townshend and Justice (1995) state that even differences between images from the same calender day may be rather complex. Processing techniques and algorithms must be developed to account for these variables and to minimise their effect. Commercialisation of remote sensing satellite programs will lead to a greater variety of sensors and, for long term monitoring, comparison of data from a range of satellite systems may be necessary.

A wide range of approaches to change detection analysis has been reported (Fung, 1990; Green *et al*, 1994; Jeanjean *et al*, 1996; Jensen *et al*, 1993; Singh, 1989). Singh (1989) considers all change detection approaches can be classified into image processing techniques that utilise simultaneous analysis of multitemporal data, or postclassification comparisons.

Simultaneous analysis of multitemporal data includes a broad range of techniques, including image differencing, image ratioing, change vector analysis, multidate image change detection and image regression (Jensen, 1996). The implementation of each of these approaches is unique, however they are all designed to minimise differences between images caused by scene-sensor variations between epochs. For example, image regression is designed to account for differences in the mean and variance between pixel values for different dates so that adverse atmospheric conditions are reduced. Additionally, these techniques emphasise residual differences between the images that may then be identified as areas of change.

Postclassification comparison relies upon the comparison of classified images produced through independent analysis of multidate imagery. Data from each date are processed

using a suitable classification algorithm and classification scheme, and changes are detected by determining the differences between the resultant land cover maps. Singh (1989) indicates the method has potential as each land cover map is produced independently, thus minimising the effects of atmosphere and sensor variations between epochs. The approach is also relatively straightforward and is able to accommodate data acquired from different sensors at successive epochs.

Perceived limitations of both of these approaches include no account of spatial characteristics of the targets during analysis. A significant component of manual interpretation of images involves analysis of site, shape, size, pattern, texture and association, which are related to the spatial characteristics and contextual arrangement of features. Artificial intelligence and knowledge-based approaches to image interpretation and change detection have been investigated by many researchers. In the context of this discussion, most artificial intelligence and knowledge-based techniques utilise either of the above techniques of temporal assessment as an integral component of the analysis system, and therefore rely upon similar fundamental image processing principles. An advantage exists in their ability to utilise contextual ancillary data, however most systems still require significant human intervention (Jensen, 1996).

Change detection will be investigated in this research using postclassification comparison techniques designed to incorporate the spectral and spatial characteristics of multisensor and multiscale imagery. Specific approaches to image processing will be discussed in Chapter 4.

Representation of change areas is important to the overall temporal assessment process and provides the final measure of land cover transitional processes. Change representation varies between masking of change/no-change areas, through to explicit change specification via cross tabulation of successive images where every land cover transition is represented. A standard date-to-date comparison of n classes results in n^2 possible change classes. To overcome this problem Jensen and Toll (1982) proposed identification of a limited range of change stages for urban development which were identifiable on Landsat MSS imagery. Alternatively, Jensen *et al* (1993) detail the concept of a *change matrix* which uses a full crosstabulation of classes in both images,

but aggregates related changes and only highlights the major change classes within the change matrix.

Due to the large number of possible transitional classes, change representation may be a complex issue for production of information products. Consequently, change reporting and representation of the results of temporal assessment will be considered in Chapter 6.

1.1.3 Multiscale Assessment

Conventional approaches to interpretation of remote sensor data endeavour to apply statistically based interpretation techniques to the most suitable imagery available. Data selection is based upon the spectral, spatial, radiometric and temporal properties of available imagery balanced against the corresponding characteristics exhibited by the target. While researchers acknowledge that improved interpretation results are a complex function of spatiotemporal characteristics of the terrain, there is much conjecture as to the precise role of spatial resolution in the accuracy of target identification. Marceau *et al* (1994) indicate that finer spatial resolution data do not necessarily improve the per-pixel classification accuracy because identification varies according to the inherent spatial structure of the target. In contrast, Townshend and Justice (1988) state that improved resolution increases the capability for accurate target identification, and Collins and Woodcock (1996b) contend that resolution should be sufficiently large that some amount of within-class averaging takes place.

Spatial resolution is understood to have several related and varied effects on the outcome of the interpretation of remote sensor data. Remote sensing represents a particular case of an arbitrary and fixed sampling grid with inherent limitations of scale and aggregation (Marceau *et al,* 1994). This manifests itself through the information content of the data being constrained by the spatial resolution of the sensor (Jensen, 1996), and consequently the ability to detect and/or recognise features is limited (Marceau *et al*, 1994). For change detection purposes this effect may be more significant because two or more images are analysed with each contributing to the overall rate of interpretation error. The effect of spatial resolution on the interpretation of land cover classes will be considered in Chapters 4 and 5 using images derived from satellites acquiring data over a range of spatial resolutions.

Singh (1989) indicates that for change detection purposes the spatial resolution and consequently the registration of component images have direct consequences on the accuracy of change images where multidate data are combined or compared. Randomly distributed errors, which develop during registration of component images, and some residual systematic effects, such as those from sensor operation and relief displacement, are propagated to the change image and are manifest as apparent increased levels of change, which may be significant.

Researchers have evaluated the cartographic accuracy of Landsat MSS and Landsat TM data (Welch and Usery, 1984) and derived methods for error modelling of SPOT HRV data (Welch, 1985). Singh (1989) suggests that techniques involving the application of fuzzy boundaries be developed which require less precise registration of images or perhaps bypass the registration process entirely. However it is anticipated that registration accuracy will always be an important consideration. Optimum results are expected where the most accurate possible registration is achieved and geometric registration is minimised as a potential source of change error. The extent of registration error for various sensors will be investigated in Chapter 3, and its effect on change measurement will be analysed in Chapter 7.

Jensen (1996) acknowledges that the thematic information content of the data is limited by the spatial resolution of the sensor, as is the capability to provide accurate geometric registration. Multiscale change analysis must therefore consider the geometric and interpretation effects of spatial resolution upon change detection products.

Thematic interpretation of multiscale remote sensor data must also consider the spatial scale or spatial frequency of the target (Townshend and Justice, 1995). This differs from the acquisition scale or spatial resolution of the sensor, as it is scene rather than sensor dependent. Several authors (Atkinson and Curran, 1997; Collins and Woodcock, 1996b; Jeanjean *et al, 1996;* Marceau *et al,* 1994) consider spatial scale or frequency of targets to be important in defining the required spatial resolution of the sensor. From global to local monitoring scales, the relationship between target structure or spatial frequency and the spatial resolution of the sensor determine the ability of the sensing system to identify the target. Marceau *et al* (1994) demonstrate this concept through variations in scene statistics by progressively aggregating high resolution airborne

multispectral scanner data. These concepts may be extended to satellite data and will be significant in the evaluation of the thematic information content of multiscale data for change detection analysis.

Integration and comparison of interpretations from multiscale analysis must be considered with respect to different levels of information content and variations in boundary locations. Townshend and Justice (1988) and Trotter (1991) consider the issues of data generalisation when dealing with multiscale interpretation, and suggest approaches for resolving the change measurement and representation problems. The most significant issues relate to positional variations and the decreasing level of boundary detail with coarser spatial resolution.

Evaluation of the effects of the spatial scale of the targets and the resultant variations in land cover boundaries will be investigated in Chapters 6 and 7 when assessing methods for reporting land cover changes and the effects of error propagation.

1.2 Problem Statement

Landsat 1 represented the first source of satellite-based remote sensor data suitable for earth resources assessment that was part of a systematic worldwide data collection program. Satellites from subsequent programs such as SPOT, National Oceanographic and Atmospheric Administration (NOAA), Earth Resources Satellite (ERS) and a number of others have been implemented, and many more are planned within the next decade. Each system is designed to collect data for identification and analysis of a range of targets and for specific functions according to particular system specifications. For change detection purposes, continuity of data collection by a particular system is of significant interest. Already the Landsat MSS instrument has been discontinued and the Landsat TM sensor is also under review.

The long-term viability of many commercial satellite remote sensing programs is not certain. Consequently, it is likely the future for high quality satellite data may be characterised by a range of systems which have only a short presence in the market place. A need exists for a thorough evaluation of the prospects for utilising multiscale satellite data for multitemporal assessment of land cover information.

A range of methods is available for change detection, however many rely upon evaluation of combined datasets derived from multiple epochs. These include PCA, tasselled-cap analysis, combined classification techniques and image differencing techniques. The basis of all of the above approaches is the consistent spatial, spectral and radiometric qualities of the data resulting from sensing with an instrument of similar specification. Where dissimilar sensors are utilised, substantial differences exist in all sensor specifications, in particular spatial and spectral resolution, and the above combined approaches are no longer appropriate.

Due to considerable differences in the spectral, spatial and radiometric characteristics of the data, analysis must involve separate interpretation of each dataset. Within this context, postclassification analysis is appropriate for evaluation of land cover changes from multiscale data.

Processing of multisensor and multiscale data provides for variations in classes derived from the resulting interpretation that must be accommodated when change analysis is undertaken. For each data source a discrete range of classes may be identified, however these classes must be reconciled for comparison of interpretations derived from different sensors in order for change detection to be effective. A coherent and consistent land cover classification scheme and change reporting methodology is required for this purpose, which must make it possible to identify important transitions that have taken place between epochs.

Accurate geometric registration of multidate data has been identified as essential for all change detection purposes. Registration errors directly affect any assessment of land cover change and result in many areas of false change recorded in change detection statistics. Comparison of multiscale data further complicates the process because each dataset contains errors of location inherent to the sensing system. Classification errors contributed by the interpretation approach and spatial errors due to the spatial resolution of the sensor and the sampling interval adopted during rectification are also important. Modelling and evaluation of these errors is necessary in order to assess the reliability of change detection statistics derived from multiscale satellite data.

1.3 Research Objectives

1.3.1 Aims of the Research

The main objective of this research is to investigate the application of change detection techniques to multiscale remote sensor data. Spectral and spatial properties of the data are investigated in order to evaluate the potential of change detection using different satellite sensors. The classification accuracy of each sensor is evaluated against known land cover distributions derived from aerial photographs, and the spectral separability of the data is evaluated to optimise the performance of the classifier. The contribution of thematic and spatial errors caused by sensor sampling and geometric registration is also evaluated. An analysis of the thematic and spatial accuracy of the final land cover change detection image is also completed.

The objective of this study is therefore to test the hypothesis:

That multiscale remotely sensed digital data can be processed to accurately determine land cover changes on the rural-urban fringe.

1.3.2 Significance and Benefits of the Research

This research provides an image processing and change assessment approach that can be applied to land cover change analysis using multiscale satellite data. Evaluation of the reliability of the multiscale approach to change detection provides future users with an alternative to the standard temporal assessment methods, and enables digital data from a range of sensors to be interpreted for derivation of land cover change statistics. This will overcome limitations on the assessment of change caused by current approaches, which rely upon analysis of digital data from the same remote sensing system. The flexibility afforded will enable users to access a combination of data sources, especially where weather conditions and reception facilities may restrict access to regular monitoring information.

Over the next decade many commercial remote sensing satellites are planned which will acquire data suitable for land cover assessment. It is likely that for long term monitoring, data from a single sensor may not be consistently available and multiscale data integration will become increasingly important. The benefit of this research is that

it identifies the specific issues involved and analyses the contributions of a range of sensors to land cover change detection.

1.3.3 Research Methodology

This research is designed to investigate the potential to utilise remotely sensed data from sensors with different spatial and spectral resolutions for temporal assessment of land cover changes in the rural-urban fringe. An assessment of the suitability of the approach is based upon an evaluation of the classification accuracy and consistency of the data derived from various sensors, and the contribution to the results of the geometric properties of the sensor and the geocoding method applied. The sources of multiscale satellite information used for this research are Landsat MSS, Landsat TM, SPOT HRV multispectral data and IRS1-D panchromatic data. The first three datasets are utilised for thematic classification and geometric assessment, while the IRS1-D panchromatic data are used only for geometric evaluation.

The study comprises the following major components:

(i) Review the use of remote sensing for information extraction applied to temporal assessment, focusing on the spectral and spatial resolutions of satellite sensors and how these affect image interpretation. Classification accuracy, spatial and thematic error propagation and change reporting will also be evaluated;

(ii) Compile relevant Landsat MSS, Landsat TM, SPOT HRV and IRS1-D satellite data for the study area in a format suitable for analysis. Interpret aerial photographs for each satellite image acquisition date for use as reference data;

(iii) Geocode all satellite data to the Australian Map Grid (AMG) using digital planimetric data provided by DOLA and MRWA, and evaluate rectification accuracy using independent control points for verification;

(iv) Define land cover classes suitable for multiscale assessment based upon a standard classification system and considering the spectral and spatial resolutions of the satellite data. Select and refine training sites for image classification;

(v) Assess the accuracy of each classification of multiscale satellite imagery and evaluate the allocation of classes considering variations in spatial resolution;

(vi) Determine land cover changes between epochs and evaluate change representation for the various spatial resolutions of satellite data by analysing the change matrices and their accuracy parameters;

(vii) Summarise the changes observed between sensing epochs and develop methods to compare classifications between datasets at multiple levels of abstraction and information content;

(viii) Evaluate the effect of combined spatial and thematic errors on change representation, and investigate appropriate methods for error modelling and assessment.

1.4 Thesis Structure

This research thesis comprises eight chapters. Chapter 1 considers the concept of change detection for near urban areas, image processing approaches for temporal assessment and analysis issues associated with multiscale assessment of remote sensor data. Problems associated with change detection using multiscale digital data are introduced and the application of image classification is discussed. The objectives of the research are stated, as well as the methodology and significance and benefits of the research.

Chapter 2 presents the study area and describes the physical characteristics of the area to be analysed. The derivation of ground reference data from aerial photographs taken in 1972/73, 1986 and 1991/92 is discussed, and the observed land cover transitions between satellite epochs is explained from a remote sensing perspective. The Landsat MSS data from 1972, Landsat TM data from 1986, SPOT multispectral (SPOT XS) data from 1991, and IRS1-D data from 1998 are discussed, along with the important characteristics of the sensors which are relevant to change detection analysis. Preprocessing of the data prior to analysis is also outlined.

The processes of rectification and resampling are detailed in Chapter 3. This chapter describes the available techniques for image rectification and outlines relevant factors to be considered in ground control point (GCP) selection and analysis of residuals. Resampling schemes are also considered and discussed with respect to establishing a common spatial resolution for the multiscale data and maintenance of a spectrally

coherent dataset. The spatial effects of image resampling are investigated and the precision of the rectified images is evaluated.

Chapter 4 reviews land cover classification strategies in the context of their application to multiscale analysis. The process of image classification is described and applied to the study area for each epoch. Detailed analysis of the spatial correlation and spectral separability of the land covers is performed. Results of the classification of each image are presented and a qualitative assessment of the classifications is undertaken.

Thematic mapping accuracy assessment methods are discussed in Chapter 5 and an assessment is made of the classification performance for each resolution of satellite data. Overall Classification Accuracy and Kappa Coefficient statistics are derived, and the optimum classification approach for each level of classification and for each image dataset is determined. Factors affecting the classification accuracy are discussed in the context of target and sensor characteristics.

Change detection techniques are reviewed in Chapter 6. The postclassification comparison approach is used to derive land cover change maps between epochs. Summary statistics of change are produced using change matrices and the land cover changes between dates are investigated. The effectiveness of change detection techniques using multiscale data is evaluated and the concept of change reporting as a means of measuring and communicating changes identified using remote sensing is considered.

Chapter 7 examines the sources of spatial and interpretive error in change maps when using multiscale remotely sensed data. An analysis of error propagation through image rectification, reference data collection and map overlay techniques used for the production of change maps is undertaken. Conclusions and recommendations for future research are provided in Chapter 8.

Chapter 2

STUDY AREA AND DATA PREPARATION

Change detection analysis requires a thorough knowledge of the characteristics of the study area in order that appropriate image processing techniques are employed and that interpretation of the results is relevant to the actual land cover changes that have occurred. An understanding of the specific properties of the satellite data utilised for change detection is also required, and they must be considered in conjunction with the image processing requirements. Reliable reference data are needed to provide guidance for selection of training samples and for comparison of classification results to evaluate the effectiveness of object identification. This chapter concentrates on the key characteristics of the satellite and reference data.

2.1 Introduction

Land cover changes are analysed in this study for an area situated within the City of Melville, which is located in the southern metropolitan area of Perth, Western Australia. The study area has undergone significant change over the time span of this study. Land covers have varied from natural vegetation and plantation forest through to remnant natural vegetation and medium density housing. Satellite imagery used within this study includes Landsat MSS, Landsat TM and SPOT HRV multispectral data, and IRS1-D panchromatic data. Additional data in the form of aerial photographs and analogue and digital maps were required to assist analysis.

Selection of a suitable study area for detailed analysis requires consideration of a range of important factors, including:

(i) Access to an archive of suitable remotely sensed data for analysis of geometric and thematic characteristics;

(ii) Convenient access to the study area for field checking of reference data and verification of image interpretation;

(iii) GCPs that are readily identifiable in the range of multiscale images investigated in this research and on the ground, which are suitable for the image rectification process;

(iv) Availability of ancillary data such as analogue and digital maps, and aerial photographs for compilation of reference data.

Compatible reference data are essential for comprehensive analysis of primary remote sensor information. Multiscale remotely sensed data were selected for this study in consideration of the limited lifespan of remote sensing systems to date and the likelihood of a multitude of satellite sensors becoming available in the future, but without any guarantees for long term data continuity. While future satellite sensors promise a comprehensive range of higher resolution data, at the time of this research the first new generation multispectral satellite sensors had only just been commissioned.

Planimetric data necessary for GCP identification and image registration are available from digital planimetric maps and hard copy topographic maps of the area. Aerial photographs corresponding to each imaging date provide detailed information regarding land cover changes within the study area over the duration of the study.

The study area and specific anniversary dates are important because they provide a temporal view of the changing land cover over an extended period of time. Location of the study area within the Perth metropolitan region facilitates access to relevant reference data in the form of medium scale aerial photography for comparison with the primary satellite data used to assess land cover change.

2.2 Characteristics of the Study Area

2.2.1 Location and Environment

The study area (Figure 2.1) comprises parts of the City of Melville municipality, which falls within the greater Perth metropolitan area. It is located approximately 11 km south west of the Perth city centre and occupies an area 4 km (E-W) by 4.5 km (N-S). In 1972 the study area was located on the southern fringe of the metropolitan region and has since been incorporated as part of the dormitory suburbs of Winthrop and Murdoch, as Perth gradually expands to the north, east and south.

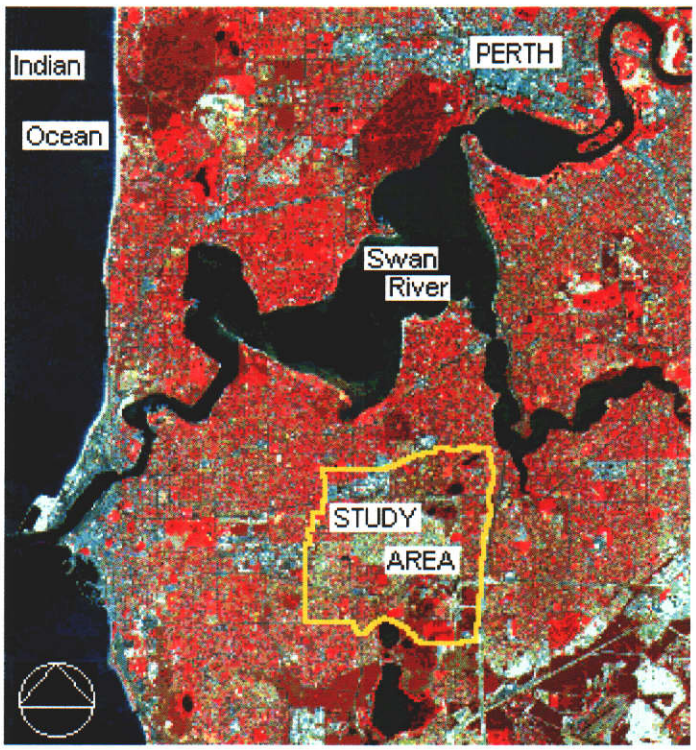


Figure 2.1 *Location of the study area. Scale 1:200 000*

Perth is a modern city of approximately one million people and is subject to continuing demands for increased land areas suitable for housing and infrastructure development.

The study area has undergone significant alteration in land covers during the period of this study, ranging from large areas of natural and plantation vegetation in 1972 through to almost complete development for urban land use in the early 1990s.

The study area is part of the Swan Coastal Plain bounded by the Darling Scarp to the east and the Indian Ocean to the west. The Plain is approximately 80 km wide and extends over 250 km from Gingin in the north to Dunsborough in the south. The Perth metropolitan area is situated in the central part of the Plain and in 2000 urban development extends from the coast to the Darling Scarp, and for approximately 40 km both to the north and south of Perth. The areas of Murdoch and Winthrop are now considered near-city suburbs.

The study area is located on the western Bassendean dune system (Seddon, 1972) formed from leached quartz sands where the soils are considered very infertile and are very low in calcium, iron and most other minerals. The topography of the area is characterised by low sand dunes up to 30 m in elevation, which have been stabilised over time by natural vegetation. The soils of the Bassendean system are excessively drained and retain little or no moisture. Between the sand ridges lie a series of interdunal depressions that form the only natural water features within the study area. Moisture is retained in these locations through an impermeable hard pan, and they are generally saturated with water for much of the year and maintain a specific range of vegetation (Seddon, 1972). Many of these areas have been retained today and several occur within the study area.

The natural vegetation of the Bassendean dune system mainly comprises woodland with a canopy of banksia species and a diverse understorey of flowering shrubs. Banksias grow to approximately 5 m and possess a very open canopy structure which, from a remote sensing perspective, provides a target comprising a complex admixture of plant and soil components. The Bassendean system has little agricultural value (Seddon, 1972), however considerable areas have been cleared for commercial pine plantations. A significant proportion of the study area has been put to this use and clearing of the plantations has been a precursor to subsequent urban development.

2.2.2 Land Cover Types

Land covers within the study area were surveyed to determine the characteristic surface features for recognition in the remotely sensed data. Site visits were made and representative areas identified and correlated with patterns in the aerial photographs to assist with image interpretation and for reference data compilation. The study area comprises the following major cover types:

(i) Woodland,
(ii) Interdunal wetlands and fringing vegetation,
(iii) Pine plantation,
(iv) Grassland,
(v) Bare ground,
(vi) Residential,
(vii) Commercial,
(viii) Recreation.

Plate 2.1 shows an area of *Woodland* with banksia trees providing the canopy and a combination of low flowering shrubs forming the understorey. All plants in this area, except for some low growing native orchids and introduced grasses, are perennials and maintain leaf cover throughout the year.

The overall density of vegetation cover of *Woodland* contrasts with the fringing vegetation of the interdunal wetlands as indicated in Plate 2.2. *Wetlands* comprise a relatively dense vegetation cover formed by a canopy of paperbarks up to 7 m tall, but a much denser understorey of ti-tree and reeds approximately 1.5 m high. Together these provide an almost 100 percent ground cover as viewed from a remote sensor. Areas of open water (*Water* class) in these wetlands are few in number and generally occur only in the larger wetlands (Plate 2.3) or due to dredging for water supply within the pine plantations.

Plates 2.4 and 2.5 show representative samples of *Pine plantations* similar to those originally found in the study area. A relatively high density of planting is shown in Plate 2.4 whereas a thinned pine plantation is shown in Plate 2.5. Mature pine plantations are almost totally devoid of understorey, although areas subjected to extensive thinning have a sparse understorey of shrubs and annual grasses.

Plate 2.1 *Woodland of the Bassendean dune system showing banksia canopy and flowering shrubs of the understorey*

Plate 2.2 *Wetlands indicating dense vegetation cover without open water*

Plate 2.3 *Wetlands indicating extent of open water areas with fringing vegetation*

Plate 2.4 *High density planting of pine plantation*

Plate 2.5 *Pine plantation following thinning operations*

Plate 2.6 *Grassland area after clearing of pine plantation or woodland, and following growth of annual grasses*

Plate 2.7 *Bare ground following clearing and prior to construction or growth of annual grasses*

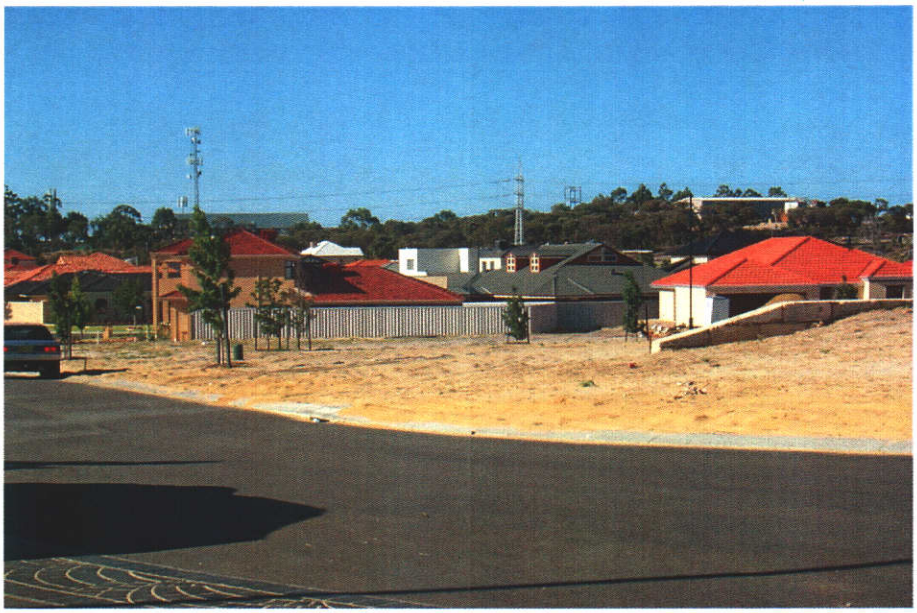


Plate 2.8 *Partially constructed residential area showing areas of bare ground*

Plate 2.9 *Residential construction showing clay and concrete roof tiles and surrounding sparse vegetation*

Plate 2.10 *Residential construction showing steel roof and surrounding dense vegetation*

Plate 2.11 *Commercial site indicating mixture of buildings and paved areas*

Plate 2.12 *Irrigated area of Recreation class associated with Commercial class and fringing trees*

Plate 2.13 *Irrigated pasture located within the Murdoch University Veterinary Science complex*

Plate 2.14 *Residential, Recreation and Wetland classes in close proximity*

Grassland areas in this region occur largely as a transitional land cover between other uses. For example, following removal of pine plantations and prior to recreational or residential development, land left fallow for at least the winter and spring seasons (June to November) are soon covered by annual grasses that grow actively until early summer. Plate 2.6 shows an example of temporary grassland development in this transitional stage. Like *Grassland*, *Bare ground* also occurs during a transitional stage. Plate 2.7 shows an area of *Bare ground* following clearing for residential development, but prior to the commencement of housing construction. Plate 2.8 indicates a further transitional phase, where residential construction is well advanced on some lots, but not others, with the area almost totally devoid of vegetation.

Residential developments within the study area generally comprise detached housing constructed on allotments varying in area from 500 to 800 sq m. Most structures have a roof area of approximately 150 to 250 sq m, predominantly of clay or concrete tile construction, but less frequently made of fibre cement or steel materials. Plates 2.9 and 2.10 indicate variations in the construction materials and adjacent land covers associated with residential developments. Variations in vegetation type and canopy cover between the two photographs are particularly worth noting.

Plate 2.11 shows an area typical of the *Commercial* class in the study area, which is characterised by large roof areas and surrounded by expansive areas of open space, generally paved with either asphalt or concrete. Roofing materials are mostly fibre cement or metal, and compared to residential areas, most commercial sites are almost totally devoid of vegetation. Factory or warehouse sites have very similar land cover characteristics to retail sites, and are typical of the example in Plate 2.11.

A number of recreation sites are included in the study area and characteristically contain a diverse range of large areas of open grassland, which are intensively managed. Additionally, many recreational areas also contain other vegetation types such as trees and shrubs. Plates 2.12, 2.13 and 2.14 show examples of the *Recreation* class that indicate the mixture of cover types present. The common element is the presence of large areas of vigorous herbaceous (grass) vegetation grown under irrigated conditions. Note also that some *Commercial* (education) development (Plate 2.12) is closely associated with various categories of the *Recreation* class, and that developments such

as sporting pavilions may be included within areas identified as *Recreation* on the reference data.

Plate 2.14 indicates the relatively close proximity and complex arrangement of land covers within the study area. Within a distance of 100 m land covers may change several times, in this case from *Wetland* to *Recreation,* and subsequently to *Residential.*

2.2.3 Transition Sequences

The previous section described the major land cover classes found within the study area during the period of data acquisition for this research. Given the spatial arrangement of features and the temporal nature of the research, there are important spatial and temporal transition sequences that influence the interpretation of remotely sensed data. Spatial transition sequences result from a continuum of cover classes between which it is difficult to identify clear boundaries of discrete nominal classes (Forster, 1985). Temporal transitions result from changes in land cover between sensing epochs, however due to the low frequency of data acquisition, the measured states of the targets may not represent adjacent states in the transition sequence (Jensen and Toll, 1982; Martin and Howarth, 1989).

Natural transitions occur between land cover types such as *Wetlands* and *Woodland.* As is the case with many thematic mapping problems, a hard boundary is placed between transitional classes that in reality may be spatially separated by a transition zone, rather than a single line of demarcation. Additionally, an area may be subject to changes in land cover with several stages occurring in the temporal transition sequence from one land cover class to another. Recognition of transitional situations by the sensor is therefore limited by the ability to define a legitimate decision boundary based upon spectral characteristics, or to temporally recognise a transitional cover class as belonging to one or other of the substantive land cover classes, and therefore reduce the mapping confusion.

For example, the spatial transition between *Wetlands* and *Woodland* needs to be considered on the basis of the spectral properties of the targets and whether they are spectrally separable. Conversely, a change from *Pine plantation* to *Residential* may not occur in a single identifiable stage. Several temporal transition stages including *Pine*

plantation, *Bare ground*, *Grassland* and engineering construction may precede the formal recognition of an area as *Residential*, having previously been categorised as *Pine plantation*. For successful interpretation, these transition stages must be uniquely identifiable or logically grouped into one or other of the land cover classes.

Other more subtle transitions may occur over a longer period of time. Not all residential areas are similar in appearance. Plates 2.8, 2.9 and 2.10 indicate the difference between new *Residential* and mature *Residential*. New residential contains a mixture of developed and undeveloped allotments, as well as bare soil and minimal vegetation. This is in contrast to more established residential areas where construction has taken place on almost all allotments and vegetation is much more dense. Depending upon garden design trends, not all established residential areas may reach the stage of relatively luxuriant vegetation which is observed in Plate 2.10.

Recognition of temporal and spatial transition sequences in the study area is an essential component for target recognition. Derivation of classification schemes and analysis techniques requires a thorough understanding of target characteristics relative to the sensor capabilities. The following section focuses on the spectral and spatial characteristics of the study area and their effects on recognition by the sensor.

2.2.4 Spectral and Spatial Characteristics

The ability of remotely sensed data to provide information regarding land covers is influenced by the spectral characteristics of targets and the spatial arrangement of target components, as well as the ability of the sensor to discriminate these factors. Sensor specifications will be discussed in Section 2.3, however it is well established that spectral resolution, which is determined by the number of spectral bands and their bandwidths, improves discrimination of targets when increased numbers of bands with narrower bandwidths are available. Additionally, sensors with more sensitive radiometric resolution also provide improved discrimination (Haack *et al*, 1987). Wharton (1987) indicates that differentiation of typical urban land covers is better based upon radiometric differences (contrast), because within class spectral properties are highly variable.

The effect of spatial resolution on interpretation performance for urban and near urban targets is less clear. Many researchers consider finer spatial resolution to be an advantage (Forster, 1980b; Gastellu-Etchegorry, 1990), because pixels contain fewer subclasses and finer resolution permits improved discrimination of within class variances. Alternatively, research results show that finer spatial resolution may contribute to poorer interpretation results due to increased within class variance, which is caused by a reduction in the smoothing effect that is present when using data with large pixel dimensions (Davis and Simonett, 1991; Haack *et al*, 1987; Martin and Howarth, 1989; Toll, 1985; Wolter *et al*, 1995).

Recognition of the spectral characteristics of targets is critical, as is an understanding of the spatial frequency of features in the scene relative to sensor spatial resolution (Marceau *et al*, 1994; Martin *et al*, 1988; Welch, 1982). The following discussion considers these factors in the context of the study area and the specific land cover classes to be identified.

At the spatial resolution of the multispectral sensors considered in this research (20 m to 79 m), the *Woodland* class comprises a relative uniform mosaic of canopy, understorey and background components that demonstrate substantial variability over distances of 3 to 5 m. This land cover class occupies relatively large areas of the scene and exhibits a relatively high spatial frequency compared to the resolution of the sensors, therefore providing a comparatively uniform response to the satellite sensors.

Within the study area interdunal wetlands vary in size from 1 ha to 10 ha, and with open water areas from 0.25 ha to 4 ha. In all cases the vegetation cover is dense and of relative uniform composition with little background soil or leaf litter reflectance. Water areas are generally small, although one wetland has an open water area of 4 ha, and within the region there are lakes up to 100 ha. Within this land cover class, the vegetation components exhibit a high spatial frequency which leads to the measurement of consistent spectral reflectance values, but due to their relatively small area and isolated locations, wetland areas are subject to considerable mixed pixel responses and edge effects, especially for Landsat MSS data.

The reflectance of *Pine plantations* measured by satellite-based remote sensors comprises predominantly interaction of electromagnetic radiation with plant components and bare soil or fallen pine needles. Where extensive thinning of the trees has occurred, some shrub and annual grass understorey also affects the response. The relative proportions of plant and background components largely depend upon the density of foliar coverage. *Pine plantations* comprise a relatively consistent mixture of components that occur at a moderate spatial frequency, and are found in relatively large areas of homogenous composition. At the spatial resolution of the available sensors, the spatial distribution of features and edge effects should not be a major factor in determining spectral response.

Grassland areas within the study area form a potentially complex spatial and temporal association with *Bare ground* areas. Spatially, *Grassland* and *Bare ground* may comprise large or small areas individually or in combination. The location and pattern may also vary with time, leading to a range of potentially complex targets for identification by the remote sensing system. While each cover type is spectrally separable, inconsistent spatial frequencies may result in variations in the ability of each sensor to identify the specific feature. Both cover types also occur as minor components in conjunction with targets such as *Residential, Commercial* and *Pine plantations*, but are considered too insignificant in a spatial sense to identify as separate targets.

From a spectral and spatial viewpoint, *Residential* and *Commercial* land cover classes are potentially the most complex targets in the area. The variety and agglomeration of target components (clay, concrete, steel and fibre cement roofing, asphalt and concrete paving, vegetation, bare soil etc), and moderate spatial frequency of objects are significant factors. Considering the spatial frequency of the components, none of the sensors are capable of identifying all individual components within the *Residential* and *Commercial* classes, and the degree of smoothing into generalised classes varies from high for Landsat MSS to moderate for SPOT HRV data. Both classes occupy relatively large areas and boundary effects should be minimal for SPOT HRV and moderate for Landsat MSS.

The *Recreation* class includes all forms of intensively managed passive and active outdoor recreational facilities, as well as areas of irrigated pasture associated with the Murdoch University School of Veterinary Science. These grassed areas provide very high spatial frequency compared to the resolution of the sensor, however they often occur in conjunction with random groups of trees and associated parking and amenity buildings, which contribute to the spectral diversity of the target. Recreation areas also occupy relatively small (<1 ha) areas and are therefore subject to considerable boundary effects during sensing, especially for coarser resolution sensors such as Landsat MSS.

2.3 Satellite Data

2.3.1 Sensor Characteristics

This research is designed to study the capability of monitoring the rural-urban fringe using remotely sensed data recorded with a range of spatial and spectral resolutions from a variety of satellite systems. Data have been acquired from four satellite systems and their specifications are summarised in Table 2.1.

	Landsat MSS	Landsat TM	SPOT HRV	IRS-1D PAN
Spectral band numbers and spectral sensitivity (μm)	1 0.50 – 0.60 2 0.60 – 0.70 3 0.70 – 0.80 4 0.80 – 1.10	1 0.45 – 0.52 2 0.52 – 0.60 3 0.63 – 0.69 4 0.76 – 0.90 5 1.55 – 1.75 6 10.40 – 12.5 7 2.08 – 2.35	1 0.50 – 0.59 2 0.61 – 0.68 3 0.79 – 0.89	1 0.50 – 0.75
Instantaneous Field of View (IFOV) (m)	79 x 79	30 x 30 band 1-5, 7 120 x 120 band 6	20 x 20	5.8 x 5.8
Quant. Levels	64	256	256	64
Altitude	919 km	705 km	832 km	~817 km
Swath width	185 km	185 km	60 km	70 km
Revisit capability	18 days	16 days	1-26 days	5 days

Table 2.1 *Sensor system characteristics (after Jensen, 1996; Saraf, 1999)*

The Landsat remote sensing program was initiated by the National Aeronautical and Space Administration (NASA) and active data acquisition commenced in 1972 with the launch of Landsat 1 as the first purpose built land resources sensing satellite. Landsat 1

was followed by the successful launch of a series of satellites during the following 12 years (Landsat 2 to Landsat 5). Landsat 6 failed during launch in 1993 and Landsat 7 was successfully launched in April 1999.

Landsat 1 through 5 all included the MSS instrument, with Landsat 4 and 5 also carrying the TM sensor. While no new satellites with MSS sensors will be launched, Landsat 7 includes the same spectral and spatial band specifications for the visible, near infrared and middle infrared bands of the TM sensor, while the thermal band has improved spatial resolution from 120 x 120 m to 60 x 60 m. In addition, a high resolution (15 m) panchromatic band has also been included (ACRES, 1999). This development program ensured the continuity of data collection by an identical sensor from 1972 when the MSS sensor was first commissioned until the demise of Landsat 5, and similarly for the TM sensor since first being launched in 1982 until the present.

The French SPOT satellite program commenced operation in 1986 with the launch of SPOT 1, and has continued with SPOT 2 and SPOT 3 being launched in 1990 and 1993, respectively. SPOT 4 is the most recent in the series having been launched in 1998. SPOT satellites have a dual capability through the HRV sensor to acquire high resolution panchromatic data or multispectral data. The first three SPOT satellites used the multispectral bands listed in Table 2.1, however, SPOT 4 now contains an additional band in the 1.58 – 1.75 μm range. The wavelength range for the panchromatic sensor has also changed from 0.51 – 0.73 μm for SPOT 1, 2 and 3, to 0.61 – 0.68 μm for SPOT 4. Revisit times may be adjusted and stereoscopic imagery acquired by pointing the sensor to adjacent orbit paths. A future SPOT 5 satellite is currently under construction and is likely to carry identical sensors to SPOT 4 to ensure continuity of the data collection program.

The ISRO has launched a range of imaging and communications satellites as part of the Indian national space program. The remote sensing program is based upon a series of satellites denoted IRS, and commenced in 1988 with the launch of IRS-1A. Three subsequent satellites have been launched with the most recent, IRS-1D, being launched in September 1997.

The multispectral imaging systems are based upon linear array sensor technology (Jensen, 1996) with spatial resolutions of 72 x 72 m, 36.25 x 36.25 m and 23 x 23 m for the LISS-I, -II and -III (Linear Imaging Self Scanning) sensors, respectively. The various models of the LISS sensor developed by the IRSO operate in the four multispectral bands almost identical to the visible and near infrared bands of the Landsat TM sensor.

The IRS-1D carries the PAN (Panchromatic) sensor with a spatial resolution of 5.8 m, which provided the first long term civilian source of high-resolution satellite data. The first such data were collected by the MOMS-2 mission flown on the Spaceshuttle in 1993. These data have a similar spatial resolution to the new generation multispectral sensors, and provide an opportunity to evaluate the geometric properties of such data.

2.3.2 Spectral and Radiometric Factors

Important factors that influence the selection of spectral bands for remote sensing include sensor design, atmospheric attenuation and energy/surface interactions. The major spectral ranges applied to earth resources sensing are between 0.4 to 12 μm and 30 to 300 mm (Richards, 1993). The 0.4 to 2.6 μm band is particularly significant for land cover sensing due to the dependence of surface reflection properties on pigmentation, moisture content and cellular structure of vegetation, and the mineral and moisture content of soil.

Figure 2.2 illustrates sample spectral reflectance curves for soil, vegetation and water in the visible and reflective infrared regions of the electromagnetic spectrum. The potential to separate targets using remotely sensed data relies upon variations in the radiometric responses of targets in spectral regions that correspond to bands available in remote sensors. Knowledge of characteristic reflectance responses of targets is essential to realise this potential.

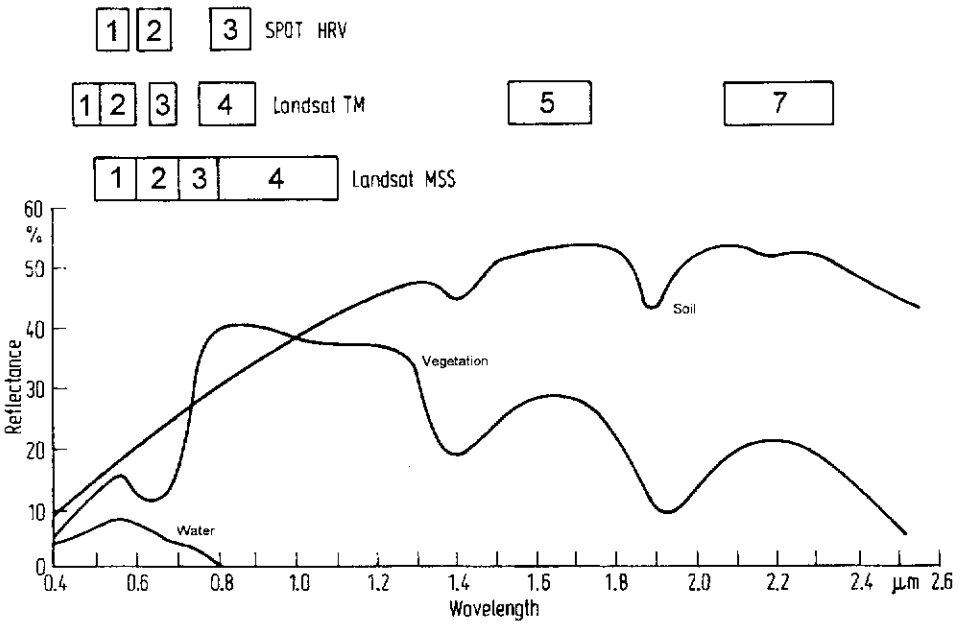


Figure 2.2 *Spectral reflectance characteristics of soils, vegetation and water in the visible and near-mid infrared (after Richards, 1993)*

The soil curve increases approximately monotonically with wavelength, and shows water absorption troughs at 1.4, 1.9 and 2.7 µm. These troughs vary in intensity according to the moisture content, and are almost absent in dry soils and sands. Additionally, clay soils exhibit characteristic absorption bands due to hydroxyl at 1.4 and 2.2 µm. Variations in these absorption bands may also be used to differentiate soil types.

Vegetation is represented by a similar characteristic reflectance curve. Water absorption bands occur at 1.4, 1.9 and 2.7 µm as for soils, while reflectance in the region from 0.7 to 1.3 µm is influenced by variations in plant cell structure. In the visible region plant pigmentation is significant with chlorophyll absorption bands dominating at 0.45 and 0.65 µm, and reflectance a maximum in the green portion of the spectrum (0.55 µm). Variations in plant species and vigour lead to variations in the water content and structure of cells and pigmentation, which results in variations in reflected radiance (Gates *et al*, 1965; Hoffer and Johannsen, 1969; Knipling, 1970). Where these variations are significant, they may be recorded by the sensor, even at

orbital altitudes. The capacity of a sensor to record these variations is determined by the presence of a suitable spectral band and the quantisation level of the sensor. Landsat MSS uses only 64 levels whereas Landsat TM and SPOT HRV utilise 256 levels, and are therefore more sensitive to variations in reflected radiation.

Superimposed on the spectral reflectance curves of Figure 2.2 are the spectral bandwidths of the Landsat MSS, Landsat TM and SPOT HRV sensors. All three sensors are primarily designed for sensing land covers containing soils and vegetation, and concentrate their spectral bands within the visible and near infrared regions. Spectral regions beyond 3 μm are less important because the level of solar energy irradiating the earth's surface is small compared to the quantity of energy in the visible and near infrared regions shown in Figure 2.2.

All Landsat and SPOT multispectral sensors possess spectral bands in the important visible and near infrared (0.7 – 0.9 μm) regions. Landsat TM provides additional bands in the visible (blue), middle infrared (reflective) and thermal portions of the spectrum designed specifically to enhance the spectral differentiability of earth surface features (Lillesand and Kiefer, 1994). This may lead to an improvement in separation of land cover targets, however the implications of the variations in the spatial resolution of sensors, especially SPOT HRV, must also be considered. The limited range of spectral bands, low quantisation level and low spatial resolution generally limit the level of detail and interpretation accuracy available from Landsat MSS in land cover mapping applications.

2.3.3 Spatial Factors

The potential effect of spatial resolution on identification of near urban land covers has been discussed in Section 2.2.4. Some researchers have achieved improved interpretation of near urban land covers due to the averaging effect of coarser spatial resolution. This improvement is due to a reduction in the within class variance of pixel response values which is caused by the inherent variability in the scene reflectance characteristics, but is suppressed with coarse spatial resolution data. It has been suggested that at high spatial resolutions (approximately 5 m) where the pixel dimension is at or below the size of common urban features such as road pavements and

house roofs, identification of individual components, based upon spectral characteristics, may be possible (Jensen *et al*, 1994). At the pixel dimension of data from the Landsat and SPOT satellites, interpretation must rely primarily upon the spectral and radiometric resolutions of the sensor, and the manner in which the available spatial resolution contributes to the effectiveness of these factors for target discrimination.

The spatial resolution of the Landsat MSS sensor represents the limit of available civilian scanning technology during the period of sensor development in the late 1960s. The Landsat TM scanner design incorporates improved spatial and spectral resolution with the objective of providing enhanced image interpretation and geometric accuracy. The Landsat TM design represents the practical limit of spatial resolution for existing optical-mechanical scanning technology. At a resolution of 20 m, the SPOT HRV system is based on a linear array sensor for data collection, and utilises similar imaging technology to that available for future high resolution satellite sensors. The designers of the SPOT HRV sensor assumed spatial resolution to be more important for interpretation than increased spectral resolution, and the system is therefore limited to three spectral bands in the visible and near infrared regions of the spectrum.

While improved spectral resolution (bandwidth) and increased quantisation levels are considered to be beneficial in defining the spectral characteristics of land covers, smaller pixel size permits sensing of targets of more pure composition and reduces the influence of mixed and boundary pixels in the interpretation process. All future satellite sensing systems designed for detailed earth surface mapping incorporate multispectral sensors with spatial resolutions of less than 10 m (Fritz, 1996). It is anticipated that at this pixel size, the potential of spectral resolution may be realised to an even greater extent for near urban targets compared to present sensing systems.

2.3.4 Temporal Factors

The purpose of this research is to investigate the potential for land cover change assessment using data from multiple satellite sensors. Jensen (1996) identifies atmospheric conditions, soil moisture and vegetation phenology cycles as the most important factors to be considered when performing land cover change detection analysis. Where dramatic differences in cloud cover or atmospheric conditions are

apparent, Jensen suggests that processes to remove atmospheric attenuation in the imagery should be applied. Data for this research were collected during periods of cloud-free and clear atmosphere, and atmospheric corrections will not be undertaken.

Most change assessment approaches compare data derived from the same sensor and therefore direct band to band comparisons are possible. Interpretation approaches that combine multidate datasets from the same sensor require application of scene normalisation algorithms in order to standardise pixel response values. In this research a multisensor approach will be applied with the consequence that each dataset will be interpreted separately for production of individual thematic classifications of land covers. Therefore any variations in pixel response caused by sensor or atmospheric differences will be accommodated because individual images will be analysed separately.

Temporal changes in land covers observed in this study between image acquisition dates are likely to be substantial due to the extended periods between sensing epochs. While these may be relatively easy to identify on aerial photographs, the main purpose of the analysis is to provide consistent interpretations from multiscale data acquired at orbital altitudes, with comparable degrees of thematic and geometric accuracy. The temporal nature of the data will therefore be considered in defining the interpretation approach, evaluating image registration processes, and developing relevant land cover change measurement and reporting methods.

2.3.5 Digital Satellite Data

The characteristics of the four satellite systems that supplied data for analysis in this study were discussed in Section 2.3.1. One scene from each satellite covering the study area will be analysed to determine the utility of the systems for change detection, and to evaluate the sources of error associated with change analysis. Details of each satellite image are shown in Table 2.2.

The data were supplied as bulk path-oriented images that had been corrected for standard radiometric distortions such as non-uniform detector response and basic geometric parameters. Image geometry had been corrected in the cross-track direction only, and included corrections for:

(i) Earth rotation,

(ii) Earth curvature distortion,

(iii) Panoramic distortion (where applicable for off-nadir pointing angle),

(iv) Satellite sensor geometry,

(v) Satellite position, velocity and attitude variations.

	Landsat MSS	Landsat TM	SPOT HRV	IRS-1D PAN
Image bands	1, 2, 3, 4	1, 2, 3, 4, 5, 7	1, 2, 3	Panchromatic
Image date	Dec 19, 1972	Oct 28, 1986	Dec 3, 1991	Feb 11, 1998
Row	082	082	414	102
Path	120	112	316	149
No. of columns	1 501	1 501	3 000	4 320
No. of rows	1 201	1 001	2 001	4 606

Table 2.2 *Satellite data utilised in this research*

False colour composite images of the Landsat MSS, Landsat TM and SPOT HRV multispectral data for the study area are shown in Figures 2.3, 2.4 and 2.5, respectively. For consistency the images are formed using the corresponding green, red and near infrared bands for each sensor. The images have been rectified to the AMG and resampled to a common pixel size of 20 m to facilitate comparisons for change analysis. Superficial examination of each image indicates the relative spatial resolution available from each sensor, however the differences in spectral resolution are not obvious from these images. The reduced radiometric resolution available from Landsat MSS data is also not immediately obvious.

2.3.6 Image Processing

All image processing of satellite imagery in this research was performed using IDRISI image processing software. IDRISI for Windows Version 2 is an integrated image processing and raster GIS package developed by the Graduate School of Geography at Clark University, USA. IDRISI operates within a Windows 95 computing environment.

IDRISI has been primarily developed as an educational package, although it is utilised in a range of government and non-government organisations for commercial purposes.

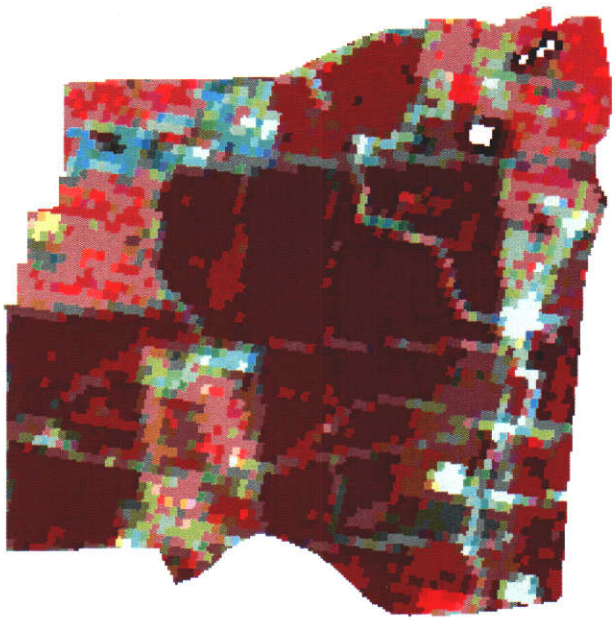


Figure 2.3 *False colour composite of the study area compiled from Landsat MSS bands 1 ,2 and 3 (1972) (Scale 1:50 000)*

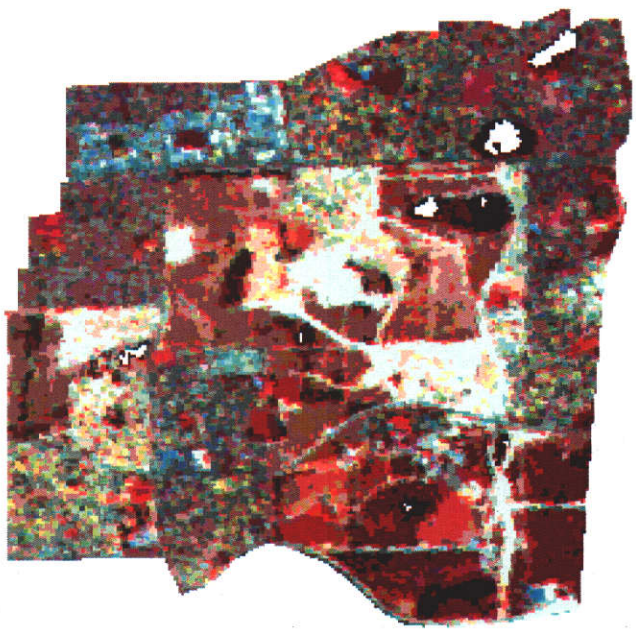


Figure 2.4 *False colour composite of the study area compiled from Landsat TM bands 2, 3 and 4 (1986) (Scale 1:50 000)*

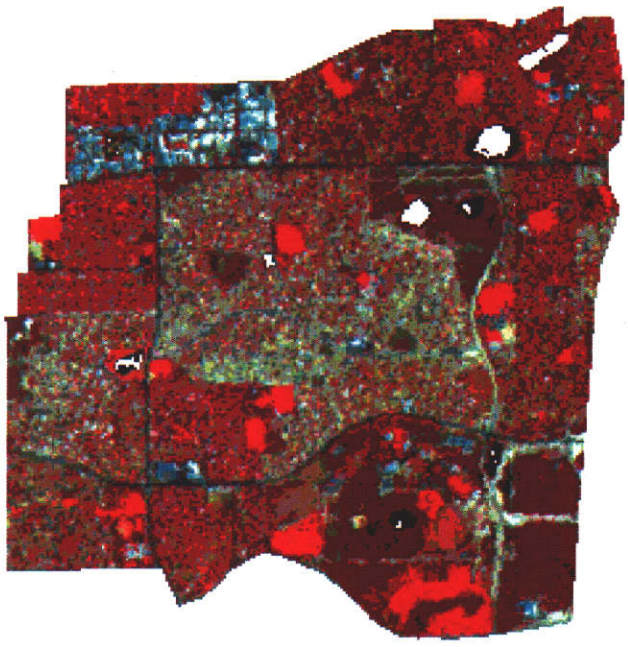


Figure 2.5 *False colour composite of the study area compiled from SPOT HRV bands 1, 2 and 3 (1991) (Scale 1:50 000)*

As a training and development package it has demonstrated significant innovative functionality in both image processing and GIS, which commercial vendors have subsequently implemented in their own products.

For this research IDRISI provides an appropriate processing platform for image analysis and change detection. The basic image processing components of image registration, multispectral classification and overlay analysis are fundamental components required for change detection. The ability to import a detailed reference map of land cover classes and to rasterise the map for comparison with interpreted satellite images is also important.

Additional evaluation concerning the spatial distribution of the data will be performed using spatial autocorrelation analysis. This module forms an integral part of the IDRISI system and is essential to this research.

IDRISI is not ideally suited to the commercial image processing environment or for large scale image analysis due mainly to the linear approach to processing and the

quality of its interface with other image processing and GIS packages. A data management overhead is introduced in IDRISI whereby each processing function produces an output data file, which then becomes the input file for the next process. In comparison, most commercial systems permit definition of a user defined algorithm, which permits several functions to be performed and only the final image is saved to disk.

While dynamic data access links to other packages are not maintained by IDRISI, the IDRISI file formats are relatively simple and may be translated to most raster-based systems. The import and manipulation of vector information is rather more problematic and limited functionality exists.

2.4 Reference Data

Reference data for the study area are required for derivation of training sites for multispectral classification and for verification of land covers derived from the classification process. Satellite image acquisition for this study occurred over the period 1972 to 1998, and aerial photographs form the only detailed and permanent record of land cover conditions against which the analysis can be compared.

2.4.1 Aerial Photography and Planimetric Maps

Aerial photography relevant to each acquisition date of the multispectral images utilised in this research were purchased from DOLA, Perth. DOLA generally acquires photography of the Perth metropolitan area annually, and photography nearest to the date of satellite overpass (Table 2.3) formed the basis of the detailed land cover map used as reference data.

A time lapse of up to six months may occur between aerial and satellite image acquisition, therefore two sets of photographs for each satellite overpass, for which thematic reference data are needed, were acquired. One set before and one set after the satellite overpass were compared in order to determine if any significant changes in land cover had occurred, between sensing by the satellite and reference data collection via aerial photography.

Date	Scale	Emulsion	Run (Photo Nos)
May 2, 1972	1:25 000	Panchromatic	12 (5197, 5198) 13 (5161,5162)
June 26, 1973	1:25 000	Panchromatic	12 (5278-5280) 13 (5306-5308)
April 19, 1986	1:20 000	Panchromatic	8 (5023, 5024) 9 (5087, 5088)
Dec 20, 1986	1:40 000	Panchromatic	2 (5198-5200)
Jan 5, 1991	1:20 000	Colour	8 (5061-5063) 9 (5126-5128)
Jan 4, 1992	1:20 000	Colour	8 (5038,5039) 9 (5213, 5214)

Table 2.3 *Details of aerial photography acquired for reference data compilation*

Changes in land cover between sensing epochs are considerable and are representative of transitions commonly observed in near urban areas. Digital mosaics from each epoch of aerial photography have been constructed to indicate the degree of change occurring in the area. Figures 2.6, 2.7 and 2.8 demonstrate the magnitude of changes that have taken place in the study area over the period 1972 to 1991.

Planimetric maps were required for image rectification and as a base for reference data compilation. Ground Control Points (GCPs) were identified and coordinates for rectification of the Landsat MSS data were extracted from the Perth and Fremantle 1:100 000 topographic map sheets which were compiled from 1976 aerial photgraphy. GCPs for rectification of the Landsat TM and SPOT HRV data were extracted from the DOLA *StreetExpress* Digital Street Directory of Perth. This format provides the most detailed planimetric information available for identification of suitable GCPs for rectification of these data.

Rectification of the high resolution IRS1-D satellite data requires determination of GCP coordinates within two to three metres. Suitable planimetric data from conventional mapping programs were not available. Survey data for the road centreline network of the greater Perth metropolitan area were available from MRWA, and proved suitable for

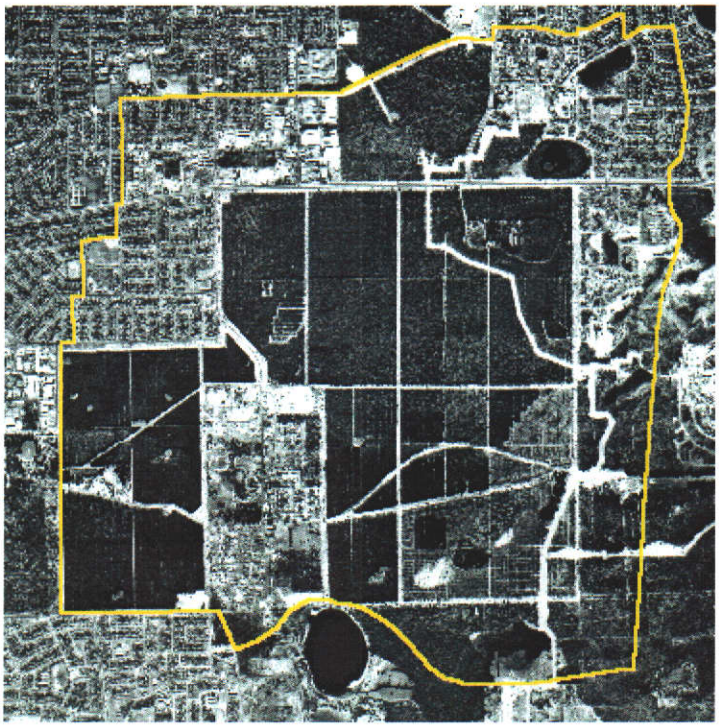


Figure 2.6 *Aerial mosaic showing the study area in May, 1972 (Scale 1:50 000)*

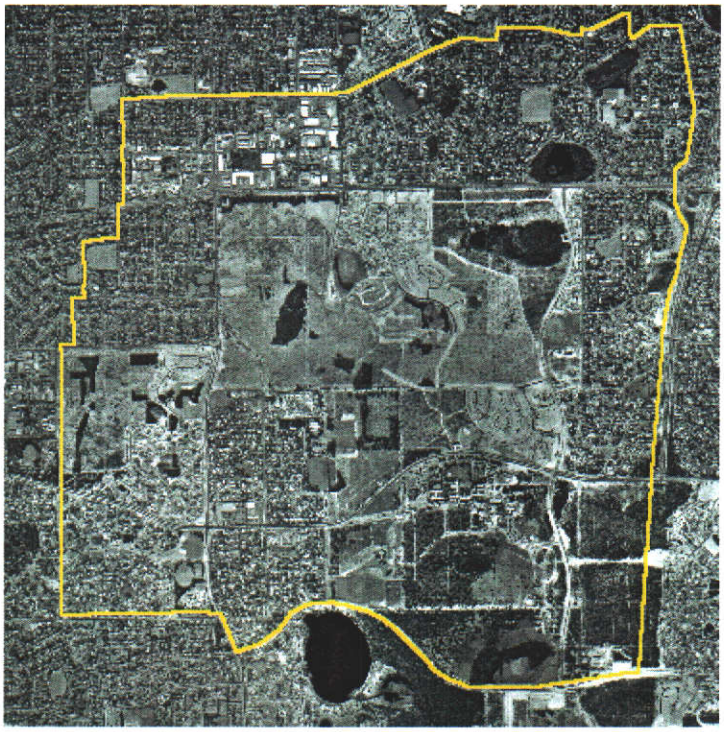


Figure 2.7 *Aerial mosaic showing the study area in April, 1986 (Scale 1:50 000)*

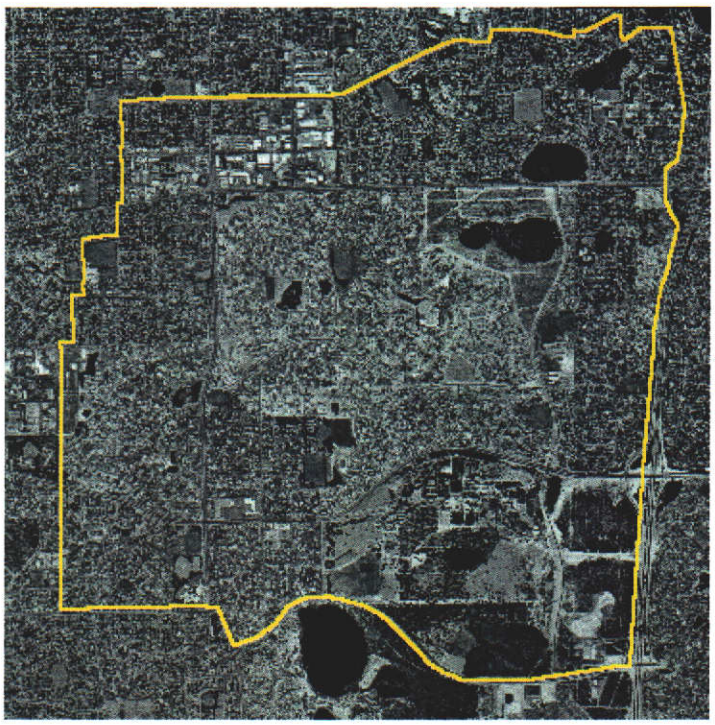


Figure 2.8 *Aerial mosaic showing the study area in January, 1992 (Scale 1:50 000)*

GCP extraction. These data were derived from a range of sources including kinematic Global Positioning System (GPS), photogrammetric and terrestrial surveys.

A series of 1:5 000 scale orthophotomaps produced by DOLA were utilised as a planimetric base for reference data compilation. These orthophotomaps maps cover the map series quadrant Perth BG 34 and include maps 04.05E, 04.06E, 05.05W, 05.06W, 05.05E and 05.06E. These large scale orthophotomaps provide a high level of detail for transfer of land cover boundaries from the aerial photography. While the satellite data were recorded over a 19 year time period, the aerial photographs used for the orthophotomap were acquired in 1984 resulting in sufficient common detail occurring in each set of aerial photography for the accurate transfer of land cover information.

2.4.2 Identification of Land Cover Types

A description of the land cover types present in the study area has been provided in Section 2.2.2. These same classes were interpreted from stereoscopic aerial photography using conventional photointerpretation techniques. Classes were identified strictly according to the observed spectral and radiometric characteristics of the land

covers, which are the same interpretive components used to identify the features during multispectral classification (Martin and Howarth, 1989). A tendency during manual photointerpretation is to utilise all facets of the image (Lillesand and Kiefer, 1994), and be distracted by land use rather than land cover classes. While consideration of all facets of the image assists in obtaining accurate photointerpretation, identified classes must be consistent between the primary satellite data interpretation and reference-data collection.

More detailed interpretation of land covers was possible from the aerial photographs than from the satellite image data, therefore the reference data record a level of detail at least equivalent to that available from the finest resolution multispectral satellite (SPOT HRV) data. For interpretations from Landsat MSS and TM data and for change assessment, aggregations of land cover classes interpreted from the aerial photographs were formed as required.

Land cover classes were identified on the stereoscopic aerial photography using a mirror stereoscope and magnifying binoculars. Class boundaries were identified on the aerial photography within 0.5 mm on the photograph, which represents $\pm$10 m in ground units. This is equivalent to 0.5 pixel at the resampled resolution of the image data, and corresponds to 0.125 of an MSS pixel, 0.3 of a TM pixel and 0.5 of a SPOT pixel. All values are within the resampling precision of the corresponding satellite images. IRS data will only be used for assessment of the geometric properties of the data, for evaluation of the rectification of high resolution satellite data. Reference data precision is therefore not relevant to this dataset.

Aerial photographs taken before and after each sensing epoch were compared to assess the reliability of land covers in the reference data. Variations may occur because the sensing date could be up to six months from the date of acquisition of aerial photography. Any variations were noted and considered in the context of the interpretation results achieved from the satellite data.

2.4.3 Reference Map Production

Reference maps of the study area were created to provide a reliable source of information for comparison with interpretations derived from the satellite information.

Reference maps were produced for each sensing epoch, with each epoch represented at Level I and Level II of the hierarchical classification scheme (see Chapter 4). The initial level of land cover interpretation represents a higher level of interpretation than is achievable from the corresponding satellite data, and classes have been aggregated in order to permit meaningful comparison of interpretations derived from the three sources of satellite information.

Compilation of the reference data comprised a number of distinct phases from interpretation of the aerial photography through to raster representation for comparison with the satellited-derived land cover results:

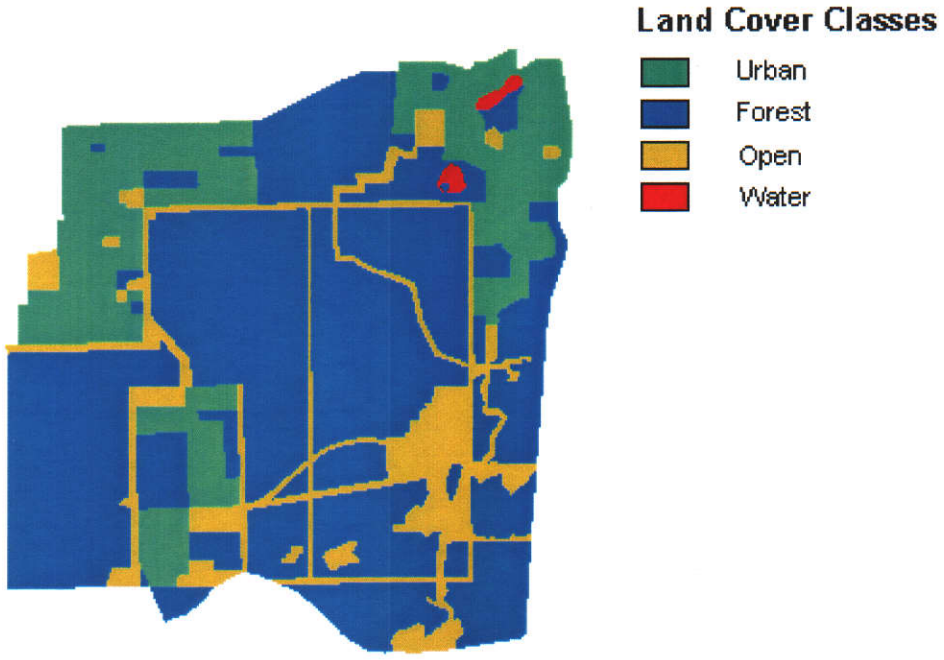
(i) *Interpretation of aerial photographs.* Overlapping aerial photographs of the study area were interpreted under a mirror stereoscope. Conventional photointerpretation techniques were employed taking care to identify land cover classes which were also relevant to interpretation of satellite data rather than land use classes which may not be identifiable in the imagery. Details of land cover classes were transferred to a 1: 5 000 orthophotmap.

(ii) *Digitising.* The completed hard copy reference map was digitised using MicroStation and exported in DXF exchange format.

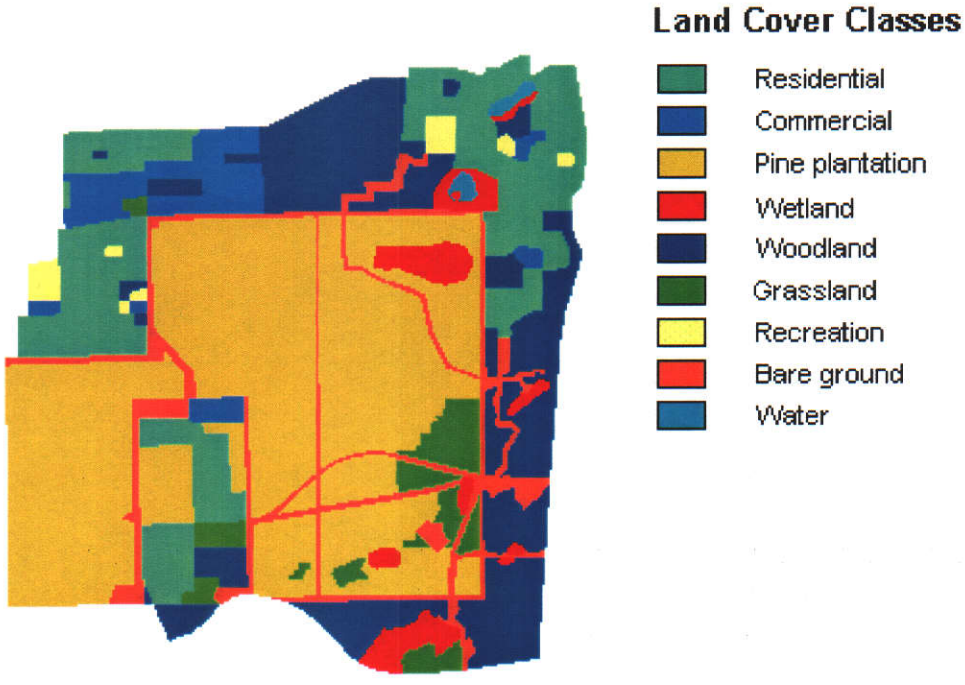
(iii) *Polygonisation.* The land cover boundary vectors were imported to IDRISI and manipulated in the vector environment to build topology and form logical polygons representing land cover classes.

(iv) *Vector to raster conversion.* Land cover polygons were transformed from vector to raster form in IDRISI to derive a raster land cover map geographically coincident with the interpreted satellite data. The pixel resolution of the land cover map was 20 m and coincides with the resolution of the resampled satellite data.

Figure 2.9 contains reference data at Level I and Level II interpreted from aerial photography acquired in 1972 to coincide with the Landsat MSS data. Four classes are represented at Level I and nine classes at Level II. Level I classes are derived from an aggregation of the Level II classes according to the hierarchical classification scheme described in Chapter 4. Figure 2.10 shows reference data for 1986 interpreted to coincide with the Landsat TM data, and Figure 2.11 represents similar reference data for 1992 that coincides with the SPOT HRV data.

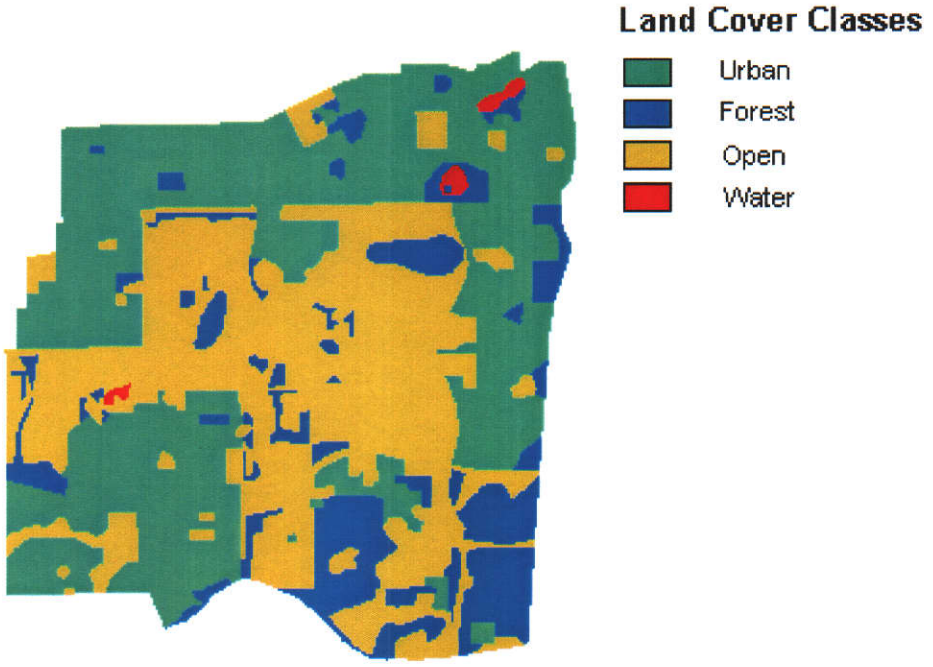


(a) Level I reference data map

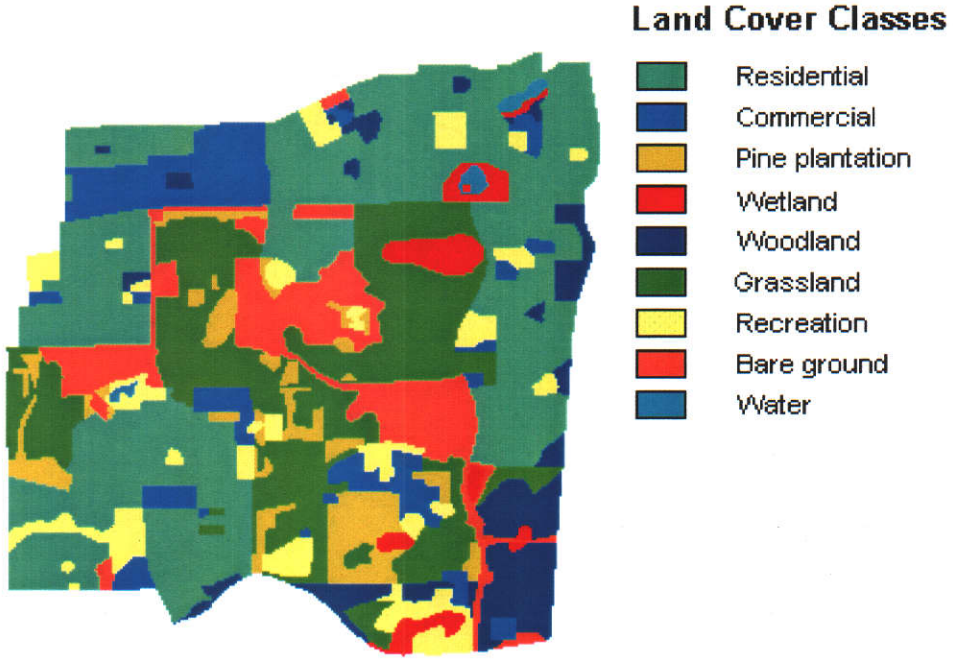


(b) Level II reference data map

Figure 2.9 *Reference data derived from interpretation of aerial photographs (1972) (Scale 1:50 000)*

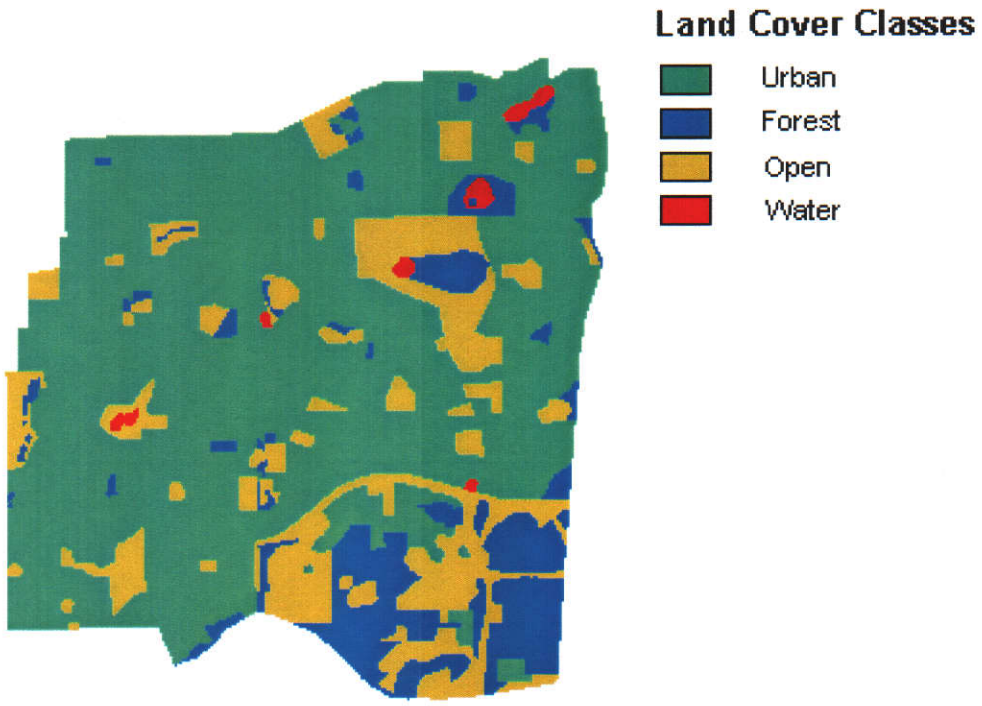


(a) Level I reference data map

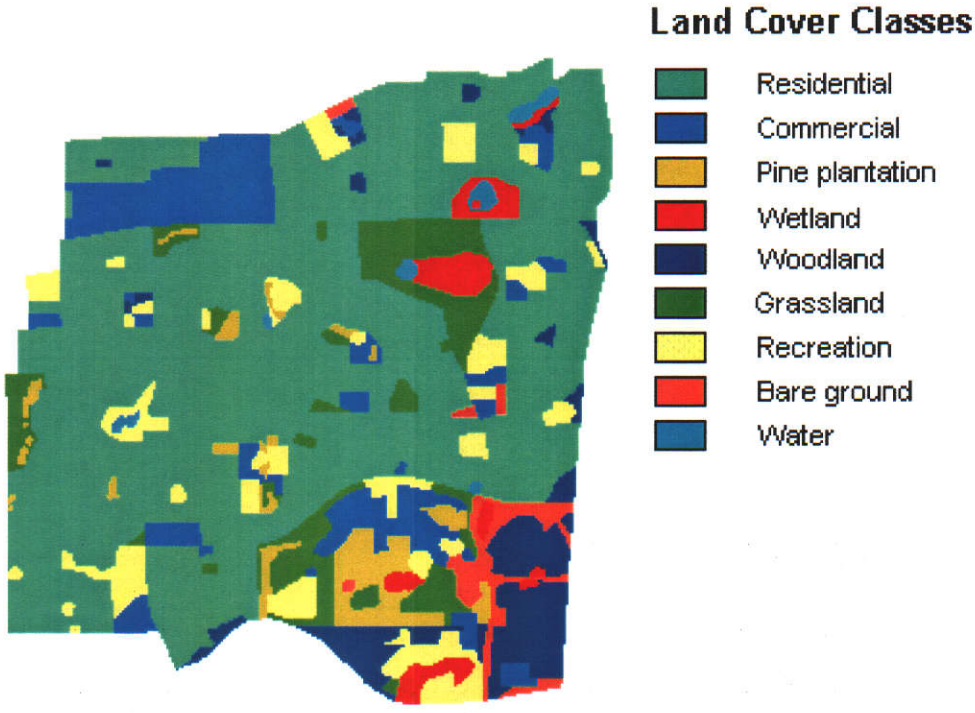


(b) Level II reference data map

Figure 2.10 *Reference data derived from interpretation of aerial photographs (1986) (Scale 1:50 000)*



(a) Level I reference data map



(b) Level II reference data map

Figure 2.11 *Reference data derived from interpretation of aerial photographs (1992) (Scale 1:50 000)*

Development of a full raster reference map was undertaken to enable direct comparison of the satellite-based land cover interpretations. This includes the comparison of boundary locations as well as for thematic classifications of the polygon areas.

2.5 Summary

The selection of a specific study area must be made in consideration of the overall objectives of the research. Where the objective is to evaluate that multiscale remotely sensed data can be used to accurately determine land cover changes in the rural-urban fringe, an area that demonstrates characteristic changes is required. The Murdoch/Winthrop study area in the Perth metropolitan region was selected because during the period under consideration the land covers changed from a mixture of urban and predominantly rural and non-urban land covers, to mostly urban land covers.

In the study area, rural and non-urban land covers range from wetlands and woodlands through to pine plantations and open grasslands, while urban land covers include residential, commercial and recreational areas. Importantly, the area is not only characterised by some static land covers, but several transition sequences between otherwise discrete classes were also recognised, which may influence the outcome of any change analysis.

As the availability of high-resolution commercial multispectral satellite data increases, the capability to monitor land cover status improves. However, the longevity of these data acquisition programs is indeterminate and the future of any one program cannot be guaranteed. A history of data is not available on which to base long term monitoring using these new generation satellite systems, and integration of disparate datasets from alternative sensors may be required for any long term monitoring program. While high-resolution multispectral satellite data were not available at that time, multiple resolution Landsat MSS, Landsat TM and SPOT HRV data were analysed in this research to test the utility of multiscale remotely sensed data. Similarly, high-resolution panchromatic IRS-1D data were investigated to evaluate the geometric aspects of potential future multiscale data integration.

Geometric rectification of the satellite data is a fundamental requirement for change analysis. Analogue and digital planimetric data derived from a combination of

photogrammetric and GPS sources were important in providing relevant control information for rectification of the satellite data. Data with a planimetric precision ranging from $\pm$3 m to $\pm$50 m were required to rectify imagery which ranges in spatial resolution from 5.8 x 5.8 m to 79 x 56 m.

Verification of change for evaluation of the utility of the remote sensing systems requires reliable reference data relevant to the interpretation of land covers. Reference data were sourced from medium scale aerial photography and interpreted to provide comparisons of thematic and positional comparisons with the primary satellite data. Conversion of the reference map to digital form facilitated this evaluation.

Chapter 3

IMAGE RECTIFICATION AND RESAMPLING

Image rectification and resampling are essential precursors to change detection analysis using remotely sensed data. Image rectification provides a robust planimetric base for comparison of data, but must be appropriate to the geometric characteristics of the satellite imaging system. The quality and selection of GCPs are fundamental to the performance of the rectification process. Image resampling is critical to the maintenance of the spectral and spatial quality of the data. The quality of the rectification and resampling processes underpin subsequent analysis for the measurement of change within the study area.

3.1 Introduction

Analysis of multiple sources of remotely sensed data is dependent upon the ability to accurately relate corresponding locations in each image through a system of spatial referencing. The basic requirement is for a geometric match to be established between data sources and for relevant brightness values to be transferred according to the derived geometric relationships. Where data sources comprise multiple images, the geometric process is termed registration or, alternatively, the process for combining an image and analogue or digital planimetric data utilising a geographic or geodetic coordinate system is termed rectification or geocoding (Ehlers, 1997). Transference of individual pixel brightness values to the registered or rectified image is called image resampling.

Uncertainties associated with information extracted from remotely sensed data are considered to fall into two categories: thematic and positional (Chrisman, 1991). While the spectral, spatial and radiometric characteristics of the data largely determine the thematic uncertainty of the land cover information, positional uncertainties are influenced by the rectification and resampling approach, as well as the geometric properties and spatial resolution of the sensor.

The requirement for rectification arises from the need to integrate information from a range of sensors, analyse changes using remotely sensed data from different epochs, locate points of interest, or to integrate remotely sensed data with GIS (Buiten and van Putten, 1997; Fonseca and Manjunath, 1996; Kardoulas *et al*, 1996). Refinements in rectification accuracy also have potential for improving results achieved from some standard information extraction approaches (Wolter *et al*, 1995).

The influence of rectification accuracy on change detection is well documented. Townshend *et al* (1992) indicate an error of 10 percent in the calculated change in Normalised Difference Vegetation Index (NDVI) between epochs may be caused by only a 0.2 pixel rectification error. In a study by Aspinall and Hill (1997), 20 percent of all changes that were observed between two land cover maps were identified to be due to geometric limitations, and Martin (1989) indicates that displacements between images of only 0.5 pixel can introduce unacceptable levels of change error. A focus on

rectification methods and quality is therefore significant in the development of appropriate change detection approaches.

The combined processes of rectification and resampling comprise a number of integrated components (Fonseca and Manjunath, 1996; Welch *et al*, 1985):

(i) *Feature identification.* Identifies a set of relevant features in the image for which ground coordinates may be derived;

(ii) *Feature matching.* Establishes correspondence between the image and ground points and derived coordinate values in both systems;

(iii) *Spatial transformation.* Determines the transformation parameters that can match the image with the map projection using the identified control points using a least squares solution;

(iv) *Spatial interpolation.* Transforms the image to the map projection using the computed mapping function parameters;

(v) *Intensity resampling.* Computes new digital values for each rectified pixel from the original image.

Multiple image data sets have been acquired at a variety of spatial resolutions for this research. Consideration to each of these components will be given in order to derive the most appropriate approach for data integration.

These data will be compared for the purpose of change detection, therefore rectification and resampling represents a critical phase of data preparation for transformation to a common planimetric base. Image resampling will also be employed to derive a common pixel resolution and to allocate appropriate digital values to each geocoded pixel. Information extracted from the remotely sensed data will be compared to reference data for thematic accuracy assessment, therefore rectification to a common ground coordinate system (AMG) will be undertaken.

Rignot *et al* (1991) indicate most rectification algorithms have been developed for passive sensors operating in the visible and near infrared regions of the spectrum, and that for integration of multiresolution data with disparate geometric characteristics, a wide variety of alternative techniques should be considered. Implementation of

automated rectification techniques have been recommended by several authors (Djamdji *et al*, 1993; Ehlers, 1997; Fonseca and Manjunath, 1996; Igbokwe, 1999), however they mainly emphasise image matching for control point selection rather than the automated performance of the rectification algorithm.

All data utilised in this research are acquired in the visible and near infrared regions of the spectrum and have reasonably consistent geometry due to acquisition by near-nadir pointing sensors. The data range in spatial resolutions from 5.8 m to 79 m, and alternative rectification techniques for optimisation of the planimetric qualities of the data will be examined. This is especially relevant due to the development of high resolution multispectral remote sensing satellites and the increasing requirement for integration of image data from multiple spatial resolutions.

3.2 Image Rectification Approaches

Image rectification methods may be categorised as either parametric or non-parametric. Parametric methods are designed to model the nature and magnitude of distortions inherent within the image and to devise specific correction formulae. Non-parametric approaches rely upon establishing an analytical relationship between the image pixels and the corresponding coordinates on the ground (Richards, 1993).

Parametric approaches rely upon knowledge of systematic errors in the satellite orbit and sensing system. This approach is unable to deal with errors such as those introduced by changes in the sensor attitude, which either cannot be estimated or it is not possible to measure them with sufficient precision to be applied to the data (Jensen, 1996).

The selection of an appropriate rectification technique is somewhat spatial resolution dependent, where the objective is to achieve residual rectification errors of less than 0.5 pixel (Labovitz and Marvin, 1986). Data from aircraft and satellite platforms are subject to similar geometric distortions, however their magnitude and significance varies (Ehlers, 1997), which affects the selection of the rectification technique. Satellite sensors are affected by the systematic effects of earth rotation and earth curvature, but are generally very stable in their altitude, attitude and velocity. As most current satellite

sensors are also of moderate spatial resolution, errors caused by altitude, attitude and velocity variations are relatively small and can be accommodated with non-parametric rectification approaches. For Landsat and SPOT satellite data most systematic errors, including defects associated with sensor design and operation, are corrected during preprocessing (Pala and Pons, 1995).

Where high spatial resolution satellite data, such as those available from the IRS1-D PAN sensor are processed, non-parametric approaches may not provide sufficient accuracy of rectification (Novak, 1992). Parametric rectification may be required, which will account for earth curvature and rotation in addition to variations in satellite altitude, attitude and velocity. It is these latter errors which are considered for correction of airborne remote sensing devices (Richards, 1993).

Fonseca and Manjunath (1996) consider the extraction of control points from remotely sensed imagery to be the most difficult component of image rectification, however they also indicate that multisensor image rectification is a difficult problem and it is unreasonable to expect a single algorithm to perform satisfactorily for all datasets.

3.2.1 Non-parametric Rectification

Non-parametric rectification depends upon the establishment of a mathematical model that relates pixel locations in remote sensing images and the coordinates of the same points on the ground. These models do not rely upon any physical relationship between the image and ground coordinates, and are suitable for correction of image geometry irrespective of the nature of the distortion (Ehlers, 1997).

Several non-parametric mathematical models have been adapted to image rectification. The simple polynomial trend model provides a global interpolation over the image and is most common in image processing systems. The polynomial approach is capable of modelling any errors (Novak, 1992), however its success depends upon control points of adequate precision and image data of consistent geometry. This applies where the sensor platform generally has a stable attitude and is subject only to flight conditions with long periodic fluctuations, as are generally found with spaceborne sensors (Buiten and van Putten, 1997).

Where the platform is not so stable, as may be found in aircraft scanners, the in-flight deviations are too severe or occur with too high frequency, and cannot be modelled with global polynomials. Ehlers (1997) indicates that a piecewise polynomial approach may be successful in these cases where the image is broken into a number of subimages and each is modelled independently. Special attention is required to accommodate discontinuities at joins between subimages, and generally first order polynomials are used to minimise complications in extrapolated parts of the individual images.

The general processing method is similar for both the global and piecewise approaches, except that the piecewise technique requires a set of GCPs for each subimage. El-Manadili and Novak (1996) indicate the polynomial approach is attractive because of its simple implementation and is independent of any satellite or sensor calibration parameters. Limitations are evident due to the requirement for extensive ground control, the lack of a physical interpretation model, the need to derive a new model for each image or subimage and strong edge effects outside the envelope of GCPs..

The general polynomial approach (global and piecewise) assumes there exists a function F, so that pixel coordinates (x, y) can be mapped to the corresponding ground coordinates (X, Y):

$$F(x, y) = (X, Y) \tag{3.1}$$

The mapping function is generally in the form of a pair of simple polynomials of the first, second or third degree (Richards, 1993) of the general form:

$$X = \sum_{i=0}^{n} \sum_{j=0}^{i} c_{ij}\, x^{i-j}\, y^{j} \tag{3.2}$$

$$Y = \sum_{i=0}^{n} \sum_{j=0}^{i} d_{ij}\, x^{i-j}\, y^{j} \tag{3.3}$$

where

c_{ij} *and* d_{ij} *are transformation coefficients*

A polynomial of the second order is of the form:

$$X = c_{00} + c_{10}x + c_{11}y + c_{20}x^2 + c_{21}xy + c_{22}y^2 \quad (3.4)$$

$$Y = d_{00} + d_{10}x + d_{11}y + d_{20}x^2 + d_{21}xy + d_{22}y^2 \quad (3.5)$$

A first order polynomial comprises only the first three terms of Equations 3.4 and 3.5. Values for the transformation coefficients are determined from a set of GCPs that can be identified in the image and on the ground. First, second and third order polynomials require three, six and ten control points respectively, for derivation of the transformation parameters. Generally at least twice this number of GCPs is utilised and least squares analysis is applied to compute the values.

The GCPs must be well defined and unambiguous so that the transformation coefficients are reliably determined and the coordinate residuals minimised. The basic hypothesis of the rectification assumes (Buiten and van Putten, 1997):

(i) Neither local discontinuities nor short-periodic fluctuations are present in the image;

(ii) The degree of the polynomial model has been selected appropriately;

(iii) There are no gross errors in the determination of the control points.

Issues concerning the selection of control points and analysis of residuals to eliminate gross errors in the determination of control points and polynomials models will be addressed in Section 3.4.

The order of the polynomial to be applied for the purpose of image rectification is based upon an assumption of the nature of errors inherent in the data. Novak (1992) indicates the polynomial method is independent of sensor geometry, but is influenced by sensor fluctuations as indicated above, the extent of the image and the amount of relief variation in the terrain. As the polynomials use a global model they are unable to account for these localised effects, especially relief displacement (Fonseca and Manjunath, 1996).

Relief displacement has the effect of causing horizontal displacement in the cross-track direction, and is more significant the further a pixel is from the nadir. Where large variations in terrain exist, a higher (third) order polynomial will reduce but not eliminate the effects of relief displacement. Selection of GCPs located at approximately the mean terrain height minimises the effects of relief displacement, as the maximum deviation in relief from the datum (mean terrain height) for any point will be minimised (Welch and Usery, 1984).

Most authors indicate that polynomials of the first or second order are appropriate for rectification of medium resolution satellite data such as Landsat and SPOT images (Jensen, 1996; Pala and Pons, 1995; Welch *et al*, 1985). The effect of higher order polynomials and an increased number of GCPs may lead to the polynomial providing a closer fit to the control, however for independent test points the root mean square error (RMSE) may increase, especially at large distances from control points. Welch and Usery (1984) found that for Landsat MSS data a complete scene rectified using 40 GCPs produced the lowest RMSE with a third order polynomial. A subscene of 256 x 256 pixels was rectified optimally with a first order polynomial using as few as five GCPs.

Similar analyses by the same authors using full scene Landsat TM data provided optimum results with second order polynomials. For subscenes up to 1024 x 1024 pixels, first order polynomials provided the optimum rectification using five to ten GCPs. This improvement in the Landsat TM rectification compared to the Landsat MSS data occurred because of the enhanced satellite attitude control available on Landsat 4 and subsequent satellites in the series. Welch and Usery (1984) conclude that following correction for systematic errors in the data, most residual errors are mainly due to scaling and translation components, and can be adequately accommodated with first or second order polynomials.

While the same level of detail is not available for analysis of the rectification of SPOT HRV data, many authors report results of high quality from global polynomial rectification using first order polynomials (Cook and Pinder, 1996; Eckhardt *et al*, 1990;

Hill and Aifadopoulou, 1990; Treitz *et al*, 1992). These results are attributable to the high geometric quality of SPOT HRV data acquired over level terrain.

Results of the polynomial regression for data used in this research are shown in Table 3.1 and indicate the RMSE values computed for first, second and third order polynomials for each of the data sources. All RMSE values for the Landsat MSS, Landsat TM and SPOT HRV data are below the 0.5 pixel threshold indicated as acceptable for rectification of image subsets (Labovitz and Marvin, 1986). For the IRS1-D image the RMSE for the first order polynomial is considerably larger than the accepted threshold, however both the second and third order results are less than 0.5 pixel.

	MSS	**TM**	**SPOT**	**IRS1-D**
1st Order	0.417	0.402	0.345	2.479
2nd Order	0.397	0.391	0.335	0.493
3rd Order	0.303	0.380	0.310	0.460
No control pts	14	25	26	40
Rows	641	1 001	1 401	4 606
Columns	671	1 501	2 301	4 320

Table 3.1 *RMSE values for first, second and third order polynomial rectification of the satellite data (RMSE in pixels)*

Computation of a third order polynomial for the Landsat MSS data results in a much lower RMSE compared to the first and second order polynomials. However, the third order polynomial was derived with only 14 GCPs whereas the minimum GCPs required are 10, and it is generally recommended that at least 20 points be utilised for reliable determination of the coefficients. Therefore, a low RMSE may be due more to a small number of GCPs than a close fit to control. It is therefore not considered highly reliable and will not be used in further analysis. Lack of detailed reference data for use with the Landsat MSS image precluded selection of a larger number of GCPs. The area involved is only 641 x 671 pixels (1/16 scene), over which higher order polynomials are not normally considered necessary (Welch and Usery, 1984). Comparison of the first and second order polynomial regression results shows only minimal variations between the

RMSE values and, consequently, a first order rectification will be applied during this research.

Computation of second and third order polynomials for the SPOT and Landsat TM data provides incremental improvement in the results compared to the first order polynomial. In each case at least 25 GCPs were used for derivation of the polynomials and are therefore considered to provide reliable estimates of the polynomial coefficients. For both the Landsat TM and SPOT data the change in RMSE when moving from a first to a third order polynomial is approximately the same as the change in the Landsat MSS image between the implementation of a first and second order polynomial. Implementation of other than a first order polynomial for the Landsat TM and SPOT HRV data is therefore not considered to offer substantially improved results.

The RMSE for Landsat MSS data decreased from 32.9 m to 31.4 m when a second order rather than a first order polynomial was applied, and for Landsat TM from 12.0 m to 11.4 m when a third order compared to a first order polynomial was used. Results for SPOT HRV data with the implementation of a third order as compared to a first order polynomial decreased the RMSE from 6.9 m to 6.2 m. Considering the likely errors involved in control point selection and determination, these differences are not considered to be significant.

Implementation of first order polynomial rectification for the Landsat MSS, Landsat TM and SPOT HRV data is based upon the results of this investigation and knowledge of the internal geometric consistency of the Landsat data (Welch, 1985; Welch and Usery, 1984). Similar evaluations of the IRS1-D data are not available and results from Table 3.1 indicate that first order rectification is not appropriate. A total of 105 GCPs were selected for rectification of the IRS1-D data and after elimination of outliers 70 points were accepted, with 40 points utilised for polynomial determination and the remainder for check point evaluation. Using a first order polynomial, a stable solution could not be determined and improvement of the RMSE using a first order rectification was not possible.

Implementation of a second or third order polynomial provided a stable solution and improved the RMSE to an acceptable level of less than 0.5 pixel. The change in RMSE from a second to third order model was greater than any incremental change observed in the other datasets, and without confirmation of the geometric integrity of the IRS1-D data, a third order polynomial was selected for rectification to provide the optimum solution possible.

The specific area under investigation in this research is less than 5 km x 5 km, which has been subset from a larger image of approximately 40 km x 45 km. According to Welch and Usery (1984), rectification of medium resolution satellite data may be accommodated using first order polynomial rectification and, while for an image area of this size second order rectification should be considered, it does not offer any advantage in this case.

Buiten and van Putten (1997) indicate that whenever a first order polynomial does not meet the specifications for Landsat TM or SPOT data, the reason for the lack of fit is most likely due to errors in GCP measurement or weaknesses in the GCP distribution. While the third order rectification of the Landsat MSS data has been disregarded, all other results using first to third order polynomials meet the generally accepted guideline with RMSE values less than 0.5 pixel. Higher order polynomial rectification did not appear to offer superior results, therefore all further analyses of the Landsat MSS, Landsat TM and SPOT images will be undertaken using data rectified with a first order polynomial. Results for the IRS1-D data were less consistent and indicate that with potentially lower geometric consistency of higher resolution data, higher order polynomials or even parametric rectification approaches may be required. However, quality results have been achieved from these data using a conventional second or third order polynomial solution.

Evaluation of the RMSE of GCPs has been used to evaluate the consistency of polynomial modelling procedures for image rectification. Values of independent check points will be compared to the results of image rectification undertaken through polynomial modelling to evaluate the geometric quality of the rectification process. A comprehensive analysis of these results is discussed in Section 3.5.

3.2.2 Parametric Rectification

The processes of parametric or deterministic rectification are based upon a thorough knowledge of the physical characteristics and magnitude of the errors and distortions inherent in the remotely sensed data. Sources of error comprise elements of the sensing system (sensor aspect ratio and non-linear scan rate), target (earth rotation and earth curvature) and platform parameters (position, velocity, attitude and altitude). Current satellite systems cannot provide the detailed data required for implementation of this deterministic approach, however with the inclusion of GPS navigation systems accurate flight path determination is possible, and interpolation of other flight parameters will assist this process (Ehlers, 1997).

Pala and Pons (1995) indicate the main advantage of deterministic approaches is the high precision and robust solution available, while the major disadvantages result from the complexity of implementation and input data requirements, and the computational intensity of the process. Consequently, the application of orbital (deterministic) models has currently been restricted to the preprocessing of raw data by satellite ground stations to provide basic corrections, such as earth curvature and rotation, panoramic distortion for off-nadir pointing sensors and non-linearity of the sensor movement.

The development of digital photogrammetric systems has encouraged ongoing research in this area, however the application has been mainly for cartographic mapping from SPOT XS stereoscopic images. Experience with digital aerial photography indicates that as the resolution of images increases, so more sophisticated rectification methods are required (Novak, 1992). Launch of high resolution commercial satellites will result in further development in this area, however most future satellite data for high precision mapping will be supplied already geocoded and it is expected few commercial image processing packages will offer deterministic rectification modules for satellite data.

The current research therefore relies upon the application of the global polynomial approach to rectification, which will be applied to all satellite datasets. While established as a routine approach for Landsat MSS, Landsat TM and SPOT HRV data, its application to the rectification of IRS1-D PAN data enables evaluation of the approach for rectification of high spatial resolution satellite data.

3.3 Image Resampling

Image resampling is required as an integral component of the rectification process whereby new pixel locations are computed based upon the rectification parameters, extent of the image and the output pixel dimension. Each new pixel is defined by real number coordinates, which do not necessarily coincide with integer pixel locations of the input image. Consequently, an interpolation procedure, termed resampling, is required to allocate brightness values for each of the output pixels. Ehlers (1997) refers to this as a *pixel-filling* approach that ensures every output pixel is addressed only once during the process, and that no gaps occur in the output image.

Important components of resampling are the selection of the resampling algorithm and specification of the output pixel resolution. Selection of an inappropriate resampling algorithm can have a deleterious effect on the intensity values of the output pixels of the rectified image, and specification of an inappropriate pixel resolution may lead to degradation of the spatial quality of the data. The following sections deal with these issues.

3.3.1 Resampling Methods

Resampling methods that are generally available in image processing systems include the nearest neighbour, bilinear and cubic convolution interpolation techniques (Dikshit and Roy, 1996).

Nearest neighbour interpolation is the simplest technique whereby the value of the output pixel is allocated the brightness value of the nearest pixel in the input image. The procedure is computationally simple, however it introduces pixel level geometric discontinuities (up to a maximum of $\sqrt{2}/2$ pixel), making the image appear visually disjointed or *blocky*. Dikshit and Roy (1996) indicate this may significantly affect extraction of texture (spatial) features, however others (Ehlers, 1997; Richards, 1993) state that retention of the original values is advantageous for subsequent classification or other spectral-based processing, or essential if the classes resampled are nominal map classes.

Bilinear interpolation is accomplished by interpolation from the nearest four adjacent neighbouring pixels in the input image using an inverse distance squared weighting algorithm of the form (Jensen, 1996):

$$BV = \frac{\sum_{k=1}^{4} \frac{z_k}{D_k^2}}{\sum_{k=1}^{4} \frac{1}{D_k^2}} \qquad (3.6)$$

The bilinear interpolator also acts as a spatial moving filter that subdues extreme brightness values throughout the output image. It may provide a much more visually appealing result, but may also degrade some image detail (Ehlers, 1997).

Where an image of high visual quality is required, such as for photointerpretation, or when the image is to be viewed under magnification, cubic convolution resampling is recommended (Richards, 1993). The most common form of cubic convolution algorithm is based upon a sin(x)/x function (Ehlers, 1997). Cubic convolution uses a 4 x 4 matrix of pixels that surrounds the output pixel in the input image to interpolate the new value. Interpolation is first undertaken in the y-direction for each of the four vertical lines in the matrix to determine the brightness value equivalent to the x coordinate of the output pixel. Interpolation is then undertaken in the x direction to determine the equivalent brightness value of the output pixel.

In this research all resampled images are to be subsequently used for multispectral classification, consequently maintenance of the original brightness values of the pixels is important. The radiometric smoothing along boundaries produced by both the bilinear and cubic convolution interpolation algorithms will also affect the classification accuracy by increasing the proportion of mixed pixels in the data. Forster and Trinder (1984) indicate that the specific resampling algorithm applied does not appear to affect subsequent percentage classification accuracy, however boundary pixels were seen to move between spatially adjacent classes as a result of changes in resampling strategy. Marceau *et al* (1994) also report that the smoothing effect can influence subsequent scale and aggregation analyses.

Richards (1993) suggests that for classification purposes images should be resampled following classification in order for the original pixel values to be preserved, and then rectified prior to map production or integration within a GIS. Alternatively, where resampling is completed prior to classification, a nearest neighbour approach should be applied. The latter approach will be utilised in this study to enable accurate integration of reference data for training site extraction and thematic accuracy evaluation, but at the same time minimising radiometric distortion of the data.

3.3.2 Resampling Interval

The resampling interval for image data determines the final dimension of pixels in the rectified image to be utilised for subsequent processing. The spatial resolution at the time of image acquisition is independent of the resampling interval, and resampling to a finer pixel size does not improve the spatial resolution compared to the original data. However, there may be implications for the spatial distribution of pixels in the resampled image when the resampling interval is significantly different to that of the original data.

Multispectral images have been acquired for change detection in this research with pixel dimensions ranging from 79 x 56 m to 20 x 20 m. For comparison of multiscale images for the purpose of change assessment, it is necessary to resample the images to a common geometric datum and pixel dimension. A suitable resampling interval must be determined, but at the same time any detrimental spatial or radiometric effects on the data must be minimised.

The following details the approach to the determination of the resampling interval of pixels for the multispectral datasets investigated as part of this study. The IRS1-D data comprise only a single panchromatic band and will not be used for land cover interpretation, but for assessment of the comparative geometric qualities of the rectification techniques and investigation of spatial error propagation in high resolution satellite data. Consequently, consideration of the resampling interval for these data in the context of the multispectral datasets is not an important issue and will not be considered further.

In a study using successive spatial aggregation of airborne MSS data, Marceau *et al* (1994) consider the following are important with respect to spatial resolution in remotely sensed data:

(i) The information content of remotely sensed images is dependent on the measurement scale determined by the spatial resolution of the sensor;

(ii) Neglecting the resolution and aggregation level when classifying images can produce unpredictable results having little correspondence in the scene;

(iii) There is no unique spatial resolution for the detection and discrimination of all targets comprising a complex natural scene.

The above comments relate particularly to the relationship between the scene and the original pixel dimension. However, resampling represents a potential source of aggregation and reduction in spatial resolution through specification of a resampling interval significantly larger than the original pixel dimension, which will degrade the radiometric and spatial information content of the data. The resampling interval must be established such that these qualities are not degraded.

Atkinson and Curran (1997) indicate the resolution of remotely sensed data should be a balance between a pixel size sufficiently large to acquire the desired information with the minimum possible data, but fine enough to capture the variation of interest within the target. Therefore, the spatial resolution of the sensor should be much finer than the resolution at which the maximum local variance of the target occurs because, if it is not, the spatial variation of interest in the target may be lost. In order to retain the radiometric and spatial qualities of each image it is necessary to resample at least to the same spatial resolution as the original data. The question whether the spatial distribution of the pixel values is retained during the resampling process is nevertheless important, and a pixel resolution finer than that of the original data should be selected. Conversely, resampling to a larger pixel size will degrade the spatial resolution and certainly alter the radiometric distribution of the pixels in the image.

In a study that compared Landsat TM and SPOT HRV images, Hill and Aifadopoulou (1990) resampled data using nearest neighbour interpolation to the coarsest pixel size of 30 m, on the assumption that all spectral and radiometric qualities of the data would be

retained. In this situation there will be one third less output pixels than input pixels for the SPOT data, resulting in the loss of both spatial and radiometric detail. Conversely, Jakubauskas *et al* (1990) compared Landsat MSS and Landsat TM imagery and resampled all data using a nearest neighbour algorithm. They also chose a pixel dimension of 30 m, but in this case there were at least as many output pixels in both the Landsat MSS and Landsat TM datasets as there were in the input data, and the spatial and radiometric contents of the data were maintained.

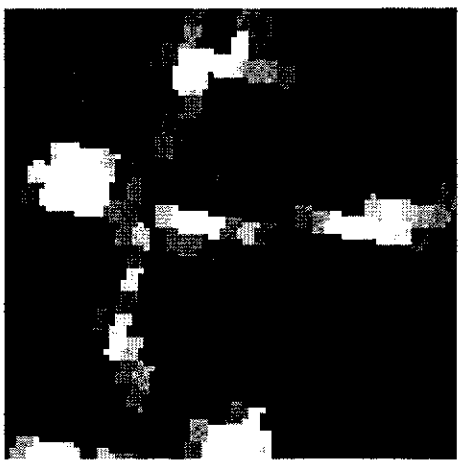
Lunetta *et al* (1991) suggest it is possible to resample to such an extent that the geometric and radiometric attributes of the resampled pixels have a poor relationship to the original data. The authors provide an example where cubic convolution interpolation was utilised to resample Landsat MSS (79 x 56 m) data to a 10 x 10 m pixel suitable for merging with SPOT PAN data at 10 m resolution. Given that cubic convolution interpolation uses a 4 x 4 kernel, the brightness value for the resampled 10 m pixel was effectively derived from a 316 x 224 m window. This approach was most likely used for integration of bands for production of a composite image where aesthetic considerations were important, rather than for analytical purposes. A more feasible solution for digital analysis could be obtained by applying a nearest neighbour interpolation that retains exact brightness values with sufficient repetition of pixel values to populate the more numerous pixel locations of the output image.

This latter approach has been applied in this research to reduce the multiscale satellite data to a common resolution for analysis purposes. All multispectral image data have been resampled to a standard 20 m pixel equivalent to the resolution of the SPOT HRV data. Any changes to the spatial and radiometric quality of the data are most likely to occur when changes in pixel dimension during resampling are most severe, as is the case with resampling Landsat MSS data to pixels of 30 m or less. Figure 3.1 illustrates the effect of resampling Landsat MSS data to pixels of dimension 5, 10, 20 and 60 m. A maximum pixel size of 60 m was selected as there are approximately the same number of pixels in the output image as the input image. This enables a near one-to-one mapping of resampled pixels to occur, and therefore does not result in any loss of information during resampling. The 20 m pixels were selected to coincide with the

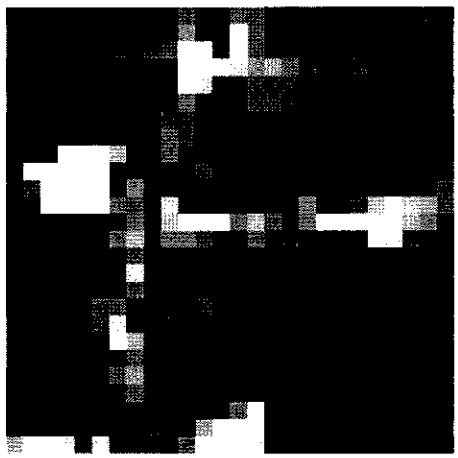
resolution of the SPOT HRV data, and 5 and 10 m resolutions to investigate effects of further reducing the pixel dimension.

(a) 5 m pixel

(b) 10 m pixel



(c) 20 m pixel



(d) 60 m pixel

Figure 3.1 *Effect of pixel resampling interval on portion of the Landsat MSS image*

Resampling to a 60 m pixel produces images where pixels are constrained to the orientation of the mapping axes. This is at variance with the actual geometry at acquisition where the track of the satellite is approximately 9 degrees off the meridian, resulting in the inclination of scan lines that are rotated from the equator by the same amount. One-to-many mapping of pixels at finer resampling intervals enables the resampled images to more closely match the geometric pattern of land surface features

at observation. This effect is illustrated in Figure 3.1 for the 5, 10 and 20 m resampling intervals where the data are rotated as they were at observation.

Representation of off-diagonal features by 60 m pixels is constrained because there is insufficient pixel redundancy in the output image to permit adequate representation of the actual direction of the linear features. Subsampling of pixels to a smaller size during rectification permits the resampled images to make an incremental representation of shape compared to the *blocky* appearance using a 60 m pixel. While not improving the primary spatial resolution of the data, resampling to a higher resolution enables the spatial distribution of pixels in the resampled image to more closely match that of the original data, and presumably also the ground scene. A similar analogy may be developed for the Landsat TM and SPOT HRV data, however these data are recorded at a higher spatial resolution and the change in pixel dimension, orientation and location during resampling is a less serious issue.

Dikshit and Roy (1996) and Forster and Trinder (1984) state that the most serious effects of resampling on the brightness structure of remotely sensed images occurs along boundaries between spectrally distinct regions. These comments were made in support of the use of cubic convolution and bilinear interpolation in order to minimise the effect of resampling on the extraction of texture components in the scene. However, for the sake of radiometric integrity and to enhance multispectral classification results that do not include texture analysis, a nearest neighbour approach should be adopted to minimise any deleterious effects on boundary pixels. Additionally, to reduce any smoothing effects or boundary displacement, a resampling interval at least as small as the smallest input pixel should be adopted as indicated above.

Landsat MSS data of the study area resampled to 20 m and 60 m resolutions were subtracted to examine the effect of the resampling interval on individual pixel values. The resultant difference image has a mean of 0 and standard deviation of 1.78, with 94.0 percent of pixels changing by less than two standard deviations (three digital numbers). Figure 3.2 indicates the pattern of differences for the same area shown in Figure 3.1. Only variations greater than two standard deviations are highlighted. Most variations are observed along boundaries between spectrally distinct areas and arise due to

differential displacement of pixels as a result of variations in the resampling interval. Based upon the findings above and those by Dikshit and Roy (1996), it is important when using multiscale data for change detection that an appropriate resampling interval is selected that permits accurate representation of ground features and minimises the degradation of spatial and radiometric qualities.

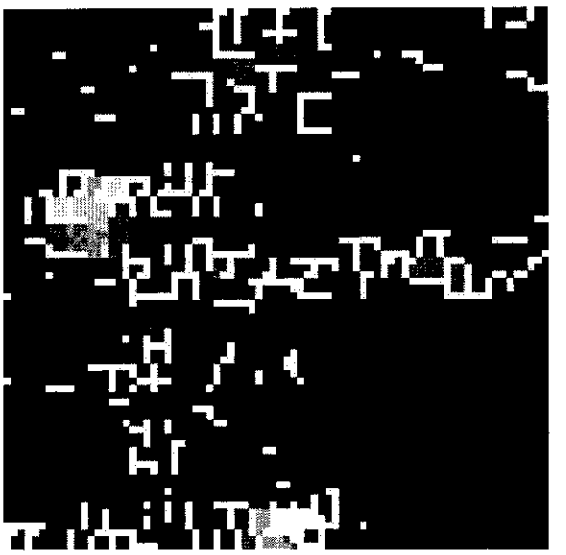


Figure 3.2 *Area shown in Figure 3.1 indicating differences (white) between 60 m and 20 m resampling interval*

The images under consideration contain a range of targets (see Section 2.2), which all contain unique radiometric and spatial characteristics and, according to Marceau *et al* (1994), may be discernible at different spatial resolutions. In order to maintain the maximum radiometric content and spatial integrity of the observed data from all sensors, a uniform resampling interval of 20 m was selected. This sampling interval is equivalent to the spatial resolution of SPOT HRV, similar to that of Landsat TM, but significantly smaller than the Landsat MSS sensor.

3.4 Ground Control Points

3.4.1 Requirements

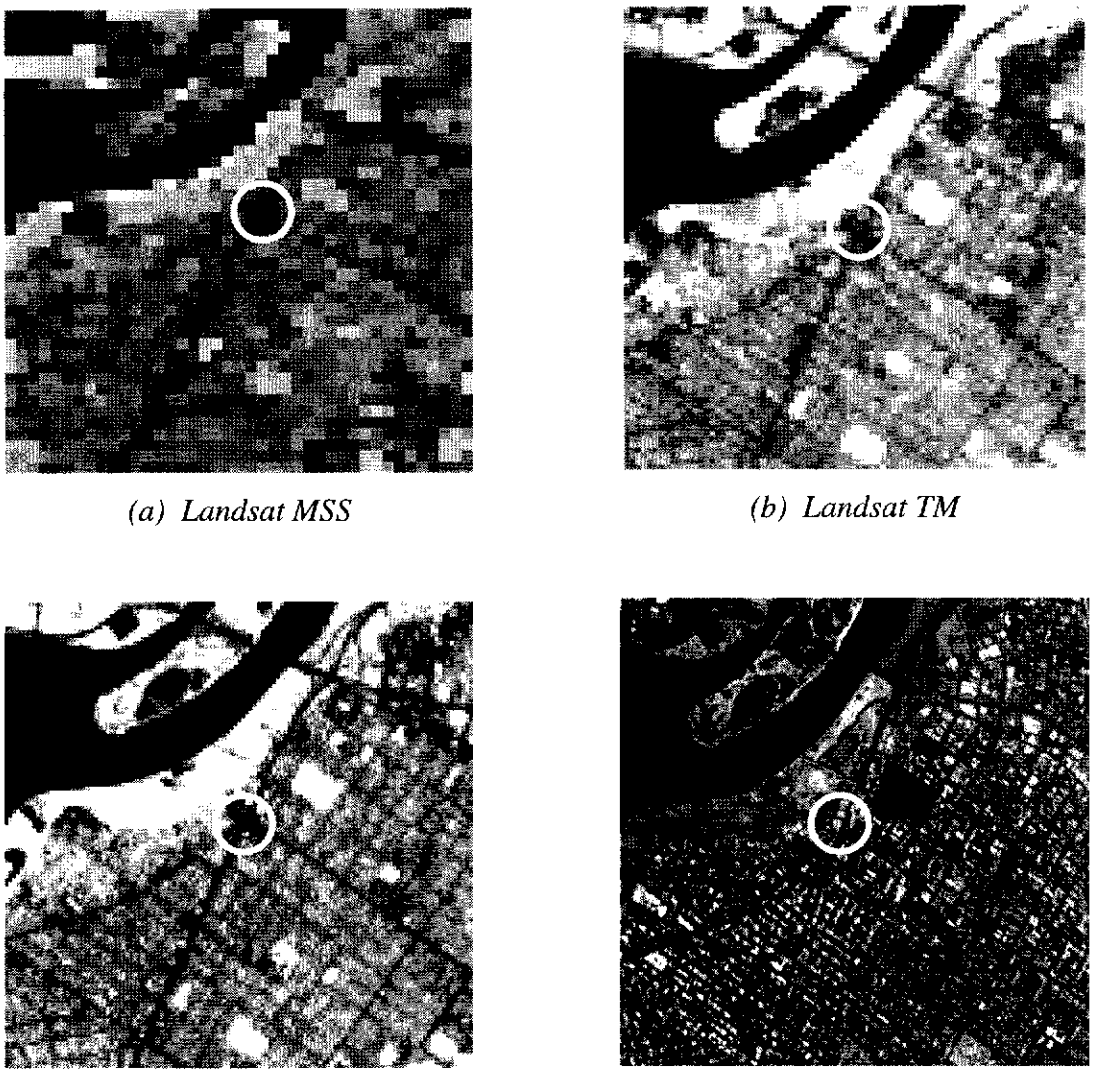
Fonseca and Manjunath (1996) indicate the most difficult step in image rectification is establishment of the transformation relationship between features on the ground and in the image, but that the computation of the mapping function is relatively

straightforward. The requirement for carefully selected GCPs of adequate precision is highlighted by enumeration of the sources of GCP-related error that affect the quality of image rectification. Important sources of error are considered to include the spatial resolution of the image, and the precision, number and distribution of GCPs (Kardoulas *et al,* 1996; Labovitz and Marvin, 1986; Welch *et al,* 1985). Image rectification errors caused by terrain relief variations have been previously discussed, and may be relevant when selecting individual control points.

The spatial resolution of the image data influences the ability to discern objects that are used as GCPs, and affects image qualities such as shape and tone. For identification purposes, control points must be at least the dimension of a pixel and have sufficient contrast with the background. For coarse resolution data such as Landsat MSS, relatively large features are required including major roads, railways and land-water interfaces. The ability to determine the precise coordinates of such features is limited and may result in random errors of position. Forster (1980a) indicates GCP location on Landsat MSS for small ground features on uniform backgrounds may be estimated to the nearest 0.5 pixel with a maximum error of $\pm$0.25 pixel. Where high contrast targets were utilised, similar results were reported by Welch *et al* (1985) in a rural-urban fringe area using visual displays of Landsat TM data on a computer screen. However, in the same study less well defined features could only be identified to $\pm$0.5 or $\pm$1.0 pixel depending upon target contrast and shape. Given the nature of most imaging and display systems, subpixel estimation of GCP locations is problematic (Welch and Usery, 1984), and heavily dependent upon scene characteristics.

Figure 3.3 shows a sample of image data from this study which has been used for selection of GCPs. The feature circled is the intersection of two four lane bitumen roads, Canning Highway and Berwick Street in the Perth metropolitan area. Shown are representations of the same areal extent for each type of data in order to indicate the influence of spatial resolution on GCP selection. In general, the same points were not selected as GCPs in each of the datasets due to decreasing levels of aggregation as the spatial resolution became increasingly finer. In the higher resolution imagery it was possible for more precisely defined GCPs to be identified. However, with the complex range of targets available, consistent identification of GCPs at the subpixel level was

not feasible in any dataset. In all cases in this research pixel locations of GCPs for the multispectral data were measured to integer values. Sufficient points were selected for regression of the rectification parameters to provide an RMSE of GCPs for polynomial derivation of less than 0.5 pixel, and for check points of 0.5 to 1.0 pixel.



(a) Landsat MSS *(b) Landsat TM*

(c) SPOT HRV *(d) IRS1-D*

Figure 3.3 *Representation of a typical GCP in multiscale images*

More advanced image matching techniques have been developed for digital photogrammetry, and have also been applied to the determination of GCP locations. These automated approaches utilise area-based and feature-based registration algorithms that compare corresponding images and determine a precise match (Fonseca and

Manjunath, 1996). This approach has received most attention for registration of multisensor and multidate images where manual methods resulted in an RMSE of 0.87 pixel, compared to an RMSE of 0.32 pixel for the automated image matching algorithm (Dai and Khorram, 1997). These techniques will be more commonly utilised in the future, however were not available for this research.

Evaluation of the precision of GCP coordinates extracted from analogue or digital maps is related to the data compilation standards, map scale and digitising accuracy. Digitising accuracy is typically $\pm$0.25 mm and the standard deviation of planimetry for most topographic maps is $\pm$0.25 mm at map scale. The combined effect at map scale is given by:

$$\sqrt{0.25^2 + 0.25^2} \quad = \quad 0.35 \text{ mm} \tag{3.7}$$

This value agrees closely with that used by Welch and Usery (1984) and Cook and Pinder (1996) for rectification of Landsat TM data where GCP coordinates were extracted from 1:24 000 United States Geological Survey (USGS) maps. Kardoulas *et al* (1996) digitised GCPs for rectification of Landsat MSS, Landsat TM and SPOT PAN data from 1:100 000 topographic map sheets with an estimated precision of $\pm$50 m. Poor results were achieved, with RMSE values for rectification of the SPOT PAN data greater than the RMSE values for the Landsat MSS datasets. The limiting factor in the Kardoulas study appeared to be the precision of the GCPs derived from the reference data. For Landsat MSS it was equivalent to 0.6 pixel, yet for the SPOT PAN data it was equivalent to 5 pixel.

GCP precision of approximately 0.5 pixel represents a viable limit for rectification of satellite data using coordinates extracted from existing maps. Improvements in GCP positioning are possible using GPS techniques, however the significance of the inclusion of these high precision data in the rectification process is yet to be determined.

GCP coordinates for this study were extracted from a variety of digital mapping products. GCPs for rectification of Landsat TM and SPOT HRV data were derived from the *Perth 1998 StreetExpress* digital street directory. These data are compiled to a

planimetric standard of ±12.5 m and can be viewed under high zoom on the computer screen. The pointing accuracy is therefore very high and the precision of extracted coordinates is equivalent to ±12.5 m, which represents just over 0.5 pixel for the SPOT HRV data and 0.4 pixel for the Landsat TM data.

GCP coordinates for rectification of the IRS1-D data were extracted from the MRWA digital road centreline dataset which is compiled from photogrammetric, GPS and ground surveys to a precision of between 1 and 3 m, with most points meeting the 1 m specification (MRWA, 1999). The precision of these coordinates is equivalent to less than 0.5 pixel in the IRS1-D dataset, and is therefore suitable for rectification purposes.

Labovitz and Marvin (1986) report on a NASA specification that 90 percent of randomly selected points on geocoded satellite image products should be within 0.5 pixel of the ground position. There appear to be no standards specific to the production of satellite products, except where they are compared to topographic maps of particular scales. For example, Welch and Usery (1984) indicate that an RMSE of 25 m for Landsat TM data meets the US National Map Accuracy Standard for 1:100 000 scale, and several researchers indicate rectification results equivalent to 0.5 to 1.0 pixel are acceptable for change detection (Jensen *et al*, 1997; Townshend *et al*, 1992). This value most likely represents the limit of rectification precision unless sophisticated image matching approaches are employed. Image rectification undertaken in this study will be assessed as acceptable if the 0.5 - 1.0 pixel standard for check points is achieved, and its effect on change detection will be assessed.

GCPs should be well distributed around the edges and generally over the extent of the image to ensure adequate control of the interpolating polynomials throughout the image (Richards, 1993). The behaviour of interpolating polynomials also varies with the order of the interpolation function and the number of control points. These two issues were discussed in Section 3.2.1. Labovitz and Marvin (1986) investigated the effect of the spatial location of control points and concluded the distribution of points was unimportant provided the points were not all clustered near one edge of the scene. Common practice is for GCPs to be evenly spread throughout the image, which is less

significant when using the more rigid first or second order polynomials, and this approach will be followed in this research.

Figure 3.4 illustrates the distribution of control points in each of the images analysed. An adequate distribution of control was provided in the vicinity of the study area indicated on the diagram, although the distribution is not uniform throughout the rectified area. While the quality of the rectification cannot be guaranteed in areas not covered by GCPs, analysis in Section 3.5 examines the effect of the number and distribution of GCPs.

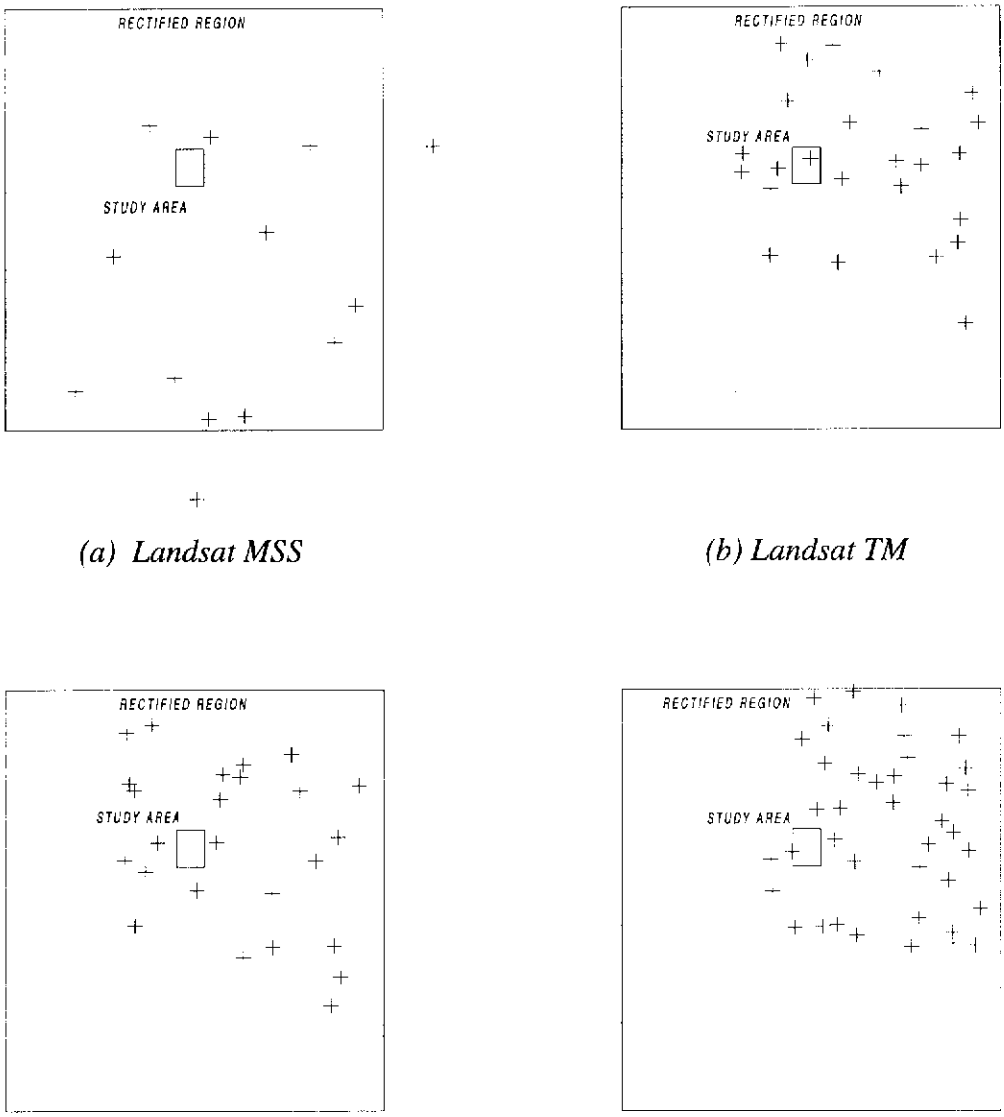
Evaluation of image rectification is usually based upon analysis of GCP residuals through the computation of the RMSE. Most studies state an RMSE for rectification, however it is not always clear whether this relates to the *a priori* RMSE computed from the GCPs used for interpolation, or the *a posteriori* RMSE computed from independent check points. Townshend *et al* (1992) utilised the same GCPs for computation of the RMSE, but cautioned that the real (*a posteriori*) RMSE value is likely to be somewhat higher.

Kardoulas *et al* (1996) state that sufficient GCPs must be selected in order to provide some points for rectification and others as check points for *a posteriori* consideration of the quality of the rectification. This permits not only evaluation of the interpolating polynomials, but also assessment of the complete rectification process including any effects of relief displacement and other error sources not considered. The major disadvantage of this assessment is that it is impossible to isolate specific sources of error contributing to the overall RMSE. Evaluation of all rectification processes in this research will be undertaken in Section 3.5 using *a posteriori* computation of the RMSE from independent check points.

3.4.2 Point Selection

GCPs were extracted from each image according to the considerations discussed in Section 3.4.1. Control points were required for the final analysis of each dataset in order to provide sufficient points for formation of the interpolating polynomials and for computation of the RMSE from independent check points. The minimum number of

GCPs required for a third order polynomial interpolation is 10, therefore it is desirable to identify at least 40 points in order to allow for exclusion of some points from the interpolating process due to excessive residuals.



(a) Landsat MSS *(b) Landsat TM*

(c) SPOT HRV *(d) IRS1-D*

Figure 3.4 *Distribution of GCPs for rectification of satellite images*

Adequate points were identified for the Landsat TM, SPOT HRV and IRS1-D data, however due to the relatively small scale of the reference information relating to the Landsat MSS data, only 22 reliable GCPs were identified.

Figure 3.5 indicates the overall process used for selection and processing of GCPs. GCPs were initially processed in a single file for identification of outliers. GCPs were eliminated from the overall GCP file if the point showed an RMSE greater than 1.0 pixel. When all RMSE values were less than 1.0 pixel, the file was split, with a majority of points used for interpolation and the balance used as check points.

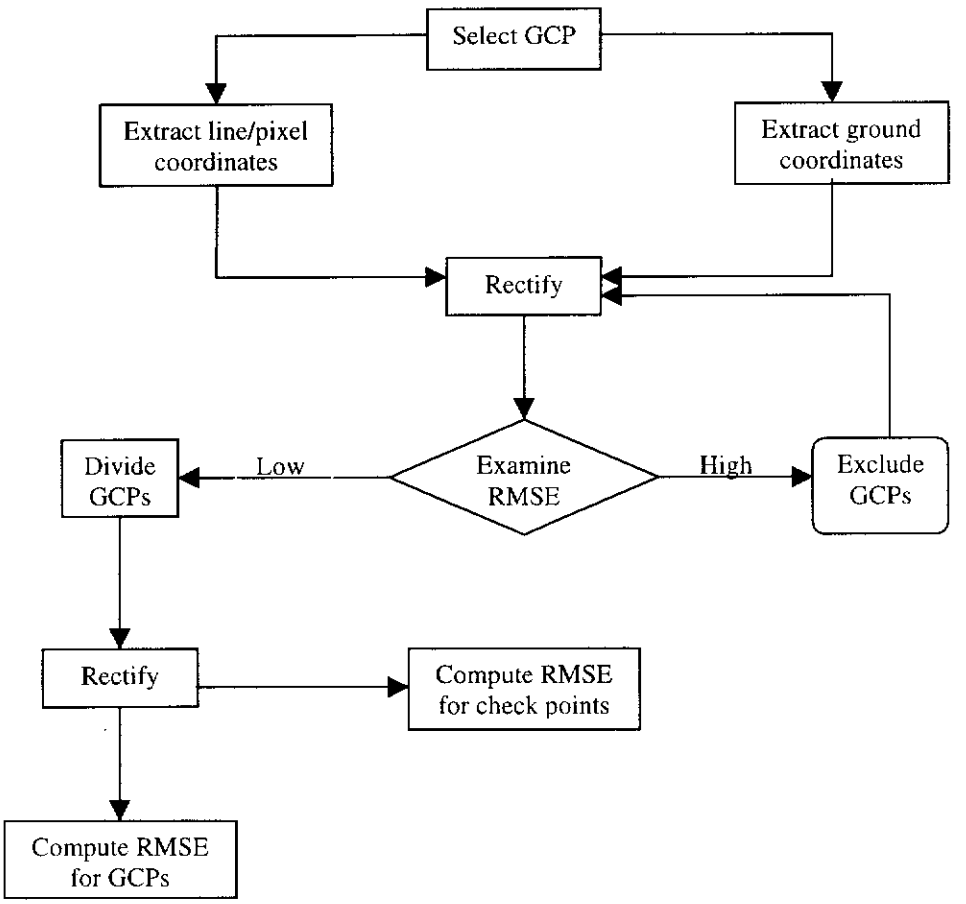


Figure 3.5 *Flowchart showing GCP selection and analysis*

GCPs were selected with a uniform distribution across the image and were based upon well defined points identifiable in both the image and the planimetric reference data. Each dataset was treated independently due to the wide range of spatial resolutions of the images, therefore a clearly defined point in the SPOT HRV or IRS1-D data may not

have been visible at all in the Landsat MSS data. As a result there are very few control points common to all images. Zoom facilities available in the image processing system were utilised to facilitate optimum determination of the row and column values for each GCP.

3.5 Georeferencing

3.5.1 Rectification Results

Rectification of each image was performed and an independent evaluation of the quality of the rectification undertaken. Table 3.2 details the RMSE values for GCPs utilised to derive the transformation parameters (*a priori*), and for independent check points (*a posteriori*). The values are provided in units of input pixels for each dataset, with the RMSE for check points consistently larger than the corresponding value for derivation of the transformation parameters. Note that for Landsat MSS, Landsat TM and SPOT HRV images a first order polynomial was used while for the IRS1-D image a cubic polynomial was utilised for rectification.

Points	MSS		TM		SPOT HRV		IRS1-D	
	No Pts	RMSE	No Pts	RMSE	No Pts	RMSE	No Pts	RMSE
GCPs	14	0.417	25	0.402	26	0.345	40	0.460
Check points	8	0.826	23	0.434	27	0.407	30	0.549

Table 3.2 *Summary of rectification results for polynomial regression and quality evaluation (RMSE in pixels)*

Tables 3.3, 3.4, 3.5 and 3.6 list details of residual values derived from *a posteriori* analysis of independent check points in the area covered by the rectification. Input pixel coordinates were measured to the integer level on the image for Landsat MSS, Landsat TM and SPOT HRV images. Measurement of GCPs to 0.1 pixel for the IRS1-D image and computation of rectified values to 0.001 pixel through the transformation process represents a characteristic of software operation rather than a requirement for higher precision. Residual values, in pixel terms, are larger for the Landsat MSS data than for the other datasets, while the residuals for the Landsat TM, SPOT HRV and IRS1-D data in pixel units, are all similar.

POINT	INPUT X	INPUT Y	RECTIFIED X	RECTIFIED Y	RESIDUAL
1	65	492	64.328	491.482	0.849
2	287	523	286.963	522.302	0.699
3	466	391	465.217	390.264	1.075
4	544	302	544.483	302.044	0.485
5	546	166	545.391	165.330	0.906
6	421	157	420.335	157.429	0.791
7	405	284	405.513	283.243	0.915
8	58	175	58.584	175.489	0.762
				RMSE	**0.826**

Table 3.3 *A posteriori RMSE of Landsat MSS image (RMSE in pixels)*

POINT	INPUT X	INPUT Y	RECTIFIED X	RECTIFIED Y	RESIDUAL
1	944	785	943.792	785.491	0.533
2	831	902	831.226	901.847	0.273
3	774	963	773.686	962.904	0.329
4	701	750	701.294	749.638	0.466
5	777	621	777.256	621.592	0.645
6	752	547	751.697	547.248	0.392
7	832	668	832.222	668.478	0.528
8	868	169	867.716	168.582	0.506
9	1087	219	1087.140	218.619	0.406
10	1069	112	1068.566	111.869	0.453
11	1145	260	1145.169	260.102	0.197
12	221	901	221.103	901.008	0.103
13	361	896	360.973	896.191	0.193
14	377	837	376.640	837.023	0.361
15	270	681	269.564	680.864	0.457
16	586	881	586.371	880.748	0.448
17	570	718	569.769	717.537	0.518
18	434	567	433.795	566.712	0.353
19	549	635	548.944	635.260	0.266
20	471	398	470.814	398.021	0.187
21	388	360	388.000	360.507	0.507
22	402	433	402.205	433.038	0.208
23	520	200	519.167	200.232	0.865
				RMSE	**0.434**

Table 3.4 *A posteriori RMSE of Landsat TM image (RMSE in pixels)*

Figure 3.6 illustrates error vectors for the *a posteriori* residual values derived from the image transformation. All residuals are shown in input pixel units with the cross indicating the control point location. The objective of these residual plots is to indicate

any systematic patterns in the magnitude and distribution of residual values. Consistently large vectors indicate a poor fit to control or insufficient GCPs in the transformation. Regions where residuals demonstrate consistent direction may suggest errors in the GCPs used in the transformation, and indicate sufficient cause for inspection of the RMSE of the GCP and its coordinates.

POINT	INPUT X	INPUT Y	RECTIFIED X	RECTIFIED Y	RESIDUAL
1	299	1269	299.226	1269.334	0.403
2	734	1364	734.561	1363.907	0.568
3	401	1115	401.221	1115.046	0.225
4	402	910	401.960	910.516	0.517
5	730	890	729.704	890.213	0.365
6	764	805	764.634	804.968	0.635
7	844	1253	844.423	1252.765	0.484
8	1051	1286	1051.268	1285.799	0.335
9	1402	1113	1402.263	1113.293	0.394
10	1119	1051	1118.796	1051.607	0.640
11	1314	1007	1313.856	1006.843	0.213
12	1028	783	1027.968	783.279	0.281
13	912	779	911.793	778.934	0.217
14	884	966	883.526	966.109	0.486
15	1634	727	1633.561	727.076	0.446
16	1459	1093	1459.345	1093.372	0.508
17	432	684	431.788	683.872	0.248
18	680	660	680.468	660.011	0.468
19	945	724	945.202	723.866	0.243
20	635	366	635.281	366.169	0.328
21	646	167	646.113	167.104	0.154
22	786	118	785.684	117.991	0.316
23	1025	677	1024.626	677.287	0.471
24	1272	226	1272.043	225.965	0.055
25	1060	436	1059.987	436.438	0.438
26	1466	372	1465.719	372.058	0.287
27	1712	340	1712.528	340.005	0.528
				RMSE	**0.407**

Table 3.5 *A posteriori RMSE of SPOT HRV image (RMSE in pixels)*

3.5.2 Discussion and Analysis

Previous research indicates that for image rectification *a priori* RMSE values should be less than 0.5 pixel, and that *a posteriori* values should be less than 1.0 pixel. Table 3.2 indicates these values have been achieved in this study with all *a priori* values substantially less than 0.5 pixel, and in the case of SPOT data, an RMSE as low as 0.345 pixel. From an operational viewpoint, the *a posteriori* RMSE is of more interest

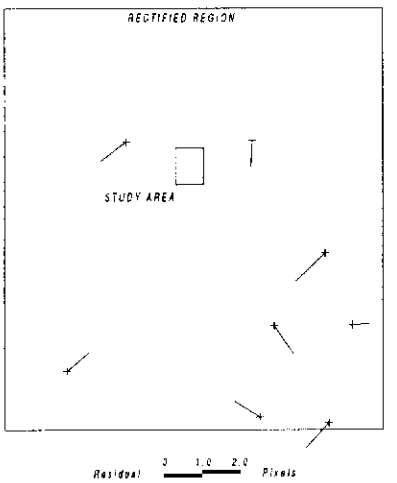
because it provides an independent estimate of rectification quality. The *a posteriori* RMSE values for the Landsat TM and SPOT HRV data were particularly low at 0.434 and 0.407 pixel respectively, however the Landsat MSS RMSE was substantially higher at 0.826 pixel.

POINT	INPUT X	INPUT Y	RECTIFIED X	RECTIFIED Y	RESIDUAL
1	158.7	3952.6	158.652	3952.381	0.224
2	832.7	4152.0	832.524	4152.138	0.224
3	1220.4	4470.0	1221.070	4470.033	0.670
4	1156.4	4039.0	1156.611	4039.154	0.262
5	862.7	3409.6	862.840	3409.257	0.370
6	1363.1	3439.9	1363.627	3439.920	0.528
7	1305.4	4256.6	1305.278	4256.684	0.148
8	1701.5	3472.9	1701.345	3472.605	0.333
9	2253.8	3361.9	2253.669	3361.127	0.784
10	2538.5	4102.0	2538.024	4102.574	0.746
11	836.4	2895.5	836.371	2895.619	0.123
12	1416.8	3027.9	1417.495	3028.098	0.722
13	284.8	1851.4	284.243	1851.726	0.645
14	605.3	2161.1	604.531	2161.109	0.769
15	1511.1	2718.2	1511.212	2717.976	0.250
16	2038.2	2803.5	2038.300	2803.776	0.294
17	2217.8	2474.8	2217.833	2475.230	0.432
18	2384.8	1892.5	2385.233	1891.969	0.685
19	2955.9	2066.5	2955.737	2065.906	0.616
20	480.7	544.4	481.028	543.935	0.569
21	1168.0	348.5	1167.940	348.662	0.173
22	789.0	1398.3	789.505	1398.311	0.505
23	1527.8	1653.5	1528.017	1652.982	0.562
24	1723.3	515.4	1723.966	514.701	0.965
25	2621.2	111.3	2621.401	111.284	0.202
26	2773.2	1791.8	2772.761	1791.868	0.444
27	2206.5	1142.7	2207.160	1141.830	1.092
28	2987.6	1652.1	2987.316	1651.819	0.399
29	3439.3	853.1	3439.670	852.522	0.687
30	3264.2	339.4	3264.452	339.350	0.257
				RMSE	**0.549**

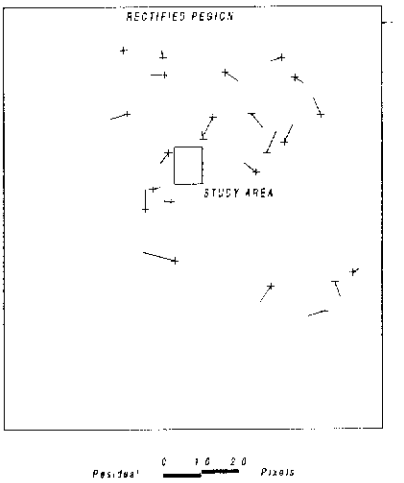
Table 3.6 *A posteriori RMSE of IRS1-D image (RMSE in pixels)*

The Landsat MSS RMSE results from moderate but uniform residuals for all check points (Table 3.3), indicating that no gross errors in GCP identification or polynomial determination are present. Figure 3.4a shows a uniform GCP distribution across the image, and examination of Figure 3.6a does not indicate any systematic direction or localised distortion in the residual vectors resulting from the rectification process.

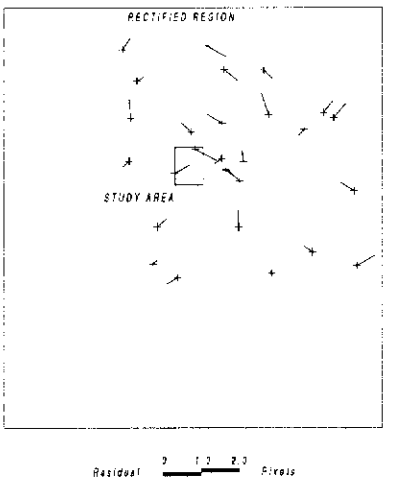
These results indicate the transformation process was geometrically sound, and a relatively high *a posteriori* RMSE is therefore due to the limitations inherent in the identification and positioning of GCPs for the Landsat MSS dataset.



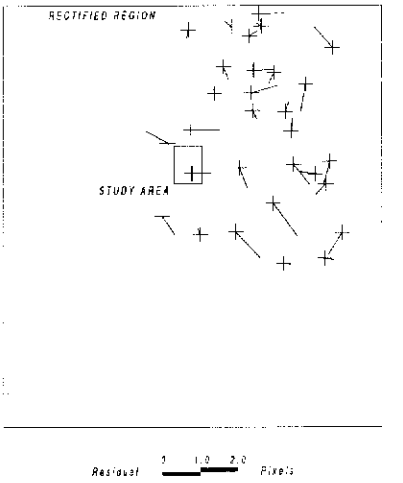
(a) *Landsat MSS image*



(b) Landsat *TM image*



(c) SPOT HRV



(d) IRS1-D

Figure 3.6 *A posteriori error distribution and residual vectors*

All GCPs for rectification of the Landsat MSS data were derived from 1:100 000 topographic maps. Estimation of coordinates to $\pm$0.35 mm at map scale (see Equation 3.7) or $\pm$0.5 pixel is possible, but identification of GCPs on the imagery is much more

difficult compared to the other data sources. Additionally, fewer GCPs were located for polynomial determination, therefore higher residuals for check points could be expected compared to other datasets.

Table 3.4 lists residuals for the Landsat TM data that range from 0.103 to 0.865 pixel, with no gross errors or inconsistencies apparent. Distribution of GCPs in Figure 3.4b is affected by the areal extent of the data. The available data did not cover the lower third of the rectified region, however they were adequate to include the study area and permit selection of suitable control. Residuals for individual check points within the SPOT HRV data are shown in Table 3.5 and range from 0.055 to 0.568 pixel, thus indicating the consistency of the rectification. The distribution of GCPs is demonstrated in Figure 3.4c. No GCPs are located in the lower one third of the rectified area as satellite data were not available for this region, otherwise a uniform distribution of GCPs is evident.

Examination of residual vectors for Landsat TM and SPOT HRV data in Figures 3.6b and 3.6c respectively does not indicate any systematic directional distortion. In both datasets a minimal level of localised distortion is evident in the northeast part of each image, however examination of Figures 3.4b and 3.4c demonstrates that adequate GCPs are available in this region for both datasets and GCP distribution is not a contributing factor. Terrain elevations are relatively uniform in the area of the image and do not increase apart from the eastern margin of the rectified area. In the most extreme case, elevation variations across the complete image are less than 150 m and relief displacement should not contribute significantly to errors in the rectification process (see Section 3.2.1). The average residual in this area for both datasets is approximately 0.5 pixel and is considered to be well within the 1.0 pixel threshold for the *a posteriori* RMSE.

Check point residuals in Table 3.6 indicate variations between 0.123 and 1.092 pixel for the IRS1-D data, and demonstrate there are no gross errors or inconsistencies in the rectification. The distribution of GCPs shown in Figure 3.4d is determined by the available data which only cover the northeastern area of the rectified region, which restricted GCP selection to this area. A plot of check point residuals in Figure 3.6d does not indicate any systematic pattern and demonstrates an apparent random distribution of large and small residuals, both in terms of direction and magnitude. The third order

polynomial is therefore appropriate for rectification of these high resolution data, and indicates that current image processing systems have the potential to undertake geometric rectification of data from the next generation of satellite scanners.

Welch and Usery (1984) indicate the internal geometric consistency of Landsat MSS to be good, but the geometric quality of Landsat TM data to be even higher. Consequently, transformation error is likely to be due more to error in determination of GCPs than propagation through the transformation process. They indicate the precision of rectification is dependent upon the precision of GCP measurement (location) error in the image, digitising error of the map, and relief displacement effects not taken into account during the transformation as follows:

$$\sqrt{(location\ error)^2 + (digitising\ error)^2 + (relief\ error)^2} \tag{3.8}$$

For Landsat MSS the maximum location error for a pixel is considered to be equal to half the diagonal of a 79 x 56 m pixel (49 m). Digitising error incorporates both the act of digitising and the precision of map compilation. From Section 3.4.1 this was determined to be 0.35 mm at map scale, and for a 1:100 000 scale is equivalent to 35 m. The effect of relief displacement in all data used in this research may be computed (from Pala and Pons, 1995):

$$\frac{L\,z}{(H - z)} \tag{3.9}$$

where

L = *distance of the pixel from the satellite path*

z = *terrain height of the pixel*

H = *altitude of the satellite.*

With the range of elevations in the area of the Swan Coastal Plain being less than 30 m, the effect of relief displacement at the edge of a Landsat image is a maximum of 3 m. From Equation 3.8 the precision of rectification of Landsat MSS due to GCP measurement errors is therefore computed to be:

$$\sqrt{49^2 + 35^2 + 3^2} \quad = \quad 60 \text{ m} \tag{3.10}$$

This value compares favourably with the *a posteriori* RMSE of 0.826 pixel, which is equivalent to 65 m as a proportion of the Landsat MSS ground resolution element. Similar evaluations for Landsat TM, SPOT HRV and IRS1-D data result in predicted values of 25, 19 and 5 m, respectively. However, the computed *a posteriori* RMSE for these data are only 13, 8 and 3 m.

The relatively high *a posteriori* value for the Landsat MSS data is due to lower geometric consistency of the data and the availability of fewer GCPs for determination of the transformation parameters. These image data are from 1972 and the only planimetric reference data available were derived from a 1:100 000 map topographic sheet produced in 1976. This map provided only generalised detail and therefore limited the opportunity for selection of control. Welch *et al* (1985) conclude the major influence on the rectification of Landsat TM data is the spatial resolution of the imagery. This also applies to a greater extent to Landsat MSS and less to SPOT HRV data, and in each case the contribution of point identification error in the image is the largest single source of positioning error.

Detailed and up to date digital map data were available for the Landsat TM, SPOT HRV and IRS1-D data (see Section 3.4.1), and selection of a relatively large number of well distributed GCPs was possible. In conjunction with improved geometric quality of the data from these sensors, substantially improved *a posteriori* RMSE values (13, 8 and 3 m, respectively) were obtained, compared to the predicted values.

Rectification precision as measured by the *a posteriori* RMSE is important, however any shift in pixel locations caused by pixel resampling must also be considered in the comparison of multidate imagery for change assessment. The displacements of pixels due to resampling are not included in the *a posteriori* RMSE calculation. Lodwick (1980) indicates that for nearest neighbour resampling average RMSE translation errors of 0.289 pixels are expected statistically, and therefore should be included as follows:

$$\sqrt{(RMSE\ transformation)^2 + (RMSE\ resampling)^2} \tag{3.11}$$

Combined with the *a posteriori* RMSE transformation error, the total estimated RMSE including spatial displacement due to resampling are 0.875, 0.521, 0.499 and 0.620 pixel for Landsat MSS, Landsat TM, SPOT HRV and IRS1-D data, respectively. These values will be used in evaluation of the geometry of the imagery for multidate analysis and provide data for error propagation assessment of the change evaluation process.

3.6 Summary

Comparison of multidate imagery for detection of land cover changes requires the development of a geometrically coherent dataset in order that spatially coincident comparisons can be made. The processes of image rectification and pixel resampling are fundamental to the derivation of these data.

With the development of new generation high resolution satellite sensors with pixel dimensions of the order of 5 m, achievement of rectification with subpixel precision requires investigation of appropriate algorithms. All data, including the high resolution IRS1-D imagery, were rectified using global polynomial rectification techniques. With improvements in the spatial resolution of future sensors, parametric techniques may be required, however investigation of polynomial techniques applied to high resolution IRS1-D data indicate that with sufficient GCPs of suitable precision, acceptable rectification results can be achieved.

First, second and third order polynomials were evaluated for rectification of each of the datasets. For Landsat MSS, Landsat TM and SPOT data there were no significant advantages in using higher order polynomials, and a first order polynomial was applied in each case. Higher resolution IRS1-D data were evaluated and a first order model was deemed to be inadequate. A third order polynomial was applied for the rectification of these data. In all cases, rectification produced *a posteriori* RMSE values of less than 1.0 pixel and, for Landsat TM, SPOT HRV and IRS1-D data, where quality reference material for derivation of GCPs were available, RMSE values of 0.5 pixel were achieved without the application of specialised processing approaches.

The polynomial rectification approaches utilised in this study have produced geometrically consistent data that are suitable for change detection analysis. While all images have been rectified to better than the dimension of one input pixel, the range of

dimensions of the input pixels means that positional variations will continue to be an issue in future analyses.

Maintenance of pixel brightness values during rectification and resampling is vital where spectral analysis such as multispectral classification is to be undertaken following geometric correction of the data. Cubic convolution and bilinear interpolation were considered, but nearest neighbour resampling was utilised to maintain radiometric consistency for subsequent classification.

Change detection analysis is facilitated through resampling all data to a common pixel dimension. To minimise the loss of radiometric information and to facilitate multiscale comparisons a 20 m pixel grid was selected so that it is identical to the SPOT HRV data, which is the smallest pixel in the multispectral dataset.

Resampling of remotely sensed data during rectification can have significant effects on the spatial location of pixels. Research indicates that resampling Landsat MSS data to a 60 m pixel results in the displacement of many pixels, which may effect the analysis of data for change detection. Spatial displacement of the Landsat MSS data has been minimised through utilisation of a resampling interval of 20 m. As differences between the spatial resolution of sensors increase, further consideration of potential spatial and radiometric distortions will be required for integration of data for change detection analysis.

Chapter 4

IMAGE PROCESSING

This chapter focuses upon the application of information extraction algorithms for identification of a range of land over classes within the study area. Supervised and unsupervised classification algorithms are used to extract land cover details from the satellite data. Successful classification of land covers is dependent upon clear recognition of ground features resulting from analysis of their spectral characteristics. Spatial autocorrelation and class differentiation analyses are used to define the spatial patterns of reflectance characteristics and to measure the spectral separability of the land cover classes for supervised classification. Stratified sampling is used to determine the optimum grouping of clusters derived from unsupervised classification into relevant land cover classes. Results from a supervised maximum likelihood classification and ISODATA unsupervised classification are presented using land cover classes equivalent to USGS Level I and Level II classifications. These results will be used for temporal comparison with similar classifications derived from data collected over a range of spatial resolutions.

4.1 Introduction

Image processing relates to the range of image enhancement and information extraction procedures that may be applied to remotely sensed data for feature identification. The goal of image enhancement is to improve the interpretability of an image by increasing the apparent distinction between features (Lillesand and Kiefer, 1994). This terminology describes processes such as contrast stretching and spatial filtering in preparation for visual interpretation, or the elimination of pixel *dropouts* and creation of principal components images as a precursor to subsequent digital image analysis. Information extraction may take the form of photointerpretation of hard copy images or quantitative analysis of digital data using automated classification approaches (Richards, 1993).

Within this study image enhancement techniques need only be investigated to the extent necessary to provide data of consistent quality suitable for classification. Data which occupy relatively small sites with minimal topographic variation are assumed to exhibit consistent atmospheric and topographic effects (Gong and Howarth, 1992). The classifications utilised in this study are applied to independent multidate datasets and are statistically invariant to linear transformations, consequently no radiometric corrections were made to the data. Destriping corrections were applied to the Landsat MSS data, but only to the extent of correcting sensor-induced degradation such that consistent levels of classification could be obtained. Contrast stretching and formation of colour composite images were also performed as an aid in the identification of training samples for supervised classification.

Evaluation of the information content of remotely sensed data and its application to land cover mapping relies upon careful definition of the land cover classes. The classification scheme must be based upon taxonomically correct definitions of the classes and be organised according to logical criteria (Jensen, 1996). Analysis of multiscale data requires further consideration of the hierarchical structure of the data and classification system. Transformation of the results from spectral classes to informational classes such as land uses is problematic, and requires careful consideration of the data spectral characteristics and land surface properties.

The approaches applicable to unsupervised and supervised classifications are somewhat different. In the context of unsupervised classification, techniques are required to assist with the allocation of spectrally based clusters to land cover classes. In the rural-urban fringe more than 20 clusters may result from the unsupervised classification, which must then be summarised by only four land cover classes. An analytical approach to cluster allocation using sampling techniques would significantly simplify the mapping process.

Much attention has also been directed towards the collection of training samples for supervised classification and their relevance for summarising the statistical parameters of the corresponding land cover classes. Spatial autocorrelation and its relationship to the *scale of observation* and the geostatistical concept of *support* (Atkinson and Curran, 1997; Collins and Woodcock, 1996b) are significant considerations in defining the manner in which training samples are collected.

Improved definition of statistical parameters for definition of land cover classes leads to enhanced classifier performance and understanding of sampling procedures. Critical parameters include the optimum size of training samples to adequately assess class variance, and assessment of class separability to ensure that spectrally separable land cover classes are identified. Analytical bases for these assessments are considered in the form of spatial autocorrelation analysis and statistical evaluation of training class separations.

4.2 Land Cover Classification Scheme

The primary objective of this research is to investigate the application of multiscale digital data for the determination of land cover changes on the rural-urban fringe. Inherent in multiscale approaches is the concept of interpretation of variable levels of detail from each dataset as discussed in Section 1.1.3. For example, forest cover may be isolated from pasture in Landsat MSS data, however, within the forest category discrimination of evergreen and deciduous forests may also be possible from SPOT HRV data. The classification scheme must be able to accommodate these various levels of interpretation. Meaningful comparison of the results of interpretation from multiscale sources is necessary for the full potential of data collection and analysis

techniques to be realised. Operational issues and the availability of monitoring data were considered in Section 1.2.

Standardised land use and land cover data are required to encourage efficient use of data collection facilities and to enable effective management of resources. Consistent and robust land use and land cover data are essential for planning and management of infrastructure development, environmental management, energy and resource development, urban planning and industrial development. Application of standardised classification schemes provides a fundamental framework for the establishment of this information for local, regional and national purposes.

The most widely utilised land use and land cover classification scheme was developed by the USGS (Anderson *et al*, 1976) and forms the basis for the approach used in this study. The USGS classification scheme comprises four levels of land use and land cover classes arranged in a hierarchical manner. Only the first two levels have been defined by Anderson *et al* (1976) and are shown in Table 4.1. It is intended that classes at Levels III and IV be developed for specific circumstances and particular combinations of characteristics, with the requirement that these classes may logically be aggregated to generalised Level I and Level II classes.

This classification scheme was designed to utilise remotely sensed data as the primary information source especially at the more generalised levels, and to incorporate ancillary or collateral data to assist in the understanding of multiple-use arrangements or for more detailed interpretation at Levels III and IV. The range of data applicable to the derivation of each level of interpretation is given in Table 4.2, which indicates that all satellite data utilised in this research is only useful for interpretation of Level I classes. A range of studies (Martin *et al, 1988;* Trietz *et al*, 1992; Wolter *et al*, 1995 and others) have derived Level II information from medium resolution satellite data, however the success has been dependent upon the target size and heterogeneity, and contrast characteristics of the objects. While it is not expected that Landsat MSS will provide reliable and consistent interpretation at Level II, all data will be evaluated for interpretation at both levels.

The USGS classification system is designed primarily to utilise interpretation of analogue imagery and does not consider the availability of higher spatial resolution satellite data nor the implications of advanced automated interpretation techniques.

LEVEL I	LEVEL II
1 Urban or built-up land	11 Residential
	12 Commercial and services
	13 Industrial
	14 Transportation, Communications, and utilities
	15 Industrial and commercial complexes
	16 Mixed urban or built-up land
	17 Other urban or built-up land
2 Agricultural land	21 Cropland and pasture
	22 Orchards, groves, vineyards, nurseries and ornamental horticultural areas
	23 Confined feeding operations
	24 Other agricultural land
3 Rangeland	31 Herbaceous rangeland
	32 Shrub and brush rangeland
	33 Mixed rangeland
4 Forest land	41 Deciduous forest land
	42 Evergreen forest land
	43 Mixed forest land
5 Water	51 Streams and Canals
	52 Lakes
	53 Reservoirs
	54 Bays and estuaries
6 Wetland	61 Forested wetland
	62 Non-forested wetland
7 Barren land	71 Dry salt flats
	72 Beaches
	73 Sandy areas other than beaches
	74 Bare exposed rock
	75 Strip mines, quarries and gravel pits
	76 Transitional areas
	77 Mixed barren land
8 Tundra	81 Shrub and brush tundra
	82 Herbaceous tundra
	83 Bare ground tundra
	84 Wet tundra
	85 Mixed tundra
9 Perennial snow or ice	91 Perennial snowfields
	92 Glaciers

Table 4.1 *Generic land use and land cover classification system for use with remote sensor data (Anderson et al, 1976)*

Higher spatial resolution generally enables more detailed categorisation of land covers, although in complex urban areas variable results have been achieved (see Section 2.2.4). Multispectral classification can provide a robust analysis technique, however most approaches rely only upon the spectral characteristics of the data and ignore components such as site, size, texture and pattern, which provide important contextual information during manual interpretation. The USGS classification scheme utilises land cover as a surrogate for land use, however contextual data often provide the final decision criterion.

CLASSIFICATION LEVEL	DATA CHARACTERISTICS
I	Medium resolution satellite data (20-100 m IFOV)
II	High altitude aerial photography (smaller than 1:80 00 scale)
III	Medium altitude aerial photography (1:20 000 to 1:80 000 scale)
IV	Low altitude aerial photography (larger than 1:20 000 scale)

Table 4.2 *Data characteristics equivalent to levels of interpretation for remotely sensed data (after Anderson et al, 1976)*

The classification scheme developed as part of this research is designed for use with a range of medium resolution satellite data to be interpreted using multispectral classification techniques. Consequently, classes are defined in terms of spectral separability and minimal intraclass variability (Martin *et al*, 1988), rather than a recognition of a specific land use category, as contextual information will not be utilised to assist with the interpretation process. The characteristics of each class were described in Section 2.2.2.

Table 4.3 illustrates the land cover classification scheme developed for this study. The classification is hierarchical with classes at Level I and Level II used to accommodate the level of detail extracted from the various sources of satellite data. *Residential* and *Pine plantation* have been further subdivided at Level III in recognition of classes which are clearly identifiable on analogue displays of the data. Classification at Level III will not be attempted, however categories in Level II and Level III are designed to be integrated into corresponding classes higher in the hierarchy. These Level III classes

will be utilised to assist with signature extraction for the supervised classification and cluster allocation for unsupervised classification.

LEVEL I	LEVEL II	LEVEL III
1 Urban	11 Residential	111 Established residential
		112 Residential under construction
	12 Commercial	
2 Forest	21 Pine plantation	211 Pine plantation (dense)
		212 Pine plantation (low density)
	22 Wetland	
	23 Woodland	
3 Open	31 Grassland	
	32 Recreation	
	33 Bare ground	
4 Water	41 Water	

Table 4.3 *Land use and land cover classification system developed for the study area*

The use of land cover rather than land use classes as the classification objective is designed to provide categorisation of features based entirely upon spectral characteristics. Consequently, some categories that are identified as separate Level II classes within the USGS system are not regarded as separable in the current system. For example, the more generic *Commercial* class includes the Level II classes of 12, 13, 14 and 15 shown in Table 4.1. Similarly, due to the spectral similarity of the targets, the *Recreation* class includes irrigated recreation reserves as well as some areas of irrigated pasture associated with the Murdoch University School of Veterinary Science. Incorporation of ancillary data would be necessary to permit separation of these land covers into specific land use types.

Residential has been subdivided into *Established residential* and *Residential under construction* at Level III to acknowledge these two distinct land cover types and associated spectral variations. Cleared land (as a transitional stage between *Woodland* and *Residential under construction*) has not been separately categorised, but is included in the generic *Bare ground* class of Level II.

Similar modified USGS classification systems have been applied elsewhere (Martin *et al*, 1988; Trietz *et al*, 1992) for use with multispectral classification approaches for medium scale satellite data. Classification to Level III has been achieved for some

specific land cover classes, however acceptable classification even at Level II is often difficult to achieve. The design of an effective land cover classification scheme relies upon recognition of the target and sensor characteristics in conjunction with the interpretation approach to be applied. In this study, multispectral classification without consideration of contextual data is to be utilised, therefore the land covers defined in the classification scheme are defined according to their spectral separability, their intraclass variability and the classification approach.

4.3 Intraclass Variability

A land cover classification scheme applicable to the interpretation of remotely sensed data was discussed in Section 4.2. The main considerations in determining the relevant land cover classes to be interpreted for an area are the selection of classes that are representative of the land covers of the area, and selection of those classes which are also spectrally separable. The former will be satisfied where the classes are based upon a well-structured generic classification scheme, and the latter where the interclass variance is maximised and the intraclass variance is minimised. Interclass variance will be considered in this research as part of spectral signature development and classification refinement, while intraclass variance will be examined in this section using spatial autocorrelation analysis.

Intraclass variance arises from a range of sources related to the specification of class characteristics and heterogeneity inherent in individual land cover classes. Common sources of high intraclass variance include:

(i) Class development which includes more than one land cover per class with different spectral characteristics;

(ii) Land covers which form a mosaic effect with background features, such as grasses and bare ground;

(iii) Land covers which contain an agglomeration of components with a range of spectral responses, such as residential areas;

(iv) Vegetation at different stages of phenological development and vigour resulting in variable spectral responses;

(v) Transition areas between otherwise uniform cover classes.

Land cover classes must be specified based upon aggregation of pixels with similar spectral characteristics and which also convey useful information related to the interpretive process. Cover classes in this study have been selected on this basis, however the interaction of the target with the sensor is also significant, especially considering the relationship between the spectral and spatial characteristics of the target, and the spatial resolution of the sensor.

Collins and Woodcock (1996b) consider the spatial resolution of the sensor to represent the *scale of observation*, which is defined as the *spatial interval over which observations are made*. Importantly, this needs to be considered in the context of the *characteristic scale*, which the authors consider to be *the range of resolutions over which that phenomenon may be characterised.* Additionally, Atkinson and Curran (1997) and others, consider the geostatistical concept of *support* to be important, which collectively describes *the size, geometry and orientation of the space over which measurements are made*. The term support is largely inclusive of the concept of the *scale of observation.*

The information contained within remotely sensed data may be gauged by the local variance of the image data, which is determined by the relationship between the size of the objects and the spatial resolution of the sensor (Woodcock and Strahler, 1987). The spatial resolution at which local variance is maximised changes as a function of the characteristic scale of the scene elements. Based on simulation studies, variance is maximised when the scale of observations is just less than the size of the objects in the scene, and declines as the pixel size increases (Woodcock and Strahler, 1987). For data which have the capacity to be collected with a spatial resolution much finer than the characteristic scale, as may be the case with airborne multispectral data, a spatial resolution smaller than the scale at which maximum variance occurs should be chosen (Atkinson, 1997). However, where such fine resolution data are not available, as is often the case with satellite data, an amount of averaging of intraclass variation is desirable (Davis and Simonett, 1991). Additionally, where scene variance is determined to be wavelength dependent as a result of comparing a range of bands of data, there may be more than one target affecting the reflectance and the optimum spatial resolution will depend on which object is of prime interest (Atkinson, 1997).

Scene variance analysis may be utilised to determine the optimum spatial resolution for planning data acquisition or for assessment of the spatial variation of scene radiance to assist with information extraction. Analysis of intraclass variance using spatial autocorrelation techniques will be undertaken in this research in order to determine appropriate dimensions for training sites to enable the extraction of spectral signatures for implementation of a supervised classification scheme.

4.3.1 Spatial Autocorrelation

Remotely sensed images exhibit a measurable spatial structure due to the characteristics of the targets and the sensing system employed to acquire the data. Targets from any class in a scene tend to be spectrally homogenous and separable from other objects and the background. Consequently, adjacent objects are spectrally similar, and this *spatial autocorrelation* is dependent upon the size, spacing and shape of the objects in the scene (Jupp *et al*, 1988). Campbell (1981) indicates the design and operation of the Landsat MSS may also contribute to spatial autocorrelation due to the nature of the scan pattern and the sampling procedure whereby the IFOV of adjacent pixels overlap. Additionally, for higher spatial resolution sensors such as Landsat TM and SPOT HRV, the pixel to pixel correlation may be further increased due to the smaller spatial sampling interval relative to the size of the targets (Arai, 1992). However, as discussed in Section 2.2.4, smaller pixel dimensions often lead to increases in within class variance and diminished spatial autocorrelation.

While the principle objective of multispectral classification is to identify spectrally homogenous classes, it is recognised that variations of spectral characteristics within classes are significant even though the differences are insufficient to justify identification of a separate class. The use of contiguous pixels to represent a land cover class tends to underestimate the variance-covariance matrix derived from the spectral bands (Campbell, 1981; Congalton, 1988b; Labovitz and Masuoka, 1984). Consequently, several independent training samples may be required to adequately describe the mean and variance of each land cover class. Spatial autocorrelation techniques may be used to determine the areal extent of correlation so that extraction of independent training statistics may be optimised.

Spatial autocorrelation represents the concept that within a set of observations separated by a specific distance x (lag), observations will exhibit a different degree of similarity than will the set of observations separated from each other by the distance $x+1$. The degree of change in similarity is a measure of spatial autocorrelation, and the distance is termed the *lag* (Cliff and Ord, 1981). For objects spaced at uniform distances, changes in spatial autocorrelation represent changes in the similarity of values at the relevant lag distance. High positive spatial autocorrelation values indicate there is a close association between the samples separated by a specified lag. For raster data the value of x is usually an integer number of pixels.

Several measures of spatial autocorrelation are available, however to explore spatial dependence in observed radiance values for remotely sensed data, *Moran's I* is appropriate (Bailey and Gatrell, 1995):

$$I^{(k)} = \frac{n\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}^{(k)}(y_i - \bar{y})(y_j - \bar{y})}{2\left(\sum_{i=1}^{n}(y_i - \bar{y})^2\right)\left(\sum_{i\neq j}\sum w_{ij}\right)} \qquad (4.1)$$

Equation 4.1 determines spatial autocorrelation at lag k where $w_{ij}^{(k)}$ are the elements of the (n x n) spatial proximity matrix at spatial lag k. The value of *Moran's I* is not constrained to the range –1 to +1, however for most real datasets exact rescaling is not necessary (Bailey and Gatrell, 1995). Values of *Moran's I* approaching +1 represent strong spatial correlation between adjacent values at a specified lag, while values approaching -1 indicate adjacent values at the specified lag are dissimilar. Values at neighbouring lags are highly correlated since the correlation at larger lags is in part a function of correlations at smaller lags, and direct inferences should not be attributed to specific I values where there are multiple peaks in I at subsequent lags (Bailey and Gatrell, 1995).

4.3.2 Correlation Results

Landsat MSS, Landsat TM and SPOT HRV data for the study area were used to compute *Moran's I* for each of the land cover classes present in the area. The purpose of the analysis is to guide the determination of the minimum size of training sites in

order to extract training statistics which provide accurate representation of the variance present within each land cover type. Gong and Howarth (1992) recommend extraction of training pixels on the basis of stratified unaligned sampling in order to overcome the effect of spatial autocorrelation and to record training data which adequately estimates image class variance. In this research, contiguous pixels defined by irregular polygons will be defined as training samples, the minimum size of which will be determined through spatial autocorrelation.

Atkinson (1997) found that where spatial variation was wavelength dependent, a mixture of features on the ground was generally the cause of variation. Consequently, only one band of data from each sensor will be analysed (Table 4.4), as this will be sufficient to guide the correlation analysis based upon spatial variation. For maximum separation of the urban and near-urban targets characteristic of the study area, data from the near infrared region of the spectrum has been used.

Sensor	Band	Data interval
Landsat MSS	4 (0.80 - 1.10 μm)	60 m
Landsat TM	4 (0.76 - 0.90 μm)	30 m
SPOT HRV	3 (0.79 - 0.89 μm)	20 m

Table 4.4 *Sensor bands utilised for spatial autocorrelation analysis*

Land cover boundaries identified from aerial photographs of the study area were used to mask the satellite data so that the values of *Moran's I* were obtained from known areas of each land cover. The satellite data utilised in the correlation analysis were rectified using the same parameters applied to the main dataset (Chapter 3) and resampled to a pixel dimension similar in size to the original data recorded by the satellite as shown in Table 4.4. Nearest neighbour resampling was applied and no other image enhancements were undertaken.

IDRISI utilises a modified implementation of *Moran's I* for lags beyond the first spatial lag. Equation 4.1 is used to compute *Moran's I* for all lags, however for second and higher lags the input matrix must be thinned by the appropriate fraction to expose the data at the relevant spatial lag. This process results in a substantially reduced dataset for

higher spatial lags and may produce unstable *I* values where the initial sample of radiance values is small. Land cover classes with small pixel numbers, such as *Water*, have therefore been excluded from the autocorrelation analysis. The implementation of *Moran's I* used in this study applies the *king's* (or *queen's*) case, where all adjacent values including diagonals are considered. Analysis of spatial autocorrelation may alternatively be performed using the *rook's* case, but this eliminates diagonal cells from the computation and further reduces the quantity of cells evaluated.

Interpretation of spatial autocorrelation within the multisensor radiance data is accomplished through construction of correlograms where the value of *Moran's I* is plotted against the corresponding spatial lag. Analysis from the first through to the sixth lag will be undertaken for each data type representing a maximum distance of 360 m and 180 m for the Landsat MSS and Landsat TM data respectively, and 120 m for the SPOT HRV data. The practical maximum analysis distance is based upon the Landsat TM and SPOT HRV data, and the approximate size of each area of the land covers under analysis. It is anticipated that most significant changes relevant to the determination of training statistics are likely to occur within these distances. The maximum lag distance for the Landsat MSS data is well beyond the size of many contiguous areas of land covers in the study area and consideration of spatial lags beyond the third may not be relevant.

Spatial autocorrelation values for spatial lags of land cover classes for all resolutions of satellite data with less than 30 cells were not computed due to the relatively small sample size. Norcliffe (1977) estimates that for spatial autocorrelation analysis samples around 20 are considered *moderately small* and samples of 40 are deemed *moderately large*, therefore 30 was considered a reasonable cut-off level for this analysis.

Comparison of corresponding spatial lag values between sensors is not relevant because at a particular lag each sensor is measuring over a different physical distance, and variations in lag values between sensors will therefore not be directly comparable. Evaluation of the spatial autocorrelation of each data set will be undertaken independently with the objective of identifying the minimum areal extent of training samples for spectral signature derivation for each resolution of data.

Results of the spatial autocorrelation analysis for the satellite data are presented in the correlograms shown in Figures 4.1, 4.2 and 4.3.

Each correlogram demonstrates the degree of similarity of the radiance value of a pixel with its neighbours at varying pixel distances corresponding to the relevant lag values. For each type of satellite data the actual ground distance corresponding to the lag value is obtained from the product of the data interval shown in Table 4.4 and the indicated lag.

Table 4.5 summarises the results of significance testing of the spatial autocorrelation of land covers calculated using *Moran's I*. The significance of the spatial autocorrelation computed from the satellite data used in this study for each land cover category was tested against the null hypothesis that no spatial autocorrelation exists at the $\alpha = 0.01$ level of significance. A one-tailed test was applied with z = 2.33.

Testing was only undertaken on those data included in Figures 4.1, 4.2 and 4.3, and therefore excluded any comparisons where there were less than 30 pixels available to compute *Moran's I*. For each category, significance testing was not considered beyond the first lag value where no significant spatial autocorrelation was observed even if data at subsequent lag values were significant. As previously discussed, spatial autocorrelation was only considered to a maximum of six lags and only when n > 30. Consequently, for some land covers all recorded spatial lag values were significant even though all six lags were not assessed. While the algorithm used by IDRISI represents a limitation of the implementation of *Moran's I*, these data provide relevant details of the spatial patterns of radiance values in the near infrared portion of the spectrum.

4.3.3 Correlation Analysis

4.3.3.1 Landsat MSS

Figure 4.1 demonstrates the spatial variability in radiance values for the selected land covers in Landsat MSS band 4. All land covers show a gradual decrease in *I* with higher spatial lags, except for *Bare ground*. The *Bare ground* category shows a very high spatial correlation for the first lag, but dips for the second and actually increases again for the third lag. Despite the observed variability, all three lag values for this

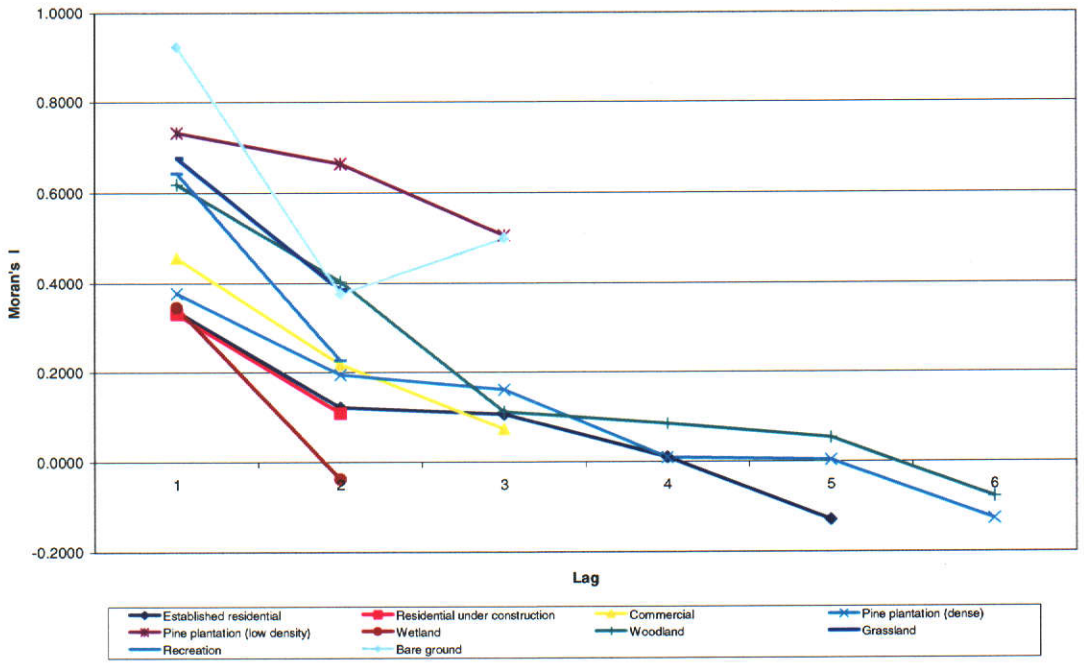


Figure 4.1 *Correlogram showing spatial variations in radiance values for Landsat MSS data*

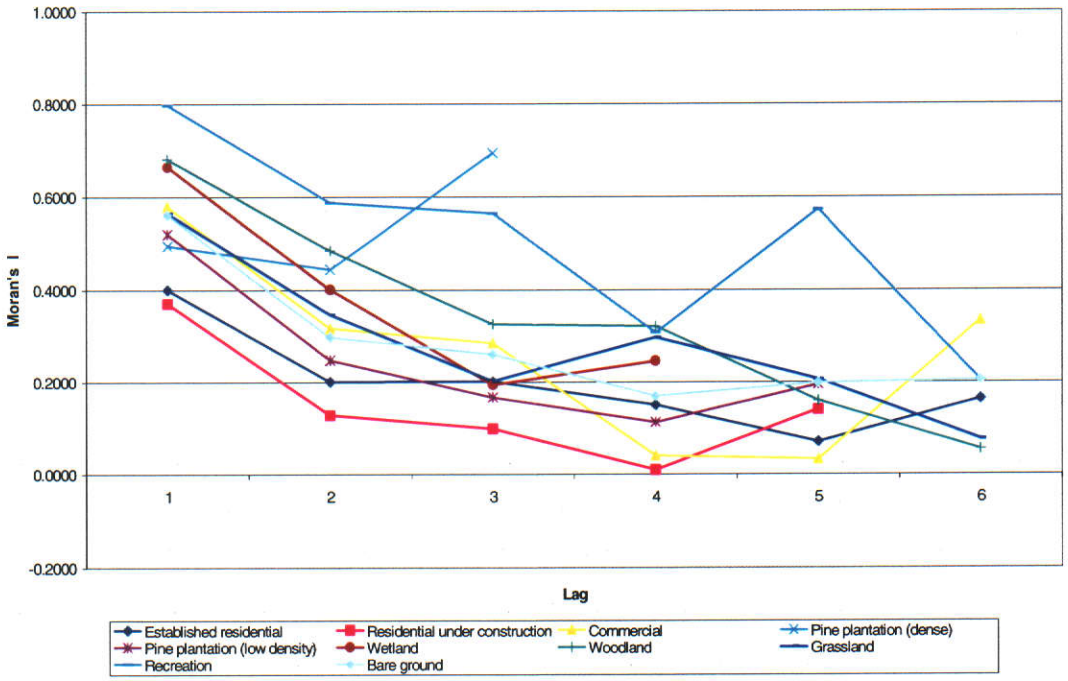


Figure 4.2 *Correlogram showing spatial variations in radiance values for Landsat TM data*

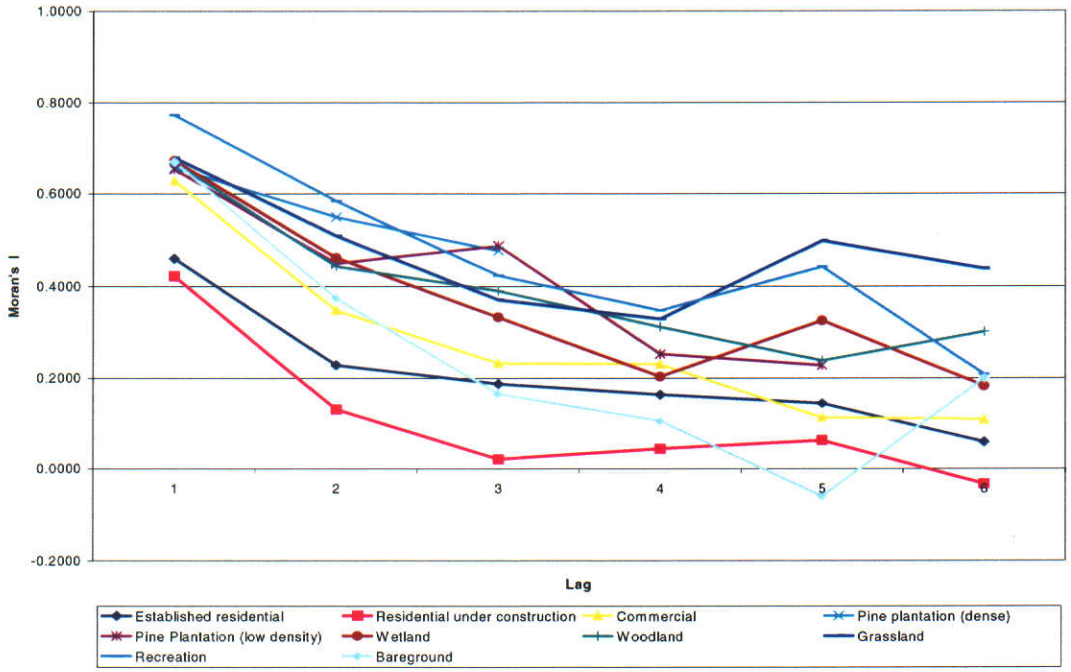


Figure 4.3 *Correlogram showing spatial variations in radiance values for SPOT HRV data*

Land Cover Category	Spatial Lag (significant values)		
	Landsat MSS	**Landsat TM**	**SPOT HRV**
Established residential	2	4	6*
Residential under construction	1	2	2
Commercial	2	3	5
Pine plantation (dense)	3	3*	3*
Pine plantation (low density)	3*	3	5*
Wetland	1	3	5
Woodland	2	4	6*
Grassland	2*	5	6*
Recreation	1	5	6*
Bare ground	3*	5	3

* *all spatial lags computed tested as significant*

Table 4.5 *Assessment of the significance of spatial autocorrelation for each land cover category in the study area*

category are considered significant. This effect may result from the presence of some low vegetation in parts of the area identified as belonging to this class, or due to the relatively narrow and disjointed nature of the areas within this category, and the inherent inconsistency of the resulting radiance values. Bailey and Gatrell (1995) state that values at neighbouring lags of a correlogram are highly correlated and caution against attributing particular significance to peaks in a correlogram at subsequent lags (third) if there are also peaks at smaller lags (first). This situation applies for the *Bare ground* category and further consideration of the peak at the third lag is not warranted.

The *Pine plantation (dense)* and *Woodland* categories occupy relatively large areas within the scene and demonstrate consistently decreasing spatial correlation. The value of *I* for *Woodland* decreases from a relatively high level (0.61) and *Pine plantation (dense)* from a relatively low level (0.38), with both classes showing equivalent levels by the third lag and low levels of correlation at the sixth lag (-0.1), indicating no spatial correlation is present. The variability within the *Woodland* category means only *I* values to the second lag are significant, while *Pine plantation (dense)* maintains significance to the third spatial lag, thus confirming the higher intraclass variability of the *Woodland* category compared to the *Pine plantation (dense)* category. *Woodland* areas characteristically comprise a mixture of trees, understorey and soil which contribute to a variable reflectance response, whereas *Pine plantation (dense)* areas provide a more consistent view to the sensor over relatively large tracts of land considering the spatial extent of the Landsat MSS pixels.

Pine plantation (low density) shows a relatively high spatial correlation over the first three spatial lags, indicating significant levels of similarity in spectral characteristics. While low density pine plantation areas comprise a variable mixture of components including pine trees, grasses and soil, the pixel averaging effect provides consistent radiance values. The *Commercial* category demonstrates a similar pattern of correlation, however only the first two *I* values are considered significant. Additionally, the strength of the spatial relationship is much lower, confirming the suspected variability of land covers and their corresponding radiance values. Commercial areas generally include a wide range of land covers such as roads, carparks and paved areas (bitumen, concrete and brick paved), reticulated grass and landscaping, various roofing materials and soil, which lead to this lower spatial correlation.

Established residential and *Residential under construction* categories demonstrate identical correlation for the first two lags. Only the first lag is considered significant for the *Residential under construction* category and, due to the low occurrence of this land cover within the study area, no further lag values were evaluated. Lag values commence at approximately 0.35 but decreased rapidly to 0.13 by the second lag and to –0.14 at the sixth lag for *Established residential.* Only the first two lags are significant for this category. These results indicate significant variability of measured radiances within these categories due to the range of component land covers. Common to both *Established residential* and *Residential under construction* are bitumen and concrete paving, and clay, concrete and steel roof areas. *Established residential* pixel responses also include significant components of reticulated vegetation while, for *Residential under construction,* pixels also include large areas of highly reflective bare soil. Within each category these conditions lead to significant variability between pixels and result in low spatial correlation within the corresponding land cover classes.

The *Recreation* category within the study area comprises many small areas used for community recreation, as well as several larger areas of reticulated grassland used for sporting activities and for pasture related to the activities of the Murdoch University School of Veterinary Science. These areas of pasture are included in this category due to the intensive management of reticulation and fertiliser treatments that are more similar to those applied to sporting facilities than to other grassland areas within the study area, and consequently are spectrally similar to the *Recreation* category. Areas of non-reticulated pasture in the study area have been placed in the *Grassland* category because the spectral response varies according to the prevailing seasonal conditions. Two spatial lags were assessed for *Recreation* which varied from 0.66 to 0.24. Only the first spatial lag is significant and indicates there is minimal consistency of response across the category, although this may be due to the distribution of *Recreation* across the study area where it occurs as a large number of isolated patches within other land cover classes, and contamination of responses by boundary pixels may be an issue.

The *Grassland* category also occurs over limited areas. Spatial autocorrelation ranges from 0.69 to 0.38 between the first and second lags, and were the only two lags assessed. Both values are considered significant thus indicating consistency of response at least over the area covered by up to five Landsat MSS pixels (centre pixel and two

pixels each side). Areas of grassland generally comprise a mosaic of native and introduced grasses and soil, which produces a relatively uniform response when averaged by the Landsat MSS sensor.

Wetland areas occur over a limited extent within the study area and the two lag values computed vary between 0.34 and –0.05. Only the first value is considered significant, with the second value indicating clearly that no spatial correlation exists at that level for this category. The large difference between the two *I* values may in part be caused by the effect of errors in the geometric rectification of the satellite data and the registration of the reference data. With *Wetland* occupying a very limited areal extent, some pixels used in the spatial autocorrelation analysis will be boundary pixels and adjacent land covers will influence the radiance values.

4.3.3.2 Landsat TM

For the Landsat TM data all categories show a gradual decrease in spatial autocorrelation except *Pine plantation (dense).* In this case all three spatial lags assessed were significant, with the first two values around 0.50 and the third at 0.70. As was the case with the *Bare ground* category for the Landsat MSS data, the *Pine plantation (dense)* category has been reduced to a number of small disjointed areas and, combined with potential correlation within the correlogram, no particular significance may be attributed to this pattern. Due to the limited extent of this category, training samples will necessarily be of limited areal extent.

Most of the spatial autocorrelation information derived from the Landsat TM data from the remaining land cover categories demonstrates a consistent pattern of results with a gradual decrease in spatial autocorrelation towards second and subsequent lags.

Established residential shows moderate spatial autocorrelation for the first lag and relatively low values for all subsequent lags. Despite these low values, spatial correlation tested significant through to the fourth lag due to the low variance of the computed *I* values, thus indicating underlying consistency of radiance values even though not at a strong level.

The *Residential under construction* category indicates significance only to the second lag with *I* quickly decreasing to values near zero. The radiometric variability of this class is characteristic of the variety of land covers present in the category and rapid change which occurs during development and construction, and is reflected in the low level of spatial consistency.

Moran's I was significant for the *Commercial* category to the third lag ($I = 0.29$), but by the fourth lag no autocorrelation was evident ($I = 0$). The smaller pixel size of the Landsat TM sensor compared to the Landsat MSS data results in reduced averaging of the target, which is especially important in heterogenous parts of the scene, such as commercial areas where sharp changes in spatial autocorrelation may be observed.

Spatial autocorrelation for *Pine plantation (low density)* tested significant only to the third spatial lag which is due to variations in the density of pine trees in the area. While no soil or understorey is visible for the *Pine plantation (dense)* category, *Pine plantation (low density)* incorporates a range of forest densities where the spacing between trees varies significantly, and understorey and soil are clearly exposed to the sensor. Consequently, the consistency of responses for low density pine plantation may vary as indicated by the correlation results for this category.

Wetland areas demonstrate significant spatial autocorrelation through to the third lag. The fourth lag is not significant and may arise from the variation of vegetation distribution according to the moisture and waterlogging regime. Moisture conditions tend to stratify the vegetation in wetland areas and effect the resultant radiometric responses.

For Landsat MSS data the *Woodland* category was previously compared with *Pine plantation (dense)*, and considered as a class demonstrating relatively high spatial variability. While the reverse is true for data recorded by the Landsat TM sensor, the extent of dense pine plantation is extremely limited in this image and spatial autocorrelation results are not as reliable. Comparing land cover classes recorded by the Landsat TM sensor, the *Woodland* category shows significant spatial autocorrelation at the fourth lag, which is less than the extent of significant autocorrelation indicated for *Grassland, Recreation* or *Bare ground* (fifth lag). These classes all represent relatively

uniform land covers and confirm that comparatively moderate variability exists in the *Woodland* category when recorded by the Landsat TM sensor.

The physical characteristics of the *Grassland, Recreation* and *Bare ground* categories were discussed previously in Section 4.3.3.1 when interpreting results from the Landsat MSS sensor. In the current analysis all three categories exhibit significant levels of spatial autocorrelation at the fifth lag. This indicates that the radiance values for each of these land cover categories exhibit relatively uniform responses over a comparatively large area. The area of each of these land covers has expanded considerably for the Landsat TM dataset, indicating the areal extent of the corresponding land covers in the Landsat MSS data may have been a limiting factor in determining reliable spatial autocorrelation characteristics.

4.3.3.3 SPOT HRV

Analysis of the SPOT HRV data for spatial autocorrelation shows that with four of the ten classes analysed, all six spatial lags are significant, thus indicating that spatial autocorrelation may extend beyond the limits of analysis undertaken in this study. This contrasts with earlier analyses when classes in which all spatial lags were significant occurred when a maximum of three lags was computed.

All six spatial lags exhibit significant levels of spatial autocorrelation for the *Established residential* and *Woodland* categories. This contrasts with the results from the Landsat MSS and Landsat TM data where these categories had been identified as providing a potentially moderate to high level of spatial variation. With the higher spatial resolution of the SPOT HRV sensor variability is expected to increase, however the pixel size of 20 m is still significantly larger than the spatial frequency of the image components and therefore permits a degree of radiometric averaging of the targets.

Grassland and *Recreation* also tested significant for all six spatial lags, however these land covers are relatively uniform and a high level of spatial correlation is expected. This also occurred for the Landsat TM data, although for the Landsat MSS data the areal extent of each was too limited and the pixels too large for a conclusive result to be obtained.

Residential under construction, Pine plantation (dense) and *Bare ground* are the only categories which have not demonstrated an increase in the extent of spatial autocorrelation (in terms of lag values as compared to distance), when comparing the analysis of Landsat TM and SPOT HRV data with the analysis of Landsat MSS data. *Pine plantation (dense)* and *Bare ground* are both of limited areal extent and are formed by a number of disjointed areas within the SPOT HRV dataset, which leads to increased variability. The degree of spatial autocorrelation measured by the number of spatial lags has not increased for the *Pine plantation (dense)* category throughout the analysis, despite the very uniform nature of the target. It is apparent the decrease in the spatial distribution of this category influences this outcome. *Bare ground* has exhibited variable responses for each of the datasets with the maximum spatial autocorrelation assessed for the Landsat TM data. The total area of this category has not shown great variation during the period under consideration, and the range of spatial autocorrelation may be due to undetected changes in surface conditions, perhaps due to the balance between seasonal herbaceous ground covers and areas of bare soil.

The consistent low level of spatial autocorrelation for the *Residential under construction* category highlights the inherent spatial variability of radiometric responses for this land cover. Results shown in Figure 4.3 indicate that beyond the third spatial lag no spatial autocorrelation is detectable. The category has moderate areal extent and is found in well-conditioned clusters, which provide conditions for spatial autocorrelation to exist. The linear extent of spatial autocorrelation has decreased from 60 m for the Landsat MSS data to 40 m for the SPOT HRV data, indicating the existence of increased variability within this category, which is enhanced by the finer spatial resolution of the sensor. This high level of variability may also lead to significant confusion during the classification process.

Commercial, Pine plantation (low density) and *Wetland* areas have gradually increased from low spatial autocorrelation for the Landsat MSS data to significant correlation at the fifth lag for the SPOT HRV sensor. *Pine plantation (low density)* and *Wetland* categories are relatively uniform land covers, but are of limited areal extent. The increase in spatial autocorrelation (by lag) is consistent with the finer spatial resolution of the sensor. As a result of the radiometric uniformity of the targets, spatial autocorrelation has not been diminished by the increased sensitivity to spatial variation

provided by the higher spatial resolution of the SPOT HRV sensor or by the limited area occupied by each land cover. The *Commercial* category extends over a relatively large area, but the variation in target characteristics for this class has resulted in significant spatial autocorrelation being limited to the fifth spatial lag.

4.3.4 Multiscale Effects

Evaluation of the autocorrelation analysis in Section 4.3.3 relied upon detailed consideration of the responses measured by each sensor for the range of land covers under consideration. Intraclass characteristics are the main focus of the analysis because assessment of spatial autocorrelation will be used to determine the minimum areal extent of training sites during the supervised classification. Comparison of spatial autocorrelation effects between datasets was not undertaken, however for a complete understanding of the outcome of the spatial autocorrelation analysis, the following factors are important:

(i) *Spatial resolution.* Spatial resolution varies from 79 x 56 m for the Landsat MSS (resampled to 60 x 60 m for the analysis) to 20 x 20 m for the SPOT HRV sensor, which produces two opposing effects with respect to spatial autocorrelation. Firstly, a large spatial resolution provides for greater integration of surface reflectance to a single value and consequently reduced variance in the image. Secondly, coarser spatial resolution increases the probability that adjacent pixels will comprise targets with significantly different radiometric properties, thereby increasing the variance within the image. These factors have mutually opposite effects, however no attempt has been made to quantify the influence on the final results.

(ii) *Sensor quantisation levels.* Landsat MSS data are quantised to 64 levels, whereas Landsat TM and SPOT HRV data are quantised to 256 levels. The effect is to reduce the total variance in the Landsat MSS data and therefore increase the apparent similarity between pixels. Reduced variance has the effect of increasing the spatial autocorrelation of the data, which must be considered in conjunction with the effects of spatial resolution discussed above. No overall analysis has been made of these combined effects, however inspection of Table 4.5 shows that in most cases the areal extent (in ground units) of statistically significant spatial autocorrelation is similar across the three datasets. *Pine plantation (dense)* and *Bare ground* are exceptions, however in both cases the

area occupied by each in the finer resolution Landsat TM and SPOT HRV data has significantly reduced compared to the Landsat MSS data. These spatial autocorrelation results derived from smaller data samples are considered less reliable.

(iii) *Support.* The concept of support and its interaction with sensor spatial resolution were discussed in Section 4.3. Support relates not to the overall integration of scene components as in (i) above, but to the comparative magnitude of the individual scene elements and the sampling interval of the sensor. Woodcock and Strahler (1987) indicate image variance reaches a maximum when pixels are just less than the size of the sensor spatial resolution. Ideally, the sensor spatial resolution should be selected such that it is smaller than the observation scale (support) at which maximum variance occurs.

Considering the nature of the targets within the study area, even the finest resolution data (SPOT HRV) has a spatial resolution larger than most scene components, which permits an amount of averaging. Exceptions may be large commercial establishments such as department stores, factories and warehouses, however they represent a minority of the targets under consideration. Therefore, in all situations considered in this study the spatial resolution is larger than the *support.*

Understanding spatial autocorrelation in multiscale satellite data relies upon knowledge of the scene-image relationships of target size and sensor spatial resolution. Within this research relative amounts of spatial autocorrelation for multiscale satellite data will not be utilised to enhance the classification process. However, knowledge of multiscale influences on spatial autocorrelation are important for the understanding of intraclass spatial correlation and for guidance in the determination of training site selection and size for implementation of supervised classification procedures. Values in Table 4.5 will be used to determine the minimum size of training samples for signature development prior to supervised classification. Training samples for each land cover class must be larger than a radius defined by the product of the maximum spatial lag defined in Table 4.5 and the pixel size defined in Table 4.4 which was used for spatial autocorrelation analysis. This applies independently for each scale of remotely sensed data and is designed to ensure training samples incorporate an appropriate degree of intraclass variance.

4.4 Information Extraction

Processing of remotely sensed data generally involves at least two important phases, image enhancement and information extraction. Image enhancement performs those functions that improve the quality or appearance of the data and correct for any defects which arise during data acquisition and preprocessing. Image enhancement includes such functions as atmospheric corrections, image destriping, contrast stretching, density slicing, principal components analysis and spatial filtering. Information extraction, however, involves the identification and labelling of objects and features with the objective of producing an information product. The following section deals with the process of information extraction applied to identification of land covers within the study area using supervised and unsupervised classification techniques applied at two levels of interpretation.

4.4.1 Unsupervised Classification

Unsupervised classification procedures utilise spectral similarities within land cover classes to allow partitioning of multispectral data into a number of spectrally separable clusters. Analyst intervention is generally restricted to *a posteriori* labelling of each cluster with an appropriate land cover class. An advantage of the unsupervised classification approach is that it establishes the spatial distribution of spectrally similar pixels and thereby identifies classes based entirely upon spectral separability. The main disadvantages relate to the requirement to define the number of clusters *a priori* and the effort required to identify each of the clusters formed by the analysis and attach appropriate labels (Deer, 1998). Many approaches to unsupervised classification are available, however most rely upon an iterative process of clustering designed to maximise the intercluster spectral separation, yet minimise the intracluster variance. The ISODATA algorithm is one such approach that will be utilised in this research and is described in the following section.

4.4.1.1 Theoretical Approach

The Iterative Self-Organising Data Analysis Technique (ISODATA) implemented within IDRISI is based upon aggregation of pixels into clusters and refinement of cluster membership on the basis of minimising the sum of the squared error of radiances of the pixels (Eastman, 1997). The ISODATA algorithm is one of many clustering

methods available, which include statistical non-supervised learning techniques, ranking, intrinsic dimensionality determination and graph theory (Jain, 1989).

The parameters on which the general ISODATA routine operates are as follows (Jensen, 1996):

(i) *Maximum number of clusters.* Usually defined by the analyst, however the final number located may be less than this value;

(ii) *Maximum percentage of pixels.* The quantity of pixels whose class assignment remains unchanged between iterations. When this value is reached the algorithm terminates;

(iii) *Maximum number of iterations.* Iterations during which the pixels are classified and the cluster means are recalculated;

(iv) *Minimum pixels in a cluster.* If a cluster contains less than this minimum percentage it is deleted and the pixels are assigned to other clusters. This also effects whether a cluster is to be split;

(v) *Maximum standard deviation.* When the standard deviation exceeds the maximum value and there are at least twice the minimum number of pixels in a cluster, the cluster is split into two clusters;

(vi) *Minimum distance between cluster means.* Clusters with a weighted distance less than this value are merged.

In the general algorithm the maximum number of clusters (K) is known. For an iteration (n) the assignment of any pixel (x) to the region (R_k) of the k^{th} cluster is given by (Jain, 1989):

$$x_i \in R_k \quad \Leftrightarrow \quad d(x_i, \mu_k(n)) = \min_{j=1,2,\ldots,K} \left| d(x_i, \mu_j(n)) \right| \tag{4.2}$$

where

$\mu_k(n)$ = *centre of the* k^{th} *cluster at the* n^{th} *iteration*

$d(x,y)$ = *distance measure between the pixel and the cluster centre.*

The cluster centres are recomputed by finding the point that minimises the distance for elements within each cluster, thus:

$$\mu_k(n+1): \qquad \sum d(x_i, \mu_k(n+1)) = \min_{y} \sum_{x_i \in R_k} d(x_i, y_i), \qquad k = 1, 2, \ldots\ldots\ldots K \qquad (4.3)$$

The procedure is repeated for each x_i, one at a time, until the clusters and their centres remain unchanged.

Within the IDRISI implementation only the number of iterations and quantity of clusters need be specified. The classification is initialised by selection of points in multispectral space that represent the location of the seed clusters. Most implementations of the basic algorithm utilise an arbitrary (systematic or random) approach to cluster seed allocation (Richards, 1993). However in IDRISI, the allocation is based upon seed clusters formed as part of a separate clustering algorithm using a histogram peak technique (Eastman, 1997). Seed clusters are derived from a three band composite image, are more representative than arbitrary allocations and tend to converge to a solution more rapidly. Additionally, the distance measure used in IDRISI relies on the maximum likelihood algorithm and makes pixel assignments based upon Gaussian maximum likelihood evaluations of class membership (see Section 4.4.2.1).

Prior to the actual seeding process the clustering algorithm computes preliminary clusters and produces a histogram of cluster membership. Significant breaks in the histogram indicate major changes in the generality of the clusters and one of these breaks is chosen as the desired level of cluster development. The values used in this study are shown in Table 4.6.

Satellite data	Clusters
Landsat MSS	20
Landsat TM	27
SPOT HRV	20

Table 4.6 *Quantity of clusters applied to unsupervised classification*

4.4.1.2 Class Definition and Refinement

The major advantage of unsupervised classification is the minimal amount of analyst intervention required for the definition of classes within a multispectral dataset. The classifier, according to the class definition rules, identifies spectrally separable classes, however it remains for the analyst to relate these spectral groups to their respective information classes. According to the classification scheme developed for this area, four land cover classes are required at Level I and up to nine classes at Level II (Table 4.3). Typically, there are many more spectral groupings identified by the unsupervised classification routine than there are land cover classes and it remains for *a posteriori* analysis to assign these spectral groups to relevant land cover classes (Jensen, 1996).

Unsupervised classification was applied to an area of approximately 5 200 ha including the study area and surrounding regions. This larger area was selected in order that, even for small clusters, sufficient pixels were available for reliable cluster formation. The surrounding areas included a similar range of land covers to those within the study area, but not necessarily in the same proportion. Following cluster formation, all subsequent analyses were undertaken within the specific study area of approximately 1 850 ha.

Spectral classes were identified within the ISODATA routine by first applying a standard clustering algorithm to a three-band colour composite. The colour composites included bands that encompassed the basic image dimensions for each data type, *viz* Landsat MSS bands 1, 2 and 4, Landsat TM bands 3, 4 and 5, and SPOT HRV bands 1, 2 and 3. The resulting frequency histogram of clusters for each image provides a guide to the degree of significance of each cluster, with natural breaks in the histogram indicating major changes in the generality of clusters within the image. It was considered that clusters beyond the limits defined in Table 4.6 did not provide any additional discrimination within the image. The same number of clusters were utilised for both Level I and Level II mapping because clustering is based primarily upon spectral separability, with the allocation to land cover classes comprising a subsequent process.

Due to the complexity of land cover patterns at the 1986 and 1991 epochs, manual allocation of clusters to land cover classes proved to be a difficult task. An alternative approach was employed for determining cluster labels that involved sampling of the

study area and allocation of clusters to land cover classes on the basis of the land cover frequency occurring in each cluster. Allocation of clusters was based on sampling of the detailed reference data that had been prepared from aerial photographs (see Section 2.4). Specification of the sampling design was based upon *a priori* knowledge of the extent of land covers derived from the reference data and stratified random sampling methodologies detailed by Fitzpatrick-Lins (1981). This methodology is further discussed in Section 5.3 and resulted in the derivation of stratified random samples as indicated in Table 4.7.

Date	Level I	Level II
1972	717	1451
1986	683	1428
1991	708	1426

Table 4.7 *Number of pixels used for stratified random samples derived for crosstabulation of clusters with land cover classes*

Crosstabulation of these samples with the derived clusters was used to guide the allocation of clusters to land cover classes for the Level I and Level II classification schemes. The number of samples is based upon the *a priori* knowledge of the land covers, with at least 75 samples per class:

$$Total\ sample = \frac{75}{Smallest\ cat./Total\ pixels} \tag{4.4}$$

Some clusters were undersampled, however these data were used only to guide the allocation process and any clusters with less than 20 samples were also compared with the aerial photographs and reference data for confirmation of allocation to a specific land cover class. This approach is straightforward for the allocation of clusters to Level I categories, although the large number of relatively small Level II categories required significant effort to verify the data with available reference information. Apparent differences in spatial location between the clusters and reference data also contributed to the uncertainty of class allocation. Results of the allocation of clusters to land cover classes are presented in Section 4.5.1.

4.4.2 Supervised Classification

Supervised classification procedures provide an opportunity for significant analyst intervention and direction of the classification process. *A priori* selection of information classes, analysis of training site statistics, and specification of sampling approaches and training site geometry are all possible using the supervised approach. The limitations of supervised classification approaches arise from the possible lack of validity of the underlying data model and that the data may not be separable into the desired classes (Deer, 1998). A wide range of classification algorithms is available which utilise a complex variety of approaches to the labelling of pixels. This study is concerned only with the application of conventional classification algorithms, and in particular the maximum likelihood decision rule. Other simple algorithms such as parallelepiped and minimum distance classification are commonly available, however the maximum likelihood classifier is the most common approach used by image processing systems (Richards, 1993).

4.4.2.1 Theoretical approach

The maximum likelihood classification algorithm assumes pixels that comprise target classes are normally distributed, and that each class may be completely described by its mean vector and covariance matrix for all bands included in the dataset (Lillesand and Kiefer, 1994). A multivariate application of the normal probability distribution function is used to model the distribution of pixels in each class, and Bayes' rule is used to optimise the allocation of pixels to the available classes (Haralick and Fu, 1983). Pixels are allocated to the class with the distribution that demonstrates the greatest probability of membership.

The multivariate application of the normal probability density function is derived from the univariate algorithm shown below (Swain, 1978):

$$p(x \mid \omega_i) = \frac{1}{(2\pi)^{1/2}\sigma_i} \exp\left[-0.5\frac{(x-\mu_i)^2}{\sigma_i^2}\right] \qquad (4.5)$$

where

$p(x \mid \omega_i)$ = *probability of a pixel at a location x being a member of class* ω_i

$\mu_i = E[x \mid \omega_i]$ = *mean value of pixels in class i*

$\sigma_i^2 = E[(x - \mu_i)^2 \mid \omega_i]$ = *variance of pixels in class i.*

Training samples are used to estimate the values of μ_i and σ_i^2 from the remote sensor data. It is important at this stage to define unimodal samples in line with the Gaussian assumption, and to ensure sufficient samples are collected for parameter estimation (Swain, 1978).

In remote sensing situations data are collected from multispectral systems which require the implementation of a multivariate probability density function as follows:

$$p(X \mid \omega_i) = \frac{1}{(2\pi)^{n/2} |\Sigma_i|^{1/2}} \exp\left[-0.5(X - U_i)^T \Sigma_i^{-1} (X - U_i)\right] \qquad (4.6)$$

where

X_i = *measurement vector containing the value of the unknown pixel i in each band*

U_i = *mean vector for class i*

Σ_i = *covariance matrix for class i.*

Equation 4.6 represents the multivariate probability density function $p(X \mid \omega_i)$ which defines the probability of pixel X being classified over n bands into a class ω_i. This function relies upon the covariance matrix being non-singular and requires at least $n+1$ training sample pixels to be evaluated.

Classification of pixel X into class ω_i occurs when the probability of belonging to this class is greater than the probability of belonging to all other classes as follows:

$$x \in \omega_i \quad \text{if} \quad p(\omega_i \mid X) > p(\omega_j \mid X) \quad \text{for all} \quad j \neq i \qquad (4.7)$$

Values of $p(\omega_i \mid X)$ are *a posteriori* probabilities and are not available, but may be estimated from the training data class probabilities - $p(X \mid \omega_i)$:

$$p(\omega_i \mid X) = \frac{p(X \mid \omega_i)\, p(\omega_i)}{p(X)} \tag{4.8}$$

where

$p(\omega_i)$ = *probability that class* ω_i *occurs in the image*

$p(X)$ = *probability of finding a pixel from any class at location X.*

The value $p(\omega_i)$ is termed the *a priori* probability and, when no other information regarding the distribution of classes is available, takes a value of 1.00 for all classes. Incorporation of prior probabilities into Equation 4.7 and removal of $p(X)$ as a common factor results in the classification rule as follows:

$$X \in \omega_i \quad \text{if} \quad p(X \mid \omega_i)p(\omega_i) > p(X \mid \omega_j)p(\omega_j) \quad \text{for all} \quad j \neq i \tag{4.9}$$

The maximum likelihood decision rule may be stated in terms of discriminant functions for X in the form:

$$\begin{aligned} g_i(X) &= \ln\left[p(X \mid \omega_i)p(X)\right] \\ &= \ln\, p(X \mid \omega_i) + \ln\, p(\omega_i) \end{aligned} \tag{4.10}$$

Therefore, substituting the discriminant functions in Equation 4.9, the maximum likelihood decision rule is stated as:

$$X \in \omega_i \quad \text{if} \quad g_i(X) > g_j(X) \quad \text{for all} \quad i \neq j \tag{4.11}$$

The multivariate probability density function defined in Equation 4.6, when operated by the natural logarithm, is stated as:

$$\ln\, p(X \mid \omega_i) = -0.5\, n \ln(2\pi) - 0.5 \ln \left|\Sigma_i\right| - 0.5(X - U_i)^T\, \Sigma_i^{-1}\, (X - U_i) \tag{4.12}$$

The constant $0.5\,n\ln(2\pi)$ may be ignored and for analyses where equal prior probabilities for all classes are assumed, Equation 4.11 is modified to provide the final form of the discriminate function for the maximum likelihood classification:

$$g_i(X) = -\ln|\Sigma_i| - (X - U_i)^T \Sigma_i^{-1} (X - U_i) \qquad (4.13)$$

Each pixel within an image will therefore be classified into one of the target classes for which training data have been defined, regardless of how small are the actual probabilities of membership for any class (Richards, 1993). Where the existence of all target classes has not been recognised, or insufficient training data are available to accurately estimate the statistics for some classes, a proportion of pixels with the lowest probability of belonging to any class may be left unclassified. Details of training statistics derived from the reference data available for the study area are discussed in the following section.

4.4.2.2 Signature Extraction

The major objective of spectral signature extraction is to summarise the multivariate spectral characteristics of each land cover class, such that when pixels are compared to the library of spectral signatures, each pixel will be allocated to the appropriate land cover class (Jensen, 1996). This approach assumes uniformity of atmospheric and target characteristics across the image, and relies upon the uniqueness of the derived spectral signatures to enable identification of targets.

The maximum likelihood classifier assumes that multivariate spectral data for each class are normally distributed and can be completely described by the mean vector and covariance matrix derived from training data (Lillesand and Kiefer, 1994). Training data which exhibit multimodal distributions may describe more than one spectral class. These training data should be further subdivided into subsidiary spectral classes and their separation in spectral space evaluated. Where discrimination is feasible, separate classes may be allocated or, if clear discrimination is not possible, the limitations of less than perfect training data must be accepted.

Spectral stratification of targets within this research has been applied to the derivation of training statistics for the *Residential* and *Pine plantation* classes. In both cases, subclasses within each global land cover class were identified. *Established residential* and *Residential under construction* were identified as separate classes, as were *Pine plantation (dense)* and *Pine plantation (low density)*. Further separation of classes was also identified for *Residential* and *Pine plantation* through the unsupervised classification process, where six spectral clusters were formed for each of these cover classes. While classification accuracies of 100 percent are rarely achieved, careful selection and refinement of training samples can provide results of an acceptable standard.

Reliable training statistics depend upon inclusion of sufficient samples to provide an accurate measure of the target mean and covariance. Swain (1978) indicates that $10n$ pixels should be used (where n is the number of spectral bands), and Richards (1993) recommends $100n$ as an appropriate number. Table 4.8 summarises the number of training pixels used for each target class within the remotely sensed data analysed for this research. According to Swain (1978) the minimum number of training pixels per class for Landsat MSS, Landsat TM and SPOT HRV data are 40, 60 and 30, respectively.

Class	Landsat MSS	Landsat TM	SPOT HRV
Established residential	66	515	1748
Residential under construction	20	168	234
Commercial	40	216	690
Pine (dense)	172	65	32
Pine (less dense)	47	113	168
Wetland	42	109	181
Woodland	145	205	433
Grassland	36	425	459
Recreation	14	205	643
Bareground	25	225	137
Water	6	21	123

Table 4.8 *Number of pixels in each class used for derivation of training statistics for maximum likelihood classification*

Acquisition of training data was constrained by the area of each target class and the spatial resolution of the sensor. With a 79 x 56 m spatial resolution, extraction of sufficient training pixels from the Landsat MSS data was impossible due to the limited area occupied by land covers such as *Residential under construction, Recreation* and *Water.* As the proportion of other land covers in the study area changed, and with the finer spatial resolution of the sensor, the influence of this factor decreased. The ability to extract sufficient training pixels for *Water* remained an issue, while the decreasing frequency of the *Pine plantation (dense)* class resulted in the minimum of training sites for this class being available in the Landsat TM and SPOT HRV data. Inability to define sufficient training pixels may have a deleterious effect on the supervised classification results for the Landsat MSS data, however the training site sampling scheme remains robust for the Landsat TM and SPOT HRV data.

The classification scheme to be applied for Level I and Level II has previously been discussed in Section 4.2. Four classes for Level I and nine classes for Level II will be utilised throughout this study. Eleven spectral classes have been identified with *Established residential* and *Residential under construction* combining to form the *Residential* class. *Pine plantation (dense)* and *Pine plantation (low density)* are combined to form the *Pine plantation* class. Both of these subgroupings have been identified as spectrally different and necessary for complete spectral description of each of the parent classes.

While the spectral structures of the data are utilised to define clusters in an unsupervised classification, and information classes are defined *a posteriori*, supervised classification relies upon *a priori* knowledge of the information classes for training site selection and spectral signature definition. Training samples must be selected from each known land cover class in order to define the information classes of interest. The following process has been adopted in this study:

(i) Identify the informational classes of interest;

(ii) Locate sample sites of the informational classes using reference data for extraction of training statistics;

(iii) Identify training samples on the imagery using the largest number of bands representing maximum variance and minimum band to band correlation in the data;

(iv) Extract training statistics from the multispectral data for each informational class.

Training sites are usually established according to a number of conflicting objectives. Firstly, they are selected to incorporate areas only containing a uniform coverage of the land cover of interest. This is rarely ideal as most land covers comprise a mixture of components, therefore it is far more useful to select areas which are representative of the land covers present. Secondly, in order for spectral signatures to be unique and to maximise separability between classes, class variance must be minimised which again suggests training samples should be collected from areas of uniform land cover. However, if representative training statistics are to be obtained, sufficiently large areas must be sampled in order to assess the natural within class variance. A measure of *sufficient size* is therefore required.

The degree of intraclass spectral variability for each land cover class has been assessed through the computation of spatial autocorrelation (see Section 4.3) as measured by *Moran's I*. Table 4.5 defines the maximum lag at which spatial autocorrelation is considered significant for a near infrared band from each of the data sources examined in this research (Table 4.4). Positive spatial autocorrelation occurs among pixels that are adjacent or close together (Gong and Howarth, 1992), and small training samples of contiguous pixels are particularly prone to this effect. Consequently all training samples were selected with one dimension (x or y) at least equivalent to the product of $2n+1$ pixels and the pixel dimension from Table 4.4, where n represents the number of pixels from Table 4.5 where *Moran's I* remained significant. The training sites need only exceed this size in one dimension so that the sample extends beyond the area of spatial autocorrelation and at least a minimal representation of sample variance is obtained. Alternative means of extracting reliable training statistics by elimination of spatial autocorrelation are to sample every n^{th} pixel or to apply filtering algorithms to the training data (Labovitz and Masuoka, 1984). Implementation of these approaches requires the value of n or the filtering parameters to be evaluated from the image dataset. Spatial autocorrelation must also be determined in this situation otherwise any sampling approach will be arbitrary in application. Whichever sampling method is utilised, taking spatial autocorrelation effects into account increases the sample variance, but results in a more relevant value for spectral signature development.

Extension of training samples over the boundaries of adjacent land covers is more prevalent when large training samples are defined or the spatial resolution of the sensor is large compared to the extent of the land cover. Training samples were defined according to the results of the spatial autocorrelation analysis detailed above, however where there was the possibility of samples overlapping adjacent cover classes samples were limited (visually) to these minimum sizes. For land cover classes that occupy relatively small or irregularly-shaped areas, training samples were defined by irregular polygons placed inside the boundaries of land cover classes that were identified from the satellite imagery or reference data. Several smaller training sites were selected in order to incorporate the class variance. These strategies were significant in developing training statistics representative of the land cover classes in the area.

Examination of the spectral independence of each land cover class was undertaken to identify the presence of any anomalies in the identification of cover classes which may result in definition of overlapping spectral classes. Tables 4.9, 4.10 and 4.11 summarise the analysis of the significance of differences between training sample mean values for all land cover classes for each image dataset. Assessment was undertaken at the 95 percent confidence level and each table indicates the number of bands where a significant difference was found to exist. The maximum possible values are Landsat MSS (4), Landsat TM (6) and SPOT HRV (3).

Table 4.9 indicates there are two combinations of land covers where a significant difference between the training sample means does not exist for any Landsat MSS bands. The similarity of *Residential under construction* (Resid con) and *Bare ground* is due to the transitional nature of the *Residential under construction* class, which at various stages of development has similar spectral characteristics to the *Bare ground* class. *Pine plantation (low density)* and *Grassland* may appear similar in all bands when the density of pine trees is sufficiently low that the background of herbaceous material, such as annual grasses, dominates the scene reflectance. Corresponding analyses for the Landsat TM and SPOT HRV data indicate a greater difference between these classes, therefore no attempt at amalgamation will be made on the basis of the Landsat MSS results.

Low differentiation of mean values is evident for *Established residential, Pine plantation (low density)* and *Bare ground* classes. For each of these classes complex mixtures of scene components, which include indeterminate mixtures of soil and herbaceous (low shrubs and annual grasses) vegetation, are present. Each class has recorded no significant difference in one or more bands with at least three other classes, however a significant difference is still demonstrated in at least one band.

Land cover	Est Res	Resid con	Comm	Pine dense	Pine low	Wet-land	Wood-land	Grass-land	Rec-reate	Bare ground	Water
Est res	0	2	4	4	4	4	4	4	3	3	4
Resid con		0	4	4	4	4	4	4	2	0	4
Comm			0	4	4	4	4	4	4	1	4
Pine dens				0	4	2	4	4	4	4	4
Pine low					0	3	2	0	4	4	4
Wetland						0	4	4	4	4	4
Woodland							0	2	4	4	4
Grassland								0	4	4	4
Recreate									0	3	4
Baregrnd										0	4
Water											0

Table 4.9 *Shaded cells indicate the number of spectrally discriminatable Landsat MSS bands for each pair of land cover classes*

Land Cover	Est Res	Resid con	Comm	Pine dense	Pine low	Wet-land	Wood-land	Grass-Land	Rec-reate	Bare ground	Water
Est res	0	6	5	6	6	6	6	3	6	6	6
Resid con		0	5	6	6	6	6	6	6	5	6
Comm			0	6	4	6	5	6	6	6	6
Pine dens				0	6	6	6	6	6	6	6
Pine low					0	6	6	6	4	6	6
Wetland						0	6	6	6	6	6
Woodland							0	6	6	6	6
Grassland								0	6	6	6
Recreate									0	6	6
Baregrnd										0	6
Water											0

Table 4.10 *Shaded cells indicate the number of spectrally discriminatable Landsat TM bands for each pair of land cover classes*

Examination of Table 4.10 indicates that significant differences exist between the mean values of all bands for most classes. However, between the *Commercial* class and the

Established residential, Residential under construction, Pine plantation (low density) and *Woodland* classes, not all bands demonstrate a significant difference. These classes include mixtures of scene components comprising manmade materials or soils, and at least some vegetation in varying proportions, which cause additional variance in training data and reduce the discriminating capabilities. Many areas of *Pine plantation (low density)* occur in conjunction with *Recreation,* which leads to low differentiation (four bands) of these classes. *Established residential* and *Grassland* are separable in only three bands in the near and middle infrared parts of the spectrum, which are responsive to vegetation type and condition.

SPOT HRV data represented in Table 4.11 show that significant differences exist between class means for all bands in most land cover classes. Comparison of *Wetland* and *Pine plantation (dense)* indicates a significant difference in only one band, which is caused by the dense nature of the vegetation and inherent similarity of reflectance characteristics. Overall, based on these analyses the SPOT HRV data indicates a high level of discrimination between class means compared to the Landsat MSS and Landsat TM datasets.

Land cover	Est Res	Resid con	Comm	Pine dens	Pine low	Wet-land	Wood-land	Grass-land	Rec-reate	Bare ground	Water
Est res	0	3	3	3	3	3	3	3	3	3	3
Resid con		0	2	3	3	3	3	3	3	2	3
Comm			0	3	3	3	3	3	3	2	3
Pine dens				0	3	1	3	3	3	3	3
Pine low					0	3	2	3	3	3	3
Wetland						0	3	3	3	3	3
Woodland							0	3	3	3	3
Grassland								0	3	3	3
Recreate									0	3	3
Baregrnd										0	3
Water											0

Table 4.11 *Shaded cells indicate the number of spectrally discriminatable SPOT HRV bands for each pair of land cover classes*

Analysis of the differences between class means provides an assessment of the independence of spectral classes and training site statistics, and indicates the separability to be expected between classes during the classification process. These comparisons are made on the basis of aggregations at the training site level. Apart from

the similarities observed between the two pairs of classes in Table 4.9, where no separation was possible in any band, all classes are separable in one or more bands. The maximum likelihood classification algorithm determines class membership of each pixel on the basis of allocation to the class demonstrating maximum membership probability computed across all bands. Consequently, individual pixels may not adhere to the exact patterns indicated above. *A posteriori* analyses of classification accuracy will be used to make this assessment in Chapter 5.

4.5 Classification Results

The process of developing a classification methodology using unsupervised and supervised approaches has been discussed in Section 4.4.1 and 4.4.2. A detailed analysis of classification accuracy will be conducted in Chapter 5. The objective of this component of the study is to present the classification results and to describe the specific considerations applied in implementing the classification algorithms for each dataset.

Variations in sensor parameters and scene characteristics during the period of this study require that both supervised and unsupervised classification approaches be investigated. Variations in the spatial, spectral and radiometric resolution of the sensors (see Section 2.3) are important due to the influence of these factors on the discrimination of spectral and radiometric characteristics of the data. The relative size, shape and distribution of the scene components have also changed throughout the period covered by the study, therefore consideration of alternative approaches to information extraction is relevant.

Rectification and resampling of the satellite data were performed as detailed in Chapter 3 in order to reduce all data to a common geometric datum. The Landsat MSS data were also preprocessed using a destriping algorithm to eliminate persistent sixth line banding. During this operation the mean and standard deviation for the entire image and also for each of the six detectors was computed. The output for each detector was then scaled to match the mean and standard deviation for the entire image. As a result, the final image demonstrates greatly reduced banding and provides more consistent radiometric responses. Apart from the elimination of banding in the Landsat MSS data, both the supervised and unsupervised classifications were performed on the raw satellite data.

Following the classification process all output files were filtered using a 3 x 3 modal filter to remove speckle and smooth the result. Only two passes were utilised in order to reduce scene variation and to minimise generalisation of detail. The increases in the Kappa Coefficient as a result of smoothing varied from 12.9 percent for Landsat TM data to 2.2 percent for SPOT at Level I. For Level II classifications the changes were in the order of 6.6 percent for Landsat TM and 2.4 percent for SPOT HRV (see Chapter 5).

Increasing diversity of land covers in the study area over the period from 1972 to 1991 increased the potential range of spectral classes required to describe the land covers present. In 1972 the study area comprised mainly large areas of spectrally homogenous dense pine plantation and some long established residential areas (Figure 2.9) while, by 1991, except for isolated stands of pine trees, most of the pine plantation had been cleared for urban development. Land covers range from cleared land through to established residential areas and associated commercial and recreational developments, which provide a relatively complex mosaic of land covers (Figure 2.11). Both supervised and unsupervised classification approaches were investigated in order to develop the highest quality land cover mapping of the area.

4.5.1 Unsupervised Classification

The groups of clusters identified through the processes described in Section 4.4.1.2 were reclassified into land cover classes at Level I and Level II, respectively. Table 4.12 indicates the number of clusters used to describe each land cover class within the study area. Within Table 4.12 some discrepancies exist for Landsat TM data between the total number of clusters required to describe classes in Level II compared to their aggregated counterparts in Level I. For example, *Residential* and *Commercial* (14 clusters) and *Grassland, Recreation* and *Bare ground* (9 clusters) at Level II, should contain the same number of clusters as the *Urban* and *Open* classes respectively, at Level I. Allocation of clusters to particular land cover classes based upon stratified random sampling is not always clear cut. In some instances two or more clusters may contain similar numbers of random samples and if a different stratified random sample is defined minor changes in cluster groupings may result.

Evaluation of the number of clusters required to describe each land cover class involves a complex relationship between the target and sensor characteristics as described in

Section 4.5. Targets that include *Residential* areas possess high spectral complexity and require the largest number of clusters for spectral description at both Level I and Level II. Most other land covers are capable of description by three or less clusters, thus indicating a degree of spectral homogeneity.

Land Cover Class		Landsat MSS	Landsat TM	SPOT HRV
Level I	*Urban*	9	12	14
	Forest	9	2	1
	Open	1	11	3
	Water	1	2	2
Level II	*Residential*	6	11	11
	Commercial	3	3	3
	Pine plantation	6	0	0
	Wetland	0	1	0
	Woodland	3	1	1
	Grassland	0	2	0
	Recreation	0	3	3
	Bare ground	1	4	0
	Water	1	2	2

Table 4.12 *Number of clusters required to describe each land cover class at Level I and Level II*

Within the Landsat MSS data, *Forest* areas (Level I) are described by nine clusters, which is necessary as a result of forest density variations and boundary effects caused by the relatively large pixel dimension of the Landsat MSS sensor. The need for 11 clusters within the Landsat TM data to describe the *Open* class (level I) does not appear to have been caused by similar influences. This class comprises moderately sized areas of grassland, bare ground and recreation land covers. Mixed boundary pixels are unlikely to be significant in increasing the number of clusters, because the areas are of regular shape and the pixels are relatively small compared to the size of the area. Therefore, variations in the spectral response of the component land covers of this class are most likely the cause of the large number of clusters required to describe the spectral characteristics.

4.5.1.1 Landsat MSS

The results of the Level I unsupervised classification of the Landsat MSS data using the ISODATA routine are shown in Figure 4.4. The four Level I classes are included, with *Forest* and *Urban* dominating the classification. Comparison with the reference data (Figure 2.9a) highlights the extensive nature of the *Forest* class. The *Open* class does not appear as extensive as expected compared to the reference data, and substantial confusion has occurred between *Open* and both the *Urban* and *Forest* classes. *Water*, while classified separately, has minimal presence within the study area.

A Level II classification of Landsat MSS data is shown in Figure 4.5. The results indicate that six of the nine classes from the reference data (Figure 2.9b) are identified. A qualitative assessment against the Level II reference data indicates reasonable correspondence. The distribution of land covers occurs in relatively large uniform areas compared to those in 1986 and 1991, therefore a classification at Level II may be successful even though the spatial resolution of the Landsat MSS sensor is not normally considered suitable for land cover mapping at this level of detail.

4.5.1.2 Landsat TM

Figure 4.6 illustrates unsupervised classification results for the Landsat TM dataset at Level I. All four Level I classes are present in Figure 4.6 with *Forest* and *Water* occupying relatively small regions. *Urban* and *Open* cover comparatively large areas, however classification noise is present due to confusion between these classes. Reference data (Figure 2.10a) indicate that *Urban* and *Open* actually occupy regions of relatively uniform land covers.

Figure 4.7 indicates that eight of the nine classes from the reference data (Figure 2.10b) are identified in the Landsat TM data at Level II. At the time of data acquisition all large stands of pine plantation had been cleared and only small isolated areas remained. Most confusion appears to have occurred between *Grassland*, *Bare ground* and *Residential*, which may occur as a result of variations in spectral responses recorded at

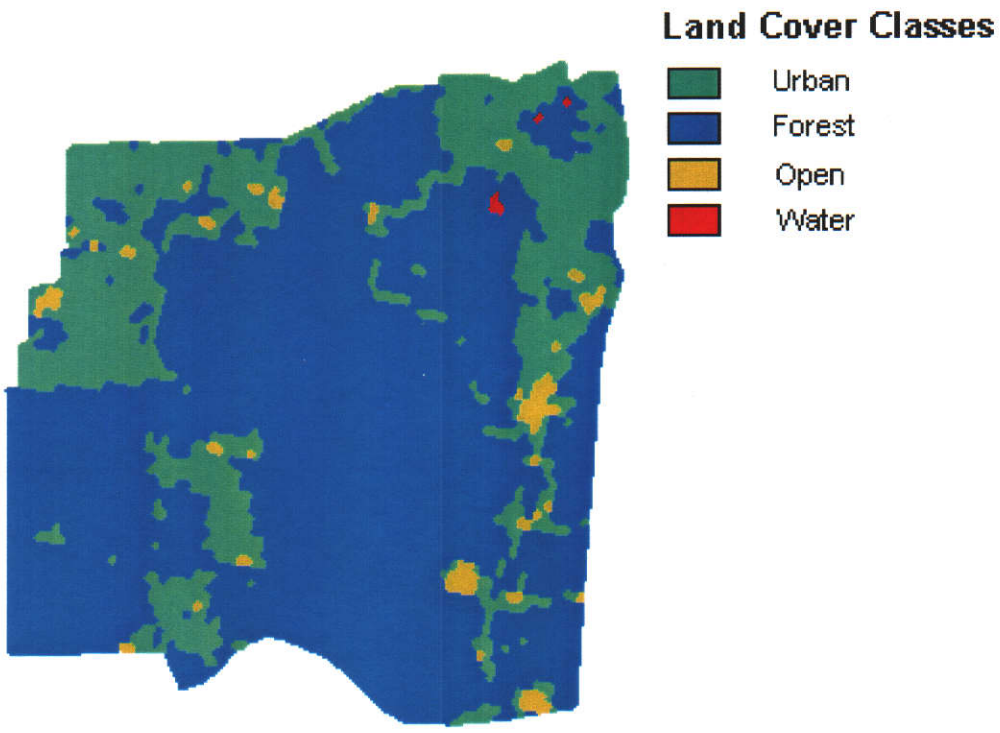


Figure 4.4 *Unsupervised classification of Landsat MSS data using Level I categories (1972) (Scale 1:50 000)*

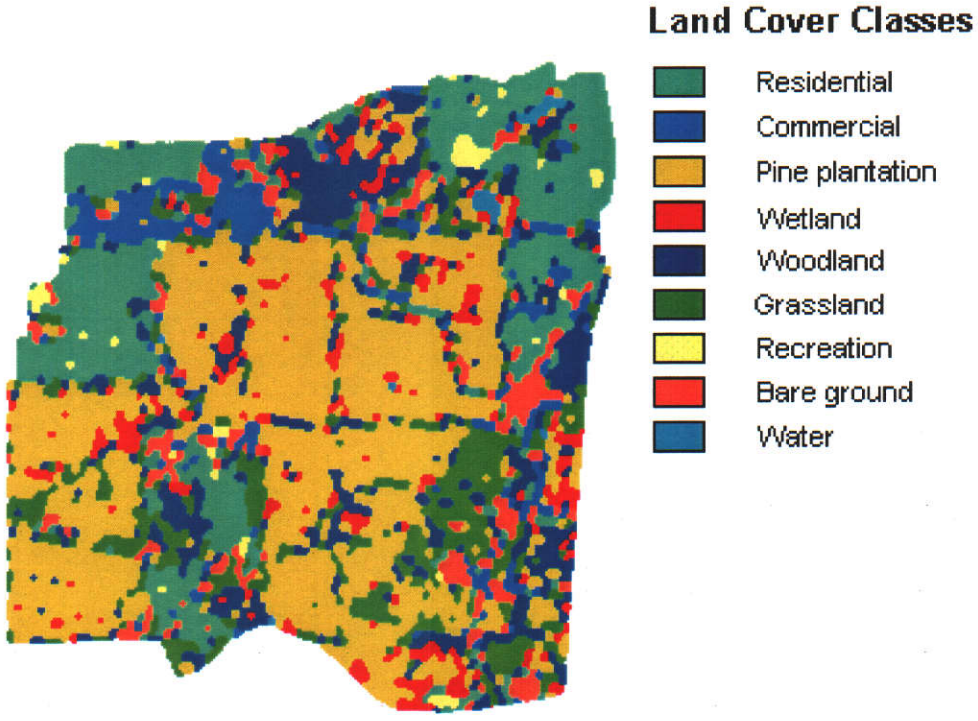


Figure 4.5 *Unsupervised classification of Landsat MSS data using Level II categories (1972) (Scale 1:50 000)*

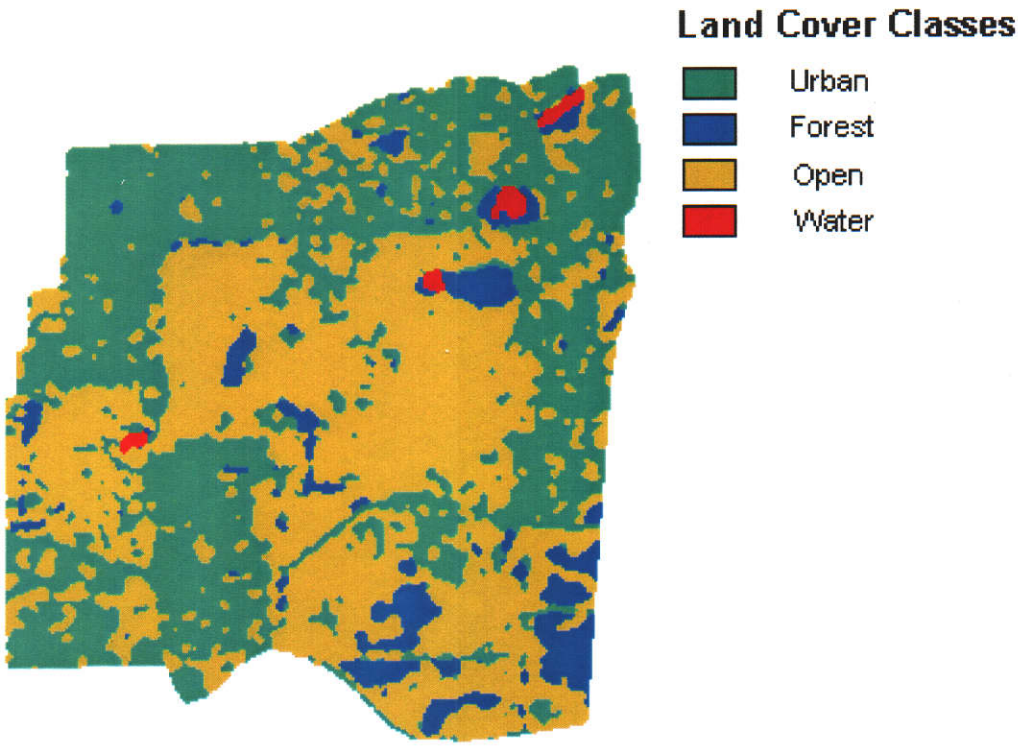


Figure 4.6 *Unsupervised classification of Landsat TM data using Level I categories (1986) (Scale 1:50 000)*

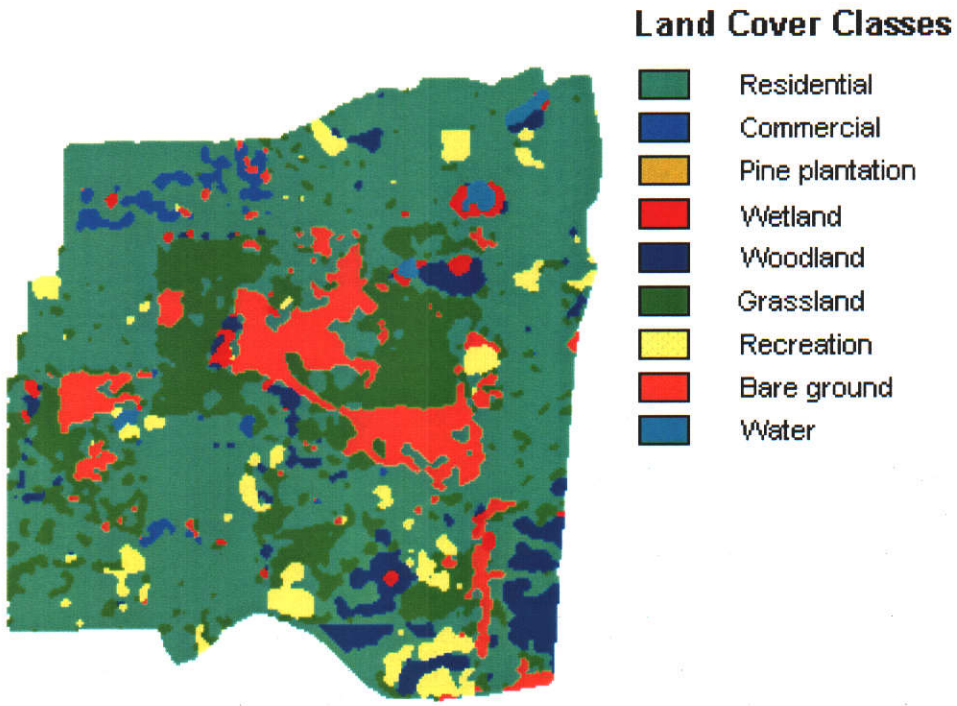


Figure 4.7 *Unsupervised classification of Landsat TM data using Level II categories (1986) (Scale 1:50 000)*

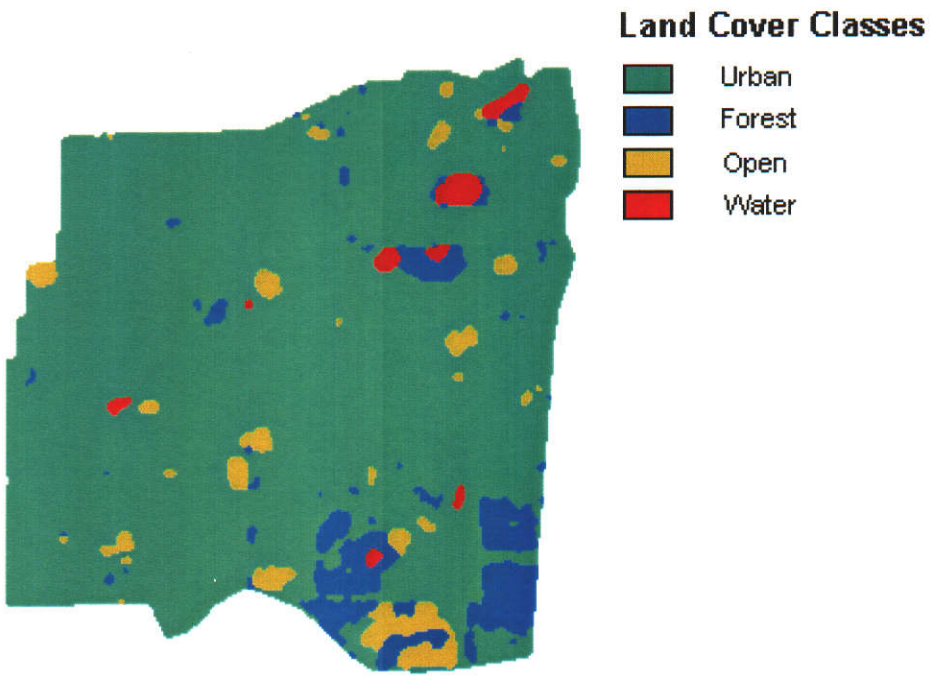


Figure 4.8 *Unsupervised classification of SPOT HRV data using Level I categories(1991) (Scale (1:50 000)*

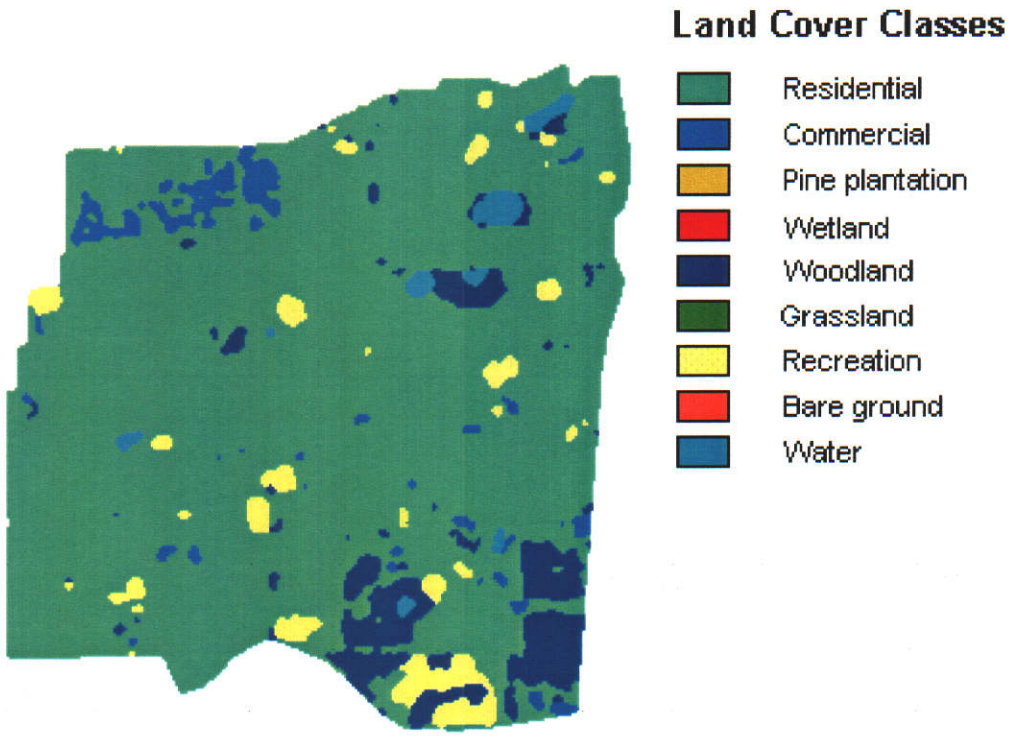


Figure 4.9 *Unsupervised classification of SPOT HRV data using Level II categories (1991) (Scale 1:50 000)*

the sensor due to the presence of transitional stages of development from broadacres (undeveloped) through to urban development.

4.5.1.3 SPOT HRV

All four Level I land cover classes are identified in the SPOT HRV data and are shown in Figure 4.8. Changes in land cover distribution since the 1972 and 1986 sensing epochs indicate almost complete conversion of the original pine plantation areas to urban land uses. At Level I this is indicated by the dominance of the *Urban* class interspersed with *Open* (mostly comprising recreational areas), and remaining areas of water, wetland, woodland and pine plantation forming lesser areas of the *Water* and *Forest* classes. Comparison of the Level I interpretation may be made with reference data in Figure 2.11a.

At Level II (Figure 4.9) only the five classes of *Residential, Commercial, Woodland, Recreation* and *Water* are identified from the SPOT HRV data. With increasing urbanisation the incidence of *Pine plantation, Grassland* and *Bareground* have decreased substantially. While the total area of *Wetland* essentially remained static, these areas appear to have been identified as *Woodland* in the unsupervised classification of the SPOT HRV data. These results may be compared to the interpretations of reference data in Figure 2.11b.

4.5.2 Supervised Classification

Spectral signatures for all eleven classes and subclasses, which were derived according to the processes described in Section 4.4.2.2, were used as input to the maximum likelihood classification. Allocation of pixels to four Level I and nine Level II classes occurred after classification and smoothing of the results with a modal filter. Classifications at each level represent an aggregation of individual classes through postclassification sorting and therefore provide for maximum discrimination based upon the unique spectral characteristics of the data. Aggregation of spectrally diverse signatures prior to classification would have decreased the scope for discrimination between classes and reduced the classification accuracy.

The area occupied by classes, changes in the patterns of land covers and variations in the spatial resolution of the sensors throughout the study have affected the ability to

define reliable spectral signatures for classes that occupied relatively small areas at the time of sensing. *Recreation* and *Water* occupy small areas and difficulties exist in defining training statistics that do not include mixed boundary pixels, especially when dealing with the larger resolution Landsat MSS pixels. Additionally, classes such as *Pine plantation* provided large uniform targets in 1972 and enabled reliable extraction of spectral signatures from the Landsat MSS data. However in 1991, *Pine plantation* occupies a very small area, and even with the higher resolution SPOT HRV data, accurate definition of targets is difficult.

Consideration must be given to these factors when examining the results of the supervised classification. Variations between datasets are expected and need to be accommodated in the change analysis process. These factors will be evaluated in Chapters 6 and 7.

4.5.2.1 Landsat MSS

Supervised classification results of the Landsat MSS data at Level I using a maximum likelihood decision rule are presented in Figure 4.10. The *Urban* and *Forest* classes are relatively uniform and the boundaries compare favourably with the reference data shown in Figure 2.9a. The *Open* class is distributed throughout each of these classes, but the nature of the category is such that it naturally occurs within any of these land covers. Two relatively large areas of *Water* are indicated, however it is likely the extreme contrast provided by the water pixels has overestimated the size of these features. Comparison with the unsupervised results in Figure 4.4 shows the supervised classification contains more classification noise, the *Water* classes are much larger, and substantially more pixels are classified as *Open*.

Figure 4.11 presents a complex arrangement of land covers resulting from a Level II supervised classification of the Landsat MSS data. All nine land cover categories were identified with *Residential, Commercial, Pine plantation* and *Grassland* being dominant. Many areas of *Woodland* appear to have been excluded and numerous areas of *Wetland* are shown throughout the study area, indicating spectral confusion of these classes. Rotation of east-west linear features caused by the flight path and sampling process of the Landsat MSS is also evident across the centre of the image, compared to the reference data in Figure 2.9b.

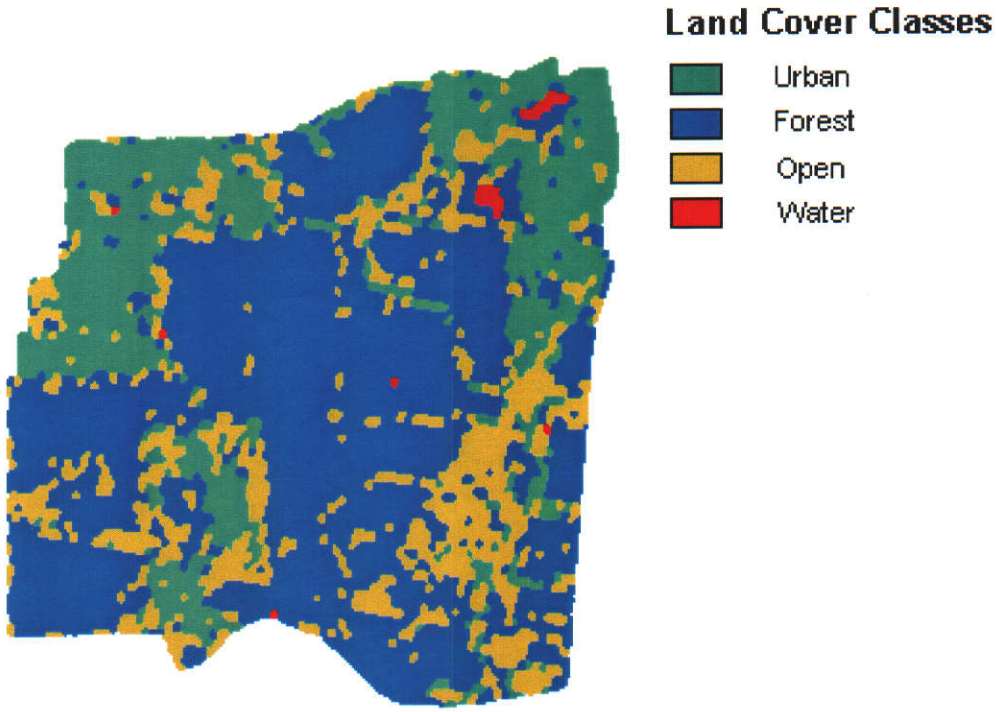


Figure 4.10 *Supervised classification of Landsat MSS data using Level I categories (1972) (Scale 1:50 000)*

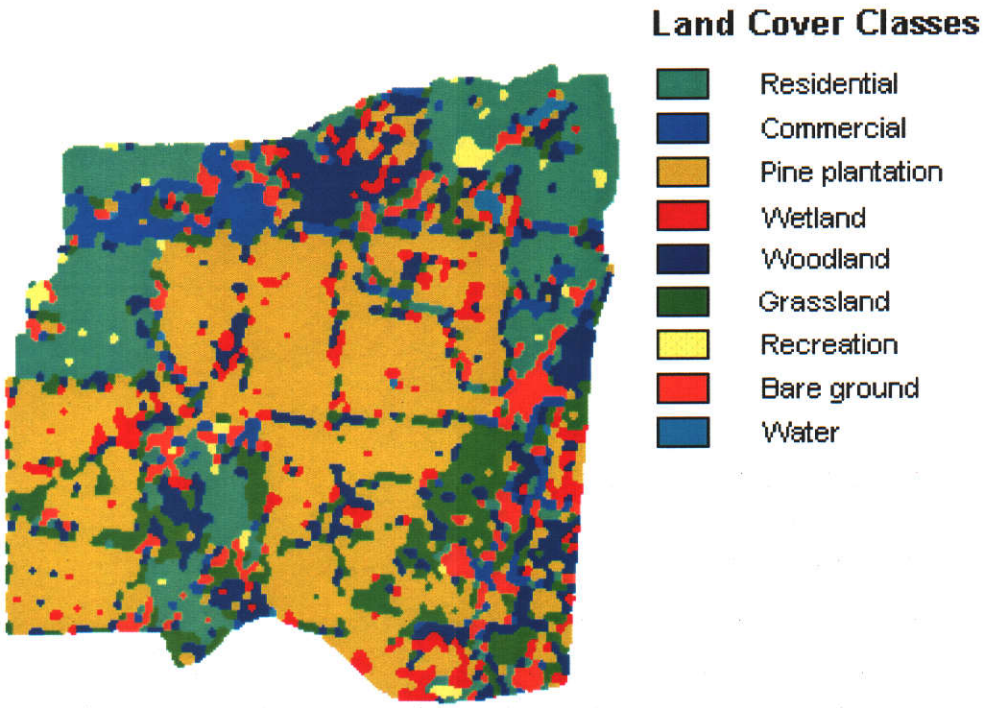


Figure 4.11 *Supervised classification of Landsat MSS data using Level II categories (1972) (Scale 1:50 000)*

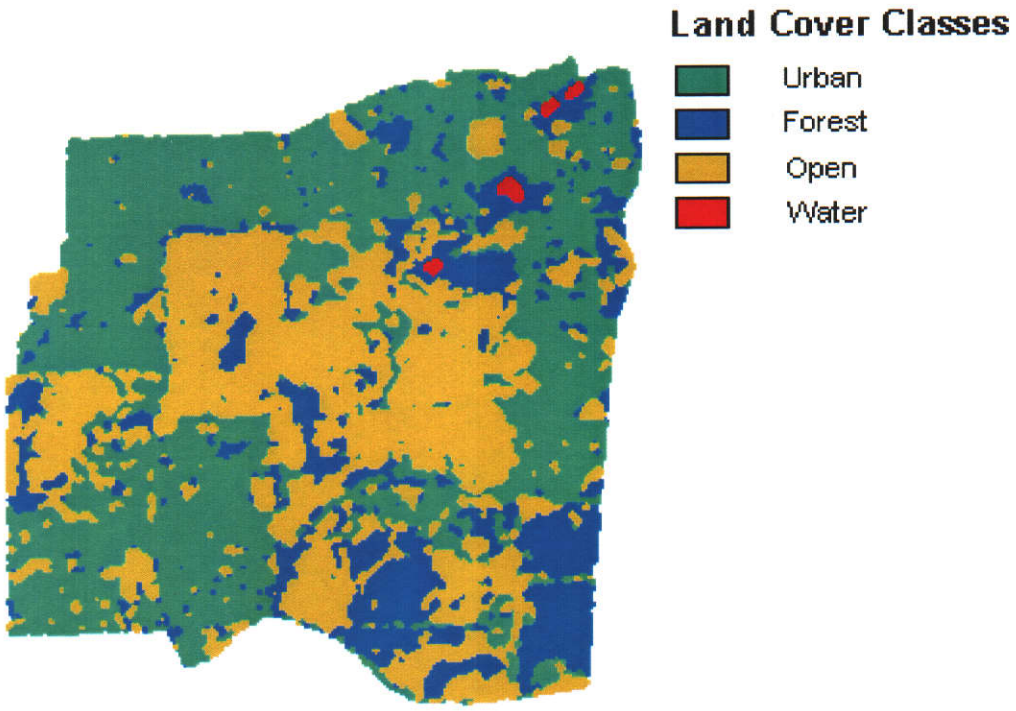


Figure 4.12 *Supervised classification of Landsat TM data using Level I categories (1986) (Scale 1:50 000)*

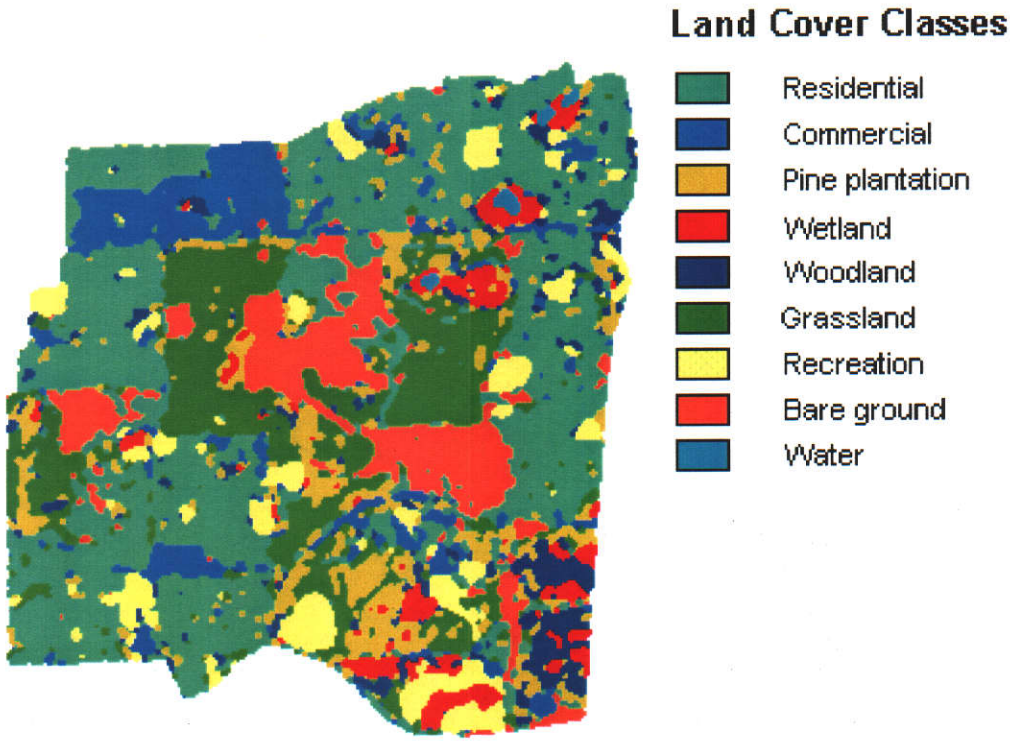


Figure 4.13 *Supervised classification of Landsat TM data using Level II categories (1986) (Scale 1:50 000)*

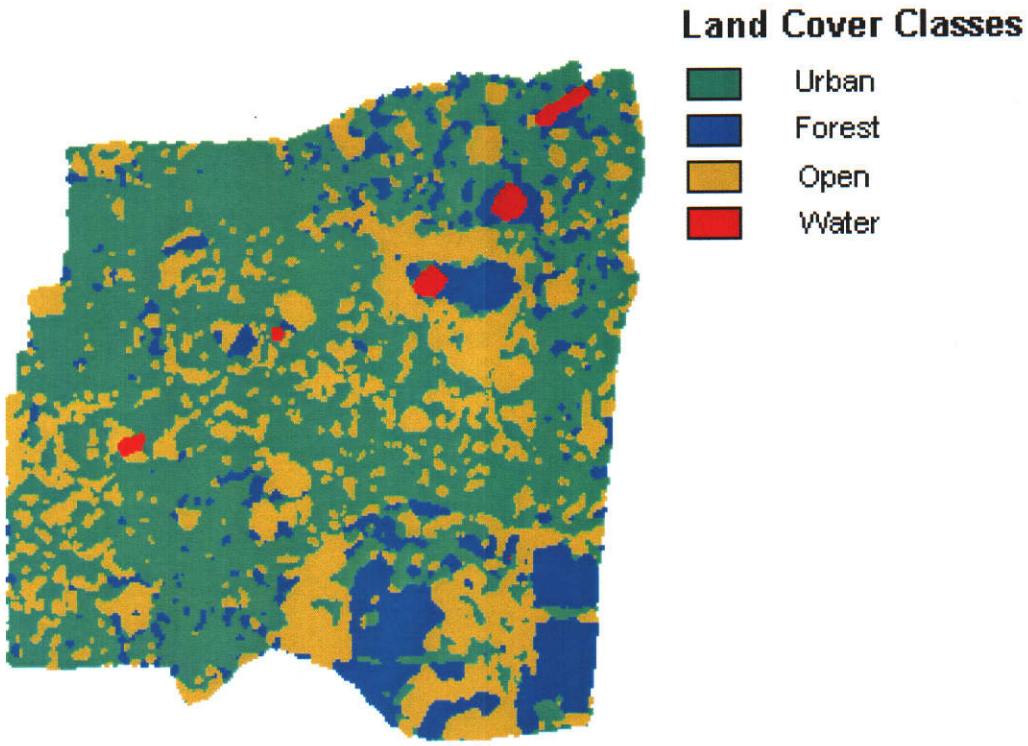


Figure 4.14 *Supervised classification of SPOT HRV data using Level I categories (1991) (Scale 1:50 000)*

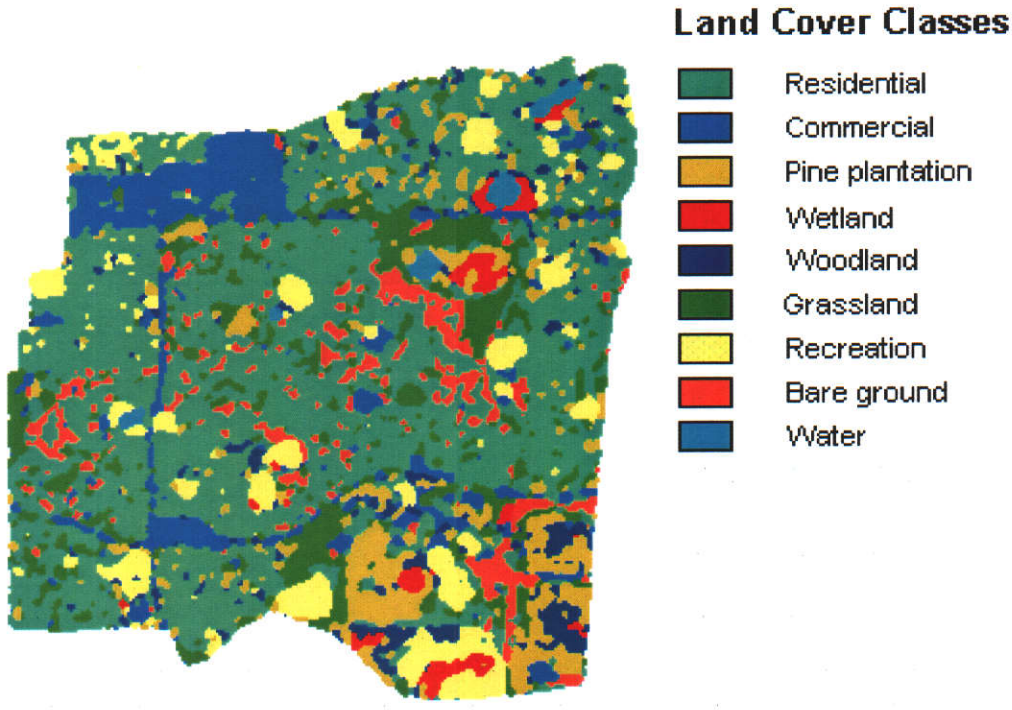


Figure 4.15 *Supervised classification of SPOT HRV data using Level II categories (1991) (Scale 1:50 000)*

4.5.2.2 Landsat TM

Figure 4.12 illustrates supervised classification results for the Landsat TM dataset at Level I. *Urban* and *Open* classes form the largest components of the image, however significant areas are still occupied by the *Forest* class that has been broken into small sections as a result of clearing for urban development. Inspection of the reference data in Figure 2.10a indicates the increasingly fragmented nature of the land covers in the area, which leads to a less uniform classification. Similar results were achieved with the unsupervised classification of Landsat TM data, however the fragmentation almost exclusively involved the *Urban* and *Open* categories, while in the supervised classification it involves all of the major classes including *Urban, Open* and *Forest.*

An emerging characteristic of the supervised classification approach is that at Level II all nine classes are represented in the classification due largely to each class being explicitly defined through extraction of spectral signatures for every category (Figure 4.13). The corresponding unsupervised classification (Figure 4.7) includes all classes except *Pine plantation.* However, a significant area of Figure 4.13 represents the *Pine plantation* class and verification with the reference image (Figure 2.10b) indicates that a substantial area of pine plantation is actually present in this locality. The unsupervised classification demonstrates that, at the level of clustering applied, *Pine plantation* is not spectrally separable. However, explicit identification of the class and derivation of spectral signatures has enabled identification of this land cover.

4.5.2.3 SPOT HRV

Comparison of Figures 4.14 and 4.15 with the corresponding images produced through unsupervised classification (Figures 4.8 and 4.9), confirms substantial differences in the distribution of classes throughout the image. In particular, the distribution of the *Open* category shows a large amount of variation in spatial distribution even though both sets of images have been filtered twice using a 3 x 3 modal filter.

Figure 4.15 again shows considerable fragmentation of classes well in excess of that present in the reference data (Figure 2.11b) or evident in the corresponding unsupervised classification (Figure 4.9). Figure 4.9 identifies only five classes compared to the nine classes identified in Figure 4.15 and contributes to the smoothed appearance of the classified image.

4.6 Summary

This chapter has investigated the derivation of an appropriate land cover classification scheme for an area of the rural-urban fringe that has been subject to substantial change throughout the duration of this project. It was necessary for the classification scheme to be hierarchical in structure such that information extracted from remotely sensed data of different spatial resolutions could be incorporated into the classification scheme (Levels I and II), for the purposes of change analysis. Traditional classification schemes are based upon manual interpretation and rely upon all facets of image interpretation, whereas the approach adopted in this research is reliant only upon automated interpretation via standard image classification procedures. Consequently, it was necessary for class definitions to be based upon spectral characteristics only. The current classification scheme is a modified form of the classification scheme developed by Anderson *et al* (1976).

Extensive evaluation of the spatial autocorrelation of image data at three spatial resolutions was undertaken. Assessment of these data indicate that at each spatial resolution significant spatial autocorrelation exists for most targets and, that if not taken into consideration, may result in an underestimation of the variance within training sites. Such underestimation results in poor parameterisation of the training data and potential confusion between classes during supervised classification. Training statistics were derived from groups of pixels with dimensions in excess of extents considered to be affected by spatial autocorrelation. Significant spatial autocorrelation was determined to exist for some land cover classes up to three pixels (180 m) distant for Landsat MSS data, up to five pixels (150 m) for Landsat TM data and up to six pixels (120 m) distant for SPOT HRV data. In all situations these values represent the limit of the autocorrelation analysis undertaken in this research.

Unsupervised and supervised classification approaches were applied to the multispectral satellite data in order to determine the optimum classification approach for each image epoch. Classifications were performed at two levels of detail according to the land cover classification scheme defined for this area. The results for these analyses have been assessed qualitatively in this chapter and will be evaluated quantitatively in Chapter 5.

Between 20 and 27 clusters were defined for the area using an ISODATA clustering algorithm, and through stratified random sampling these clusters were merged to form four land cover classes at Level I and nine classes at Level II. Unsupervised classification was used to identify spectrally separable land covers without the constraints applied by analyst intervention as occurs with the supervised classification process. At Level I the *Urban* class was identified as the spectrally most complex land cover with 14 clusters representing this class in the SPOT data, and at Level II the *Residential* class comprised 11 clusters in each of the Landsat TM and SPOT data. At both levels of interpretation the finer spatial resolution Landsat TM and SPOT HRV data resulted in a larger number of clusters for the *Urban* (Level I) and *Residential* (Level II) classes, where only nine and six clusters respectively, were recorded for these classes in the Landsat MSS data. The lower quantisation of the Landsat MSS system may be a contributing factor, however the major contribution is due to the degree of spatial averaging which occurs within these coarse resolution data.

At Level I the *Forest* class required nine clusters (Landsat MSS) and the *Open* class 11 clusters (Landsat TM) to describe the respective land covers. All other classes at Level I and Level II (apart from the spectrally complex *Urban* and *Residential* classes), required six or less clusters for representation. However, at the more detailed Level II, some land cover classes were not represented by any clusters. This result is a combination of the small area occupied by some land covers at the later sensing dates and the minimal spectral separability between these classes.

Training statistics were derived for 11 land cover classes which were subsequently coalesced into the nine classes defined at Level II. Detailed analysis of training data as part of the supervised classification process indicates that most land cover classes are spectrally separable at the 95 percent confidence level. The least amount of separation between the extracted spectral information (11 classes) is evident in the Landsat MSS data, with progressively greater separation for the Landsat TM and SPOT HRV data.

At a qualitative level of assessment, unsupervised classification using the ISODATA algorithm and supervised classification by a maximum likelihood approach produced results containing substantial spatial variability. Application of two passes of a modal filter reduced the spatial variability, and from a visual comparison with the reference

data, more consistent classification results were achieved. Some clear differences in land cover patterns exist between the classification approaches, and these will be investigated in the next chapter.

Chapter 5

CLASSIFICATION ACCURACY ASSESSMENT

Assessment of the classification accuracy of the primary remote sensor data is essential if a thorough evaluation of change detection mapping is to be undertaken. A range of map accuracy and sampling schemes will be investigated for evaluation of the quality of the classification results. The classification accuracy of each spatial resolution satellite dataset will be evaluated. The accuracy of the supervised and unsupervised classifications will be evaluated using land cover classes equivalent to USGS Level I and Level II to determine the optimum approach to information extraction for each dataset, which will then be utilised in subsequent change assessment mapping. Significant factors contributing to the outcome of the classification accuracy will also be discussed.

5.1 Introduction

Integration of geographic information derived from a range of sources including remote sensing has led to the requirement for increased knowledge of errors and their contribution to the overall quality of the final product. Error accumulation in remote sensing and GIS is difficult to monitor and, for remote sensing, variations in the target, sensor, sensing geometry and ambient environmental conditions create specific problems (McGwire and Goodchild, 1997). Both quantitative and qualitative errors are important in remote sensing. Quantitative errors relate to the positional accuracy of map data, while qualitative errors are concerned with the correctness of labelling or thematic classification of features within the data (Maling, 1988). It is this latter category of errors that will be considered in this chapter.

Classification differences between remotely sensed and reference data arise for a range of reasons (Davis and Simmonett, 1991):

(i) Misregistration of satellite data to the cartographic coordinate system;
(ii) Misregistration of reference data to the cartographic coordinate system;
(iii) Spectral confusion between information classes;
(iv) Inappropriate classification algorithm;
(v) Poor definition of information classes for training and test data;
(vi) Information classes containing several spectral classes;
(vii) Subpixel variations causing mixed pixel and boundary effects.

Error assessment is required to quantify the classification accuracy and to guide the process of analysis to determine the sources of classification error. Understanding of the above factors can lead to refinement of the classification approach and improvements in the quality of classification. Analysis of overall classification performance and analysis of performance by class will be used to evaluate the contribution of these factors, and determine the optimum classification approach for each satellite dataset. Zhang and Foody (1998) take the analysis one step further in considering the within-class spatial variability of errors using *fuzzy* membership approaches. Although these approaches consider the probability of multiple class membership, both at the class allocation and validation stages of analysis, conventional *hard* classification and accuracy analyses are appropriate in this study to provide a statement of land covers for input to the change analysis stage of the research.

Site specific and non-site specific methods of accuracy assessment are possible, however non-site specific methods only provide an assessment of the total area occupied by a specific land cover without performing any location by location comparison. If all classification errors between categories in a non-site specific assessment balance out, it is possible to achieve very high results for the accuracy assessment, however the results will be misleading. Site specific approaches make a one to one comparison of samples of the classified data and provide an assessment of the classification errors with regard for the location of the classified and reference data (Mead and Szajgin, 1982). Only site specific methods will be considered for analysis of classification accuracy in this research.

Statistically sound approaches to sample size and sampling design are required to perform valid assessments of classification accuracy for landscapes of varying spatial diversity (Congalton, 1991). Consideration of a range of random and systematic approaches will be made to ensure a robust sampling scheme is implemented for evaluation of the classified data.

5.2 Accuracy Assessment Approaches

Precision is defined as the degree of detail in reporting of a measurement which is often determined by the characteristics of the measuring equipment, while *accuracy* is defined as a measure of the difference between a measured value and a known or true value (McGwire and Goodchild, 1997). From a thematic mapping perspective, precision is related to the level of detail (or generalisation) inherent in the thematic mapping classification system (Janssen and van der Wel, 1994). In the context of thematic mapping quality assessment, accuracy relates to the agreement of the classified image with a source of reference data of greater accuracy than the primary remote sensing information. It is often derived from aerial photography or ground-based investigations, or a combination of both.

Analyses in this study will be directed towards assessment of the classification accuracy achieved through supervised and unsupervised classification of multiscale satellite data as derived in Chapter 4. The level of the classification scheme described in the previous chapter determines the precision of the classification. As the degree of detail increases from Level I to Level II, the possibility of errors also increases, which may lead to more

uncertain results and, logically, a lower classification accuracy (Janssen and van der Wel, 1994).

5.2.1 Descriptive Techniques

The advent of standardised land cover classification schemes for remotely sensed data (Anderson *et al*, 1976) generated significant interest in approaches to classification accuracy assessment. Early empirical accuracy assessment techniques relied upon simple ratios of correctly interpreted points to the total number of samples (Fitzpatrick-Lins, 1978). Although data were recorded in matrix form for each sampled category, undersampling and oversampling were common place and category-based accuracy statements were not derived. Large variations in accuracy between categories were possible, even though the overall result may have been acceptable (Story and Congalton, 1986).

Application of robust sampling schemes (see Section 5.3) enables acquisition of representative samples of every land cover class, and provides relevant data for population of the error matrix. An example of an error matrix (also known as a confusion matrix or contingency table) is shown in Table 5.3. The Overall Classification Accuracy is computed by dividing the sum of the diagonals by the total of samples checked in the accuracy assessment, and provides a measure of the proportion of all sampled pixels that are classified correctly. The mapping accuracy of each class may be derived in two ways, either by dividing the number of pixels identified correctly in the category by the total number in the reference data for that category or from the number of pixels actually assigned to the category by the classifier. The former approach provides a measure of accuracy called the Producer's Accuracy, and the latter measure is termed the User's Accuracy (Story and Congalton, 1986). These terms are related to errors of omission and commission as follows (Janssen and van der Wel, 1994):

(i) Error of commission = 100 – User's Accuracy,

(ii) Error of omission = 100 – Producer's Accuracy.

The User's Accuracy is a measure of the reliability of the classification because it measures the proportion of pixels that are classified as one category, but actually belong to other categories. Errors of commission express the severity of this situation. The

Producer's Accuracy gauges the proportion of pixels that actually belong to a category, but have been classified as other features. Errors of omission express the degree to which this type of error occurs. The User's and Producer's Accuracies also permit a more complete understanding of the between-class confusion for the purposes of signature refinement in supervised classification, or cluster evaluation in unsupervised classification.

The greatest significance may be attached to these separate measures of accuracy when the Producer's and User's Accuracies are dissimilar (Story and Congalton, 1986). For example, if the Producer's Accuracy for an area of forest is 93 percent it indicates that nearly all the reference data in the forest class were classified correctly. However, at the same time the User's Accuracy may be 49 percent, which means that of the area classified as forest, only about half is actually forest and the remainder comprises other classes.

Aronoff (1982) has integrated these values through statistical analysis for comparison of classifications with thematic map accuracy standards. The User's Accuracy (risk) specifies the probability that a map of unacceptable accuracy will pass the accuracy test, and the Producer's Accuracy (risk) specifies the probability that a map of some acceptable accuracy will be rejected. Acceptable levels of risk for the user and producer are defined as part of the analysis.

5.2.2 Analytical Techniques

Error matrices form the basis of several analytical statistical techniques developed to evaluate classification accuracy of remotely sensed data. Most approaches utilise discrete multivariate analysis because remotely sensed data are discrete rather than continuous. Most data demonstrate properties of binomial or multinomial distributions, therefore many methods based upon normal probability theory are not appropriate (Congalton, 1991).

Normalisation of an error matrix facilitates comparison of classification results, both overall and by category, for classifications by different algorithms or analysts. Without normalisation the variation in sample parameters (number of pixels) between error assessments renders direct comparison of error matrices irrelevant. Conversion of pixel

counts to percentages is possible, however uncertainty exists about whether the divisor should be the row or column total. An iterative procedure is available (Congalton, 1991) which normalises all rows and columns of the error matrix. Differences in the number of samples are eliminated and individual cells within the matrix are directly comparable. Final determination of classification accuracy must still be undertaken using other accuracy estimation approaches described in this section.

Ginevan (1979) applied the binomial probability density function to define testing procedures for assessment of thematic maps against map accuracy specifications. Specification of classification accuracy, confidence intervals, and acceptable User's and Producer's Accuracies enabled determination of the number of samples required and the permissible error rate. Further development by Aronoff (1985) provided a testing procedure to determine the actual level of thematic accuracy, rather than simple assessment of success or failure against a particular benchmark. The index derived by this method expresses statistically the uncertainty involved in the accuracy assessment, however the method is limited to a single overall accuracy value for the map or for each class, rather than using the entire error matrix for assessment (Congalton, 1991).

Accuracy assessments which include all elements of the error matrix may be undertaken using the Kappa Coefficient of Agreement (Cohen, 1960). The Kappa Coefficient was developed for comparison of data grouped by different observers (or interpreters or classification algorithms), according to nominal scales. The overall level of agreement for an error matrix (Kappa Coefficient) is based upon the difference between the actual agreement of the classification compared with the reference data (measured by the matrix diagonal), and the chance agreement, which is indicated by the product of the row and column margin values.

The application of the Kappa Coefficient to the analysis of classifications of remotely sensed data was first proposed by Congalton *et al* (1983), and has been widely reported since (Fitzgerald and Lees, 1994; Fung and LeDrew, 1988; Gong and Howarth, 1990; Lo and Watson, 1998; Rosenfeld and Fitzpatrick-Lins, 1986). The method may be used to evaluate an error matrix as a whole or for individual classes, or it may be used to statistically compare error matrices derived from different interpreters or using a variety of classification techniques.

The value of the overall Kappa Coefficient ($\hat{K}$) is computed from (Congalton, 1991):

$$\hat{K} = \frac{N\sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r}(x_{i+} * x_{+i})}{N^2 - \sum_{i=1}^{r}(x_{i+} * x_{+i})} \tag{5.1}$$

where

r = number *of rows/columns in the error matrix*

x_{ii} = *number of observations in row i and column i*

x_{i+} = *total of row i*

x_{+i} = *total of column i*

N = total *number of observations.*

A pairwise assessment of the significance of the differences between two independent error matrices may be undertaken using the normal curve deviate determined from the corresponding Kappa statistics and their variances (Cohen, 1960). The variance of Kappa is computed from the following (Bishop *et al*, 1975):

$$\hat{\sigma}^2_{\hat{K}} = \frac{1}{N}\left\{\frac{\theta_1(1-\theta_1)}{(1-\theta_2)^2} + \frac{2(1-\theta_1)(2\theta_1\theta_2 - \theta_3)}{(1-\theta_2)^3} + \frac{(1-\theta_1)^2(\theta_4 - 4\theta_2^2)}{(1-\theta_2)^4}\right\} \tag{5.2}$$

where

$$\theta_1 = \sum_{i=1}\frac{x_{ii}}{N} \qquad \theta_2 = \sum_{i=1}\frac{x_{i+} * x_{+i}}{N^2}$$

$$\theta_3 = \sum_{i=1}\frac{x_{ii}}{N^2}(x_{i+} + x_{+i}) \qquad \theta_4 = \sum_{\substack{i=1\\j=1}}\frac{x_{ij}}{N^3}(x_{j+} + x_{+i})^2$$

The test statistic for significance in large samples ($N > 100$) is given by:

$$z \sim \frac{\hat{K}_1 - \hat{K}_2}{\sqrt{\hat{\sigma}_1^2 + \hat{\sigma}_2^2}} \tag{5.3}$$

An alternative means of measuring the improvement in classification above random assignment is the Tau coefficient (Ma and Redmond, 1995). Tau is conceptually similar to Kappa, but the major difference between the two coefficients is that Tau is based on an adjustment of the Overall Classification Accuracy using *a priori* probability of class membership. Tau may be computed using equal probabilities, as occurs with unsupervised classification and supervised classification with equal *a priori* class membership, or using unequal probabilities which occur when supervised classifications with unequal probabilities of class membership are undertaken. For equal probability of class membership, as used in this research, the adjustment of Tau is directly proportional to the number of classes, whereas Kappa utilises the *a posteriori* probabilities derived from the actual distribution of pixels over all classes within the error matrix.

Table 5.1 illustrates examples of comparative values for the Overall Classification Accuracy, Tau and Kappa parameters derived from data provided by Congalton *et al* (1983) and Ma and Redmond (1995), for a range of supervised and unsupervised classification approaches. These data indicate reasonably consistent variations between the three parameters across all classifications. The Overall Classification Accuracy is considered to significantly overestimate classifier performance and has resulted in the general acceptance of the Kappa statistic (Congalton *et al*, 1983). Foody (1992) and Ma and Redmond (1995) consider Kappa to underestimate classifier performance due to overestimation of chance agreement. In their assessment, Ma and Redmond (1995) assume equal probabilities for computation of the Tau coefficient from the results produced by Congalton *et al* (1983), whereas the Kappa Coefficient was determined from the chance agreement estimated from the actual classified data which considered unequal distribution between classes. Consequently, a more relevant comparison is to compare the Kappa results to those computed for Tau assuming unequal (but unknown) probabilities of occurrence.

The Kappa statistic provides statistically valid assessments of the quality of classification and enables tests of significance between classifiers for determination of optimum algorithm performance (Fitzgerald and Lees, 1994). The Tau coefficient enables a similar level of evaluation, but the use of *a posteriori* probabilities by Kappa that are derived from the complete error matrix is advantageous (McGwire and

Goodchild, 1997). This may also be significant compared to a general assumption of equal probability of class membership. While this may lead to underestimation of classification accuracy, these values are mainly utilised for comparative purposes to evaluate classifier performance, and the value on an absolute scale is less relevant. Landis and Koch (1977) have used the qualitative descriptors shown in Table 5.2 to describe the strength of agreement based upon Kappa statistics.

Accuracy parameter	Unsupervised 10 clusters	Unsupervised 20 clusters	Modified supervised	Modified clustering
Overall Class. Acc.	0.77	0.78	0.71	0.86
Tau	0.69	0.71	0.62	0.81
Kappa	0.60	0.59	0.48	0.72

Table 5.1 *Examples of classification accuracy parameters (from Congalton et al, 1983; Ma and Redmond, 1995)*

Kappa Statistic	Strength of Agreement
< 0.00	*Poor*
0.00 – 0.19	*Slight*
0.20 – 0.39	*Fair*
0.40 – 0.59	*Moderate*
0.60 – 0.79	*Substantial*
0.80 – 1.00	*Excellent*

Table 5.2 *Qualitative descriptors for the strength of agreement for the Kappa statistic (after Landis and Koch, 1977)*

The Kappa Coefficient will be used throughout this research for comparison of classification approaches using statistical analysis. Classifications at Level I and Level II for satellite data at each spatial resolution will be compared in Section 5.4 to determine the degree of significance of the observed differences between supervised and unsupervised classifications. These assessments will be utilised to determine the most appropriate classification approach. Overall Classification Accuracy will be considered where evaluation of the actual rate of correct classification is required.

5.3 Thematic Accuracy Sampling Design

5.3.1 Sampling Scheme

Assessment of the quantitative or qualitative aspects of map accuracy rely upon a sampling scheme with a common set of criteria based upon (Ginevan, 1979):

(i) A low probability of accepting a map of low accuracy,

(ii) A high probability of accepting a map of high accuracy,

(iii) Requiring a minimum number of reference data samples.

The sampling scheme represents the protocol for selecting the sampling units that constitute the reference sample. A sampling scheme may be developed to estimate the effectiveness of different accuracy parameters, provide information on land covers of limited areal extent, evaluate different mapping procedures or assess change detection accuracy (Stehman, 1999). Important considerations for sampling design are indicated by Mead and Szajgin (1982) and Stehman (1999):

(i) Satisfies sampling design protocol,

(ii) Simple to implement and analyse,

(iii) Low variance for estimates of high priority accuracy measures,

(iv) Permits adequate variance estimation,

(v) Provides samples which are spatially well distributed,

(vi) Cost effectiveness.

The sampling scheme must ensure statistical validity and provide a practical means of implementation. The actual sampling procedure employed and specification of at least an adequate number of sample points will determine statistical validity.

Fundamental to the design of a suitable sampling scheme is the selection of reference data independently of the data used to develop the classification (Mead and Szajgin, 1982). This particularly relates to the selection of independent samples for classifier training and accuracy assessment, in the same manner that independent samples are selected for image registration and assessment of positional accuracy.

The quality of classification accuracy assessment is also affected by the thematic and positional quality of the reference data. Reference data are generally derived from aerial photographs and are therefore subject to their own sources of error (Congalton,

1991; Mead and Szagjin, 1982). Processes to correct for reference data-induced errors have been proposed (Kalkhan *et al*, 1998), but will not be applied in the sampling design for this study. Reference data have been derived from medium scale aerial photography acquired at intervals either side of each sensing date in order to minimise these errors.

Sampling schemes generally follow simple random or systematic selection protocol and utilise population, strata or cluster sampling structures. The most relevant approach to selection protocol and sampling structures are subject to considerable debate and will be reviewed in the following (Congalton, 1988a; Stehman, 1999).

(i) *Simple random sampling*. Simple random sampling may be used to sample pixels in an image and represents a straightforward approach, although location of the samples in the field may be difficult. The method is adaptable to augmenting or reducing the sample size if required, and estimation and standard error formulae are less complex compared to other approaches. The variance in simple random sampling tends to be large where categories have small samples. Systematic sampling is more precise and cluster sampling may result in smaller variances depending on the spatial pattern of classification errors.

(ii) *Systematic sampling*. Systematic sampling may be undertaken on a rectangular or square grid based upon the random location of the starting pixel. The method provides good spatial coverage and is easy to implement, and because it is an equal probability sampling design it shares the same advantages and disadvantages as simple random sampling, such as small subregions will remain undersampled. Many authors (Aronoff, 1985; Congalton, 1991; Lo and Watson, 1998; van Genderen and Lock, 1977) consider systematic sampling susceptible to biased estimates. However, the equal probability sampling characteristic means that systematic sampling estimates are no more susceptible to bias than corresponding simple random samples, unless the spatial distribution of the errors is periodic and the sampling interval of the systematic design is in phase with this periodicity. Verification of this circumstance may make systematic sampling a viable (and more desirable) option. Where spatial autocorrelation (clustering) of errors exist, systematic sampling provides a lower variance of the accuracy estimate. Systematic unaligned sampling is often applied to avoid the possibility of periodicity if systematic sampling would otherwise be employed.

Congalton (1991) indicates a method for introducing random variation into systematic sampling design. This approach diminishes any potential effect of periodicity in the errors, but also precludes the full advantage of systematic sampling being realised when it would otherwise be appropriate.

(iii) *Stratified random sampling*. Each pixel within the population is assigned to a stratum prior to the application of simple random sampling within each stratum. Strata may be defined on a category or geographic basis, with each random sample derived independently for each stratum. Stratified random sampling is used to ensure that each stratum is adequately represented within the random sample, but care must be taken to ensure the strata remain valid throughout the analysis. Where analysis objectives change, such as when land cover classes are combined, a new stratified random sample should be extracted in order to retain relevance (Stehman, 1999). When independent classifications at Level I and Level II were evaluated in this research, such as for unsupervised classifications, a separate stratified random sample was extracted for each dataset. Samples are defined in this study by random sampling within a geographic strata (grid), which provides the spatial distribution advantage of systematic and systematic unaligned sampling, but is less susceptible to the effects of periodicity in the data (Stehman, 1999).

(iv) *Cluster sampling*. Cluster sampling employs two sizes of sampling units. The primary sampling unit which is the cluster itself, and the secondary sampling unit which is represented by the individual pixels within the cluster. The location of each cluster may be defined by any of the above sampling schemes, with each pixel in a cluster forming part of the sample. Cluster sampling is performed mainly for the purposes of convenience and cost, as only each cluster need be evaluated rather than individual pixels (Congalton, 1991). It may also be useful for dealing with uncertainties induced through small misregistration errors between datasets (Congalton, 1988a). Limitations of cluster sampling occur when classification errors are spatially correlated (Congalton, 1988b), and gives rise to greater sampling error than for simple random sampling of the same size that is based on point samples (Lo and Watson, 1998).

Concerns regarding the statistical validity of systematic sampling designs have made researchers consider stratified random sampling the most suitable approach for

sampling spatial data for the purpose of accuracy assessment (Aronoff, 1985; Congalton, 1988a and 1991; Ginevan, 1979; Lo and Watson, 1998; van Genderen and Lock, 1977). Recent research indicates that systematic sampling in the absence of data periodicity is the most statistically valid sampling approach to employ (Stehman, 1999). This applies especially where spatial autocorrelation is present, because it provides for maximum average spatial separation of samples. Systematic sampling designs also provide a uniform spatial distribution of samples, a factor which is often used to justify stratified random sampling.

Experimental evaluation by Lo and Watson (1998) of each of the sampling designs described above against a total classification of the reference data, determined there was no significant difference between any of the sampling designs at the 95 percent confidence level. However, the most similar results were obtained from the stratified random sample, followed by the systematic unaligned sampling design. Given the caveat of periodicity applicable to the implementation of systematic sampling, stratified random sampling is the most reliable approach for general application to classification accuracy analysis.

Stratified random sampling will be utilised in this research with the strata defined geographically on the basis of division of the image into uniform rectangular regions defined by the number of samples required. Subsampling of these regions will be achieved by random selection of one point within each region.

5.3.2 Sample Size

The major objective of selecting a sample of appropriate size is to provide sufficient data to enable a reliable estimate of the precision at the required confidence level. Where error matrices are used as part of the evaluation and large numbers of categories are sampled, the requirement to adequately sample each category becomes important (Congalton, 1991).

Determination of the total sample size based on the binomial theorem (Equation 5.4) is commonly used (Fitzpatrick-Lins, 1981). Equation 5.4 is useful for specifying the minimum number of sample points for reliable determination of Overall Classification Accuracy, but may result in inadequate sampling of categories occupying smaller areas.

$$N = \frac{z^2 (P)(Q)}{E^2} \tag{5.4}$$

where

N = *minimum sample size*

Z = *standard normal deviate for 95 percent confidence level*

P = *expected percent accuracy*

Q = *100 - P*

E = *allowable error.*

Congalton (1991) suggests the above value as a minimum number of samples overall, and to augment this amount by selecting at least 50 samples per category and 75 samples where categories occupy large areas. There is considerable variation in the specification of class sample sizes with Richards (1993) suggesting values between 30 and 60 for *most* situations. Limitations on sample sizes are usually invoked in order to economise on fieldwork expenditure. In this research a complete reference map of the area has been produced, which obviates the need to consider fieldwork costs. Therefore, assuming n samples to be the minimum number of samples per category and applying a stratified random sampling scheme, the total number of samples for evaluation of each classification is determined by:

$$T = n / A_x / A_T \tag{5.5}$$

where

T = *Total number of pixels to be sampled*

n = *number of samples per class*

A_x = *Area of smallest class*

A_T = *Total area classified.*

The computation of T was undertaken using only categories that occupy a reasonable sized area and a minimum of 75 samples per class was selected. For example, the *Water* category produces little variability and occupies a relatively small area, and if included in the sample calculation would indicate collection of an inordinately large

sample. As a result, *Water* was ignored in the determination of sample sizes. Due to the structured nature of the supervised classification a sample of approximately 1 400 pixels was selected to assess each dataset at both Level I and II. The same size sample was utilised at both levels because generalisation from Level I to Level II is based upon postclassification class aggregation derived from classification using the same spectral signatures. As a result of the variability in the clustering process, samples of approximately 650 were derived for assessment at Level I, and 1 400 at Level II for unsupervised classification. In this case the clusters were aggregated separately for each level of classification and the sampling was also undertaken independently.

5.4 Classification Accuracy Assessment

Classifications performed using supervised and unsupervised classification algorithms will be evaluated for each of the multispectral datasets of the study area. Overall Classification Accuracy and the Kappa Coefficient will be computed to provide measures of the success of the classification process. Comparative values of Kappa were discussed in Section 5.2.2 and provide a relevant benchmark for comparison (Landis and Koch, 1977). The relevant values are (in percent):

(i) Fair agreement $20 < K < 39$,

(ii) Moderate agreement $40 < K < 59$,

(iii) Excellent agreement $60 < K < 79$.

The User's and Producer's Accuracy, and elements of the error matrix will be evaluated to assess error patterns within each classification. Assessment of both the most appropriate level of classification (Level I or Level II) and the optimum classification approach will also be made in Section 5.4.3, which will be utilised for subsequent change assessment in Chapter 6.

5.4.1 Unsupervised Classification

5.4.1.1 Landsat MSS

The results of unsupervised classification of the Landsat MSS data are shown in Tables 5.3 and 5.4. The Kappa Coefficients indicate that classification at Levels I and II show *moderate* agreement, although at Level II the Kappa value is relatively low at 45.9 percent. This reflects the effect on the detail of the classification of the relatively coarse spatial resolution of the Landsat MSS data. At 79 x 56 m resolution, significant overlap

Classified data	Reference data					User's Acc.
	Urban	Forest	Open	Water	Total	%
Urban	**133**	23	23	0	*179*	*74.3*
Forest	28	**400**	59	1	*488*	*82.0*
Open	6	2	**7**	0	*15*	*46.7*
Water	0	0	0	**1**	*1*	*100.0*
Total	*167*	*425*	*89*	*2*	*683*	
Prod Acc	*79.1*	*94.1*	*7.9*	*50.0*		
					Overall Classification Accuracy (percent)	*79.1*
					Kappa Coefficient ($\hat{K}$) (percent)	*57.3*

Table 5.3 *Error matrix for unsupervised classification at Level I using Landsat MSS data*

Classified data	Reference data										User's Acc.
	Res	Com	Pine	Wet	Wod	Gras	Rec	Bare	Wat	Tot.	%
Residential	**201**	19	8	1	2	9	8	18	1	*289*	*69.5*
Commerc	4	**33**	5	1	11	3	1	17	0	*75*	*44.0*
Pine plant	2	6	**510**	42	100	31	1	43	4	*739*	*69.0*
Wetland	0	0	0	**0**	0	0	0	0	0	*0*	-
Woodland	5	7	37	9	**76**	7	0	17	0	*158*	*48.1*
Grassland	0	0	0	0	0	**0**	0	0	0	*0*	-
Recreation	0	0	0	0	0	0	**0**	0	0	*0*	-
Bare grnd	15	10	15	0	25	16	0	**24**	0	*105*	*22.9*
Water	0	0	0	0	0	0	0	0	**1**	*1*	*100*
Total	*229*	*75*	*575*	*53*	*236*	*66*	*10*	*119*	*6*	*1367*	
Prod. Acc.	*87.8*	*43.4*	*88.7*	*0*	*32.1*	*0*	*0*	*20.2*	*16.7*		
										Overall Classification Accuracy (percent)	*61.6*
										Kappa Coefficient ($\hat{K}$) (percent)	*45.9*

Table 5.4 *Error matrix for unsupervised classification at Level II using Landsat MSS data*

between classes is expected, especially for classes that occupy relatively small areas such as *Water*.

At Level I the User's and Producer's Accuracies for *Urban* and *Forest* are relatively high. These values indicate a high level of agreement with the reference data and express high confidence in their identification. The *Open* class shows very poor accuracy with most of the reference pixels classified as *Urban* and *Forest*, and very few actually identified as *Open*. The *Open* category includes the Level II classes of *Grassland, Recreation* and *Bare ground*, and inspection of Table 5.4 indicates that all of these subclasses demonstrate substantial confusion with *Residential, Commercial* and *Pine plantation* which form parts of the *Urban* and *Forest* classes. The extent of *Water* is very limited in all data used in this study and detailed analysis of the results for the Landsat MSS data is not appropriate.

Examination of Table 5.4 shows that no clusters from the ISODATA algorithm were allocated to the *Wetland, Grassland* or *Recreation* land covers. *Wetland* pixels were allocated mostly to *Pine plantation*, *Grassland* to *Pine plantation* and *Bare ground*, and *Recreation* to *Residential*. *Wetland* comprises native forest and shrubs, and may exhibit spectral characteristics similar to dense pine forest. *Grassland* comprises areas of scattered trees with herbaceous understorey as well as areas without any trees. Depending upon the season, some areas may appear similar to low density pine forest and in summer, areas without trees may have the appearance of *Bare ground*. *Recreation* areas often contain substantial numbers of trees and, in combination with the irrigated grass and adjacency to residential areas, confusion with the *Residential* class is possible.

Generally, *Residential* and *Pine forest* achieved moderate levels of classification accuracy from a Producer's and User's perspective. *Commercial* misclassified many pixels as *Residential*, and incorrectly included many pixels from *Woodland* and *Bare ground* in the classification. High levels of reflectance from *Commercial* and *Bare ground* and the extent of variability from each of these classes most likely leads to this degree of confusion. *Woodland* shows low User's and Producer's Accuracies with mutual confusion with the *Pine plantation* class. The spectral similarity of these land covers was not evident on a class basis (see Section 4.4.2.2), but variation between

individual pixels results in widespread confusion. A very high degree of confusion exists between *Bare ground* and the *Residential, Commercial, Pine plantation* and *Woodland* classes from both a Producer's and User's perspective. This indicates poor spectral definition of these land covers and a high degree of within-target variation, which produces the variable spectral responses. This is a characteristic bought about by the large spatial resolution element of the Landsat MSS sensor and resultant mixed pixel effect. This conclusion is supported by the analysis of spectral separability in Section 4.4.2.2.

5.4.1.2 Landsat TM

The unsupervised classification of Landsat TM data provides almost identical results to those achieved using Landsat MSS. Tables 5.5 and 5.6 show *moderate* agreement with Kappa values of 58.7 percent and 46.9 percent for Levels I and II, respectively. While improved classification results were expected due to the enhanced spectral and spatial resolution of the Landsat TM data, the data have been collected 14 years apart and the patterns of land covers and their spatial complexity have changed to a great extent. The main objective of this research is to evaluate the potential for multiscale data integration for change assessment, and does not focus on the comparative information extraction capability of the sensors.

At Level I, the *Urban, Open* and *Water* classes all show high Producer's and User's Accuracies that provide confidence in the consistency of the classification. *Forest* indicates a high User's Accuracy, but confusion of the *Forest* reference data with the *Urban* and *Open* classes reduces the Producer's Accuracy to 44.6 percent. Areas of *Pine plantation* are reasonably extensive, however Figure 2.10 shows many areas of this land cover type are of linear form. Any minor misregistration with the reference data results in a high level of apparent misclassification with adjacent areas such as *Grassland.*

The Level II classification indicates a high Producer's Accuracy (89.4 percent) for *Residential,* but confusion with all other classes results in a User's Accuracy of 57.1 percent. A broadly defined spectral signature causes this effect for *Residential* which incorporates characteristics from all other classes except *Water.* The *Commercial* class demonstrates the opposite situation. *Commercial* land covers have been described in a

Classified Data	Reference Data					User's Acc.
	Urban	Forest	Open	Water	Total	%
Urban	**192**	15	38	0	*245*	*78.4*
Forest	1	**37**	5	0	*43*	*86.0*
Open	44	30	**173**	0	*247*	*70.0*
Water	0	1	0	**3**	*4*	*75.0*
Total	*237*	*83*	*216*	*3*	*539*	
Prod. Acc.	*81.0*	*44.6*	*79.7*	*100.0*		
					Overall Classification Accuracy (percent)	*75.0*
					Kappa Coefficient ($\hat{K}$) (percent)	*58.7*

Table 5.5 *Error matrix for unsupervised classification at Level I using Landsat TM data*

Classified Data	Reference data										User's Acc.
	Res	Com	Pine	Wet	Wod	Gras	Rec	Bare	Wat	Tot.	%
Residential	**464**	80	37	12	49	89	46	36	0	*813*	*57.1*
Commerc	0	**26**	0	0	0	0	0	0	0	*26*	*100*
Pine plant	0	0	**0**	0	0	0	0	0	0	*0*	-
Wetland	2	0	4	**5**	1	1	2	0	1	*16*	*31.2*
Woodland	0	0	23	18	**42**	6	1	1	0	*91*	*46.2*
Grassland	38	3	13	2	7	**179**	11	13	0	*266*	*67.3*
Recreation	3	3	1	1	2	4	**51**	0	1	*66*	*77.3*
Bare grnd	11	5	2	0	3	13	2	**97**	0	*133*	*72.9*
Water	0	0	0	2	0	0	1	0	**4**	*7*	*57.1*
Total	*518*	*117*	*80*	*40*	*104*	*292*	*114*	*147*	*6*	*1418*	
Prod. Acc.	*89.4*	*22.0*	*0*	*12.5*	*40.0*	*61.3*	*44.7*	*66.0*	*66.7*		
										Overall Classification Accuracy (percent)	*61.2*
										Kappa Coefficient ($\hat{K}$) (percent)	*46.9*

Table 5.6 *Error matrix for unsupervised classification at Level II using Landsat TM data*

narrow range of values indicated by the high User's Accuracy (100 percent). However, the spectral range of the clusters is not sufficiently broad to include the range of radiances encompassed by the reference data for the class, which results in a Producer's Accuracy of only 22.0 percent.

No clusters were defined for the *Pine plantation* class (and therefore no User's Accuracy could be computed) and reference data for this class were classified mostly as *Residential* and *Woodland,* providing a Producer's Accuracy of zero. This outcome indicates that *Pine plantation* did not form a unique spectral class based on the decision rules of the classifier. Moderate class accuracies were recorded for *Grassland, Bare ground* and *Water* with balanced Producer's and User's Accuracies between 60 and 80 percent. *Grassland* is confused mostly with *Residential,* and *Bare ground* is confused mostly with *Grassland* and *Residential,* which is largely consistent with the results shown in Table 4.10. *Water* is confused with the *Wetland* class, which may be due to some *Wetland* reference data actually containing surface water during seasonal changes in the water table that were not present in the aerial photographs. Ground investigations subsequently confirmed the existence of an area of seasonal surface water.

Wetland demonstrates a low User's Accuracy due to confusion with the *Pine plantation* class which contains a similar density of vegetation, and an extremely low Producer's Accuracy due to confusion with the *Woodland* and *Residential* classes. In both cases the observed pixel values are similar to the spectral ranges characteristic of these classes, probably caused by dense vegetation in the more established residential areas and densely vegetated native woodland areas.

The Producer's Accuracy of 40.0 percent for *Woodland* is very low and shows a high level of misclassification with the *Residential* class caused by the similarity of *Woodland* pixels to the spectral characteristics of heavily vegetated *Residential* clusters. The User's Accuracy is also low (46.2 percent) with most confusion occurring with pixels of the heavily vegetated land covers of *Pine plantation* and *Wetland* being erroneously classified as *Woodland.*

The *Recreation* class presents a relatively high User's Accuracy of 77.3 percent, which indicates the spectral characteristics defined by the clustering are quite dissimilar to

other land covers. However, with a Producer's Accuracy of only 44.7 percent, the *Recreation* class is too tightly defined and more than half of the *Recreation* pixels were not allocated correctly. The major confusion occurred with *Residential* due to the similarity to residential gardens and spatial confusion of overlapping pixels. To a much lesser extent confusion also occurred with *Grassland,* which would be similar to some of the less well irrigated areas of the *Recreation* class.

5.4.1.3 SPOT HRV

Unsupervised classifications of the SPOT HRV data are shown Tables 5.7 and 5.8, and indicate *moderate* agreement at Level I ($\hat{K} = 50.4$ percent), but only *fair* agreement at Level II ($\hat{K} = 36.9$ percent). With reference to Figure 2.11 the pattern of land covers has now changed to incorporate a very large proportion of the total study area as *Residential,* with substantial proportions of *Commercial.* Moderately sized areas of *Woodland* and *Grassland* still exist, but are less contiguous than in the Landsat MSS or Landsat TM data. *Recreation, Bare ground* and *Pine plantation* now appear in a large number of smaller isolated locations. Overall, the study area is now a complex assemblage of land covers dominated by the *Residential* and *Commercial* classes with fragmented areas of all other classes. Although the SPOT HRV data are of relatively fine spatial resolution, the issues of spectral separability and registration accuracy make accurate delineation of all land covers a difficult task.

Classified Data	**Reference Data**					**User's Acc.**
	Urban	**Forest**	**Open**	**Water**	**Total**	*%*
Urban	**495**	30	86	1	612	*80.9*
Forest	0	**48**	4	0	52	*92.3*
Open	1	2	**27**	0	30	*90.0*
Water	0	5	0	**4**	9	*44.4*
Total	496	85	117	5	703	
Prod. Acc.	*99.8*	*56.5*	*23.1*	*80.0*		
					Overall Classification Accuracy (percent)	81.6
					Kappa Coefficient ($\hat{K}$) (percent)	50.4

Table 5.7 *Error matrix for unsupervised classification at Level I using SPOT HRV data*

At Level I the *Residential* class shows very high User's and Producer's Accuracies indicating a high class mapping accuracy. Some *Forest* and *Open* pixels are included in the *Residential* class due to their presence as subclasses within the *Residential* class. This situation is more likely to arise given the finer spatial resolution of the SPOT HRV sensor.

Forest and *Open* show high User's Accuracies, but moderate and very low Producer's Accuracies, respectively. The result for *Forest* is due to classification of 35 percent of the *Forest* reference data into the *Urban* class, indicating the *Forest* clusters are too finely defined. The *Open* class has extremely low Producer's Accuracy due to classification of 73 percent of the *Open* reference data into the *Urban* class. This is due to the spectral responses of many *Open* pixels that are similar to some characteristics of bare ground, and also vacant land parcels that are also found within the *Urban* class.

Classified Data	**Reference Data**										**User's Acc.**
	Res	**Com**	**Pine**	**Wet**	**Wod**	**Gras**	**Rec**	**Bare**	**Wat**	**Tot.**	**%**
Residential	**864**	98	16	6	40	104	40	32	1	*1201*	*71.9*
Commerc	2	**34**	0	0	2	0	2	6	0	*46*	*73.9*
Pine plant	0	0	**0**	0	0	0	0	0	0	*0*	-
Wetland	0	0	0	**0**	0	0	0	0	0	*0*	-
Woodland	3	1	26	20	**39**	4	4	0	0	*97*	*40.2*
Grassland	0	0	0	0	0	**0**	0	0	0	*0*	-
Recreation	1	1	0	1	4	1	**51**	0	0	*59*	*86.4*
Bare grnd	0	0	0	0	0	0	0	**0**	0	*0*	-
Water	2	0	0	7	0	0	0	0	**11**	*20*	*55.0*
Total	*872*	*134*	*42*	*34*	*85*	*109*	*97*	*38*	*12*	*1423*	
Prod. Acc.	*99.0*	*25.4*	*0*	*0*	*45.4*	*0*	*52.6*	*0*	*91.7*		
										Overall Classification Accuracy (percent)	*70.1*
										Kappa Coefficient $(\hat{K})$ (percent)	*36.9*

Table 5.8 *Error matrix for unsupervised classification at Level II using SPOT HRV data*

Water provides a high Producer's Accuracy, but a low User's Accuracy in both Level I and II classifications. This low value is due to inclusion of some *Forest* pixels as *Water*, which may be due to registration issues associated both with the satellite data and reference data as both *Water* and *Forest* occur adjacent to each other. Seasonal variations in water level as discussed in Section 5.4.1.2 may also be relevant.

The low level of agreement $(\hat{K})$ at Level II is due to the classes of *Pine plantation*, *Wetland*, *Grassland* and *Bare ground* not being represented by any clusters at the cluster allocation stage. Further investigation of the reference data suggests the classes of *Pine plantation* and *Wetland* could justifiably be generalised and grouped with the *Woodland* class, which would improve the overall result. Evaluation of the *Grassland* and *Bare ground* classes indicates that spectral confusion of these classes occurs mainly with the *Residential* class rather than with any structurally similar classes such as *Recreation*, and therefore no improvement in classification performance could be achieved through generalisation of these classes. Generalisation to this level produces a Kappa Coefficient of 42.9 percent and an Overall Classification Accuracy of 73.4 percent, which is less than the result achieved for Level I, yet the level of generalisation is similar.

Residential recorded an extremely high (99 percent) Producer's Accuracy due to well defined cluster characteristics, but a moderate User's Accuracy (71.9 percent), due to inclusion of pixels from other classes. Considering the spectral variation of pixels normally experienced in urban areas and the fine spatial resolution of the SPOT HRV sensor, this level of accuracy represents an excellent result. The reverse situation occurred for the *Commercial* class with a User's Accuracy of 73.9 percent and a Producer's Accuracy of 25.4 percent. Only a limited number of pixels from other classes were classified as *Commercial* except for *Bare ground*, which supports the results shown in Table 4.11 and confirms the separability of *Commercial* from other classes. However, within the *Commercial* reference data, 73 percent of the pixels were classified as *Urban*, indicating they showed similarities to the *Residential* class due to the spectral characteristics of roof materials, road surfaces, paving and some vegetated and bare areas.

Woodland demonstrates very poor separability, both from a Producer's and User's perspective. *Woodland* was confused mainly with pixels from similar vegetation classes such as *Pine plantation* and *Wetland,* indicating clusters from all three classes possess similar characteristics. The *Woodland* reference data were mainly misclassified as *Residential,* which was likely to have been caused by *Woodland* areas appearing similar to clusters representing heavily vegetated gardens.

Recreation pixels appear to be well defined with a User's Accuracy of 86.4 percent and no major confusion with other classes. The Producer's Accuracy is 52.6 percent and indicates classification of *Recreation* reference pixels into the *Residential* class as the main cause of errors. This may be caused by registration errors between the satellite and reference data, and therefore assessment of some *Recreation* clusters against the *Residential* class, or due to similarities in the cluster characteristics for *Recreation* and parts of the *Residential* class. *Residential* is a spectrally complex class and comprises six (Landsat MSS) or 11 (Landsat TM and SPOT HRV) clusters, whereas *Recreation* comprises zero and three clusters respectively, for these data sources.

5.4.2 Supervised Classification

5.4.2.1 Landsat MSS

Tables 5.9 and 5.10 detail the results of supervised classification for the Landsat MSS data of the study area. The Kappa Coefficients indicate *moderate* agreement with values of 56.2 percent and 46.3 percent for Levels I and II, respectively. The largest classes of *Urban* and *Forest* in the Level I classification show reasonably balanced User's and Producer's Accuracies, with the major source of confusion occurring in the *Forest* reference data with misclassification into the *Open* class. This same effect did not occur for the unsupervised classification (Table 5.3) and indicates the clusters were more broadly defined than the spectral signatures of the supervised classification.

Open has relatively low User's and Producer's Accuracies with 50 percent of the *Open* reference data classified as *Urban* or *Forest,* and large proportions of refrence pixels from these classes being classified as *Open.* These results indicate the *Open* class displays high spectral variability which leads to difficulty in differentiating it from other classes. Table 4.9 supports this assertion indicating that the *Open* subclasses of

Classified Data	Reference Data					User's Acc.
	Urban	Forest	Open	Water	Total	%
Urban	**260**	52	49	0	*361*	*72.0*
Forest	19	**708**	44	3	*774*	*91.5*
Open	38	138	**94**	1	*271*	*34.7*
Water	0	4	0	**3**	*7*	*42.9*
Total	*317*	*902*	*187*	*7*	*1413*	
Prod. Acc.	*81.8*	*78.2*	*50.0*	*42.9*		
					Overall Classification Accuracy (percent)	*75.4*
					Kappa Coefficient $(\hat{K})$ **(percent)**	*56.2*

Table 5.9 *Error matrix for supervised classification at Level I using Landsat MSS data*

Classified Data	Reference Data										User's Acc.
	Res	Com	Pine	Wet	Wod	Gras	Rec	Bare	Wat	Tot.	%
Residential	**202**	13	9	0	20	7	4	20	0	*275*	*73.4*
Commerc	8	**37**	8	0	15	3	0	15	0	*86*	*43.0*
Pine plant	1	1	**430**	25	40	4	0	15	0	*516*	*83.3*
Wetland	0	0	43	**7**	25	3	0	0	3	*81*	*8.6*
Woodland	15	2	39	5	**94**	4	0	18	0	*177*	*53.1*
Grassland	7	3	67	6	42	**26**	0	29	1	*181*	*14.4*
Recreation	6	2	0	0	0	1	**6**	0	0	*15*	*40.0*
Bare grnd	8	12	6	0	17	8	0	**24**	0	*75*	*32.0*
Water	0	0	1	3	0	0	0	0	**3**	*7*	*42.9*
Total	*247*	*70*	*603*	*46*	*253*	*56*	*10*	*121*	*7*	*1413*	
Prod. Acc.	*81.4*	*52.9*	*71.1*	*15.2*	*36.9*	*46.4*	*54.6*	*19.8*	*42.9*		
										Overall Classification Accuracy (percent)	*58.4*
										Kappa Coefficient $(\hat{K})$ **(percent)**	*46.3*

Table 5.10 *Error matrix for supervised classification at Level II using Landsat MSS data*

Grassland, Recreation and *Bare ground* show low significant differences with several *Urban* and *Forest* subclasses. Similar low classification results were achieved with the unsupervised classification of these data (Tables 5.3 and 5.4), which confirm this result. However, with the supervised and unsupervised classification of the Landsat TM data, an improved classification rate for *Grassland, Recreation* and *Bare ground* was achieved. This improvement may be spatially or spectrally related.

The Producer's and User's Accuracies for *Water* are both low, however inspection of Table 5.9 indicates that most errors result from confusion with *Wetland* which may have occurred due to seasonal water level changes or minor misregistration of the satellite or reference data.

Interpretation of the Level II data indicates the *Residential* and *Pine plantation* classes are well defined with relatively high Producer's and User's Accuracies. Consideration of Table 4.9 indicates that class difference assessment is not a perfect predictor of class confusion for *Residential.* However it does provide a useful guide to the outcome in the case of *Pine plantation,* where Table 4.9 confirms difficulty in separating *Wetland, Woodland* and *Grassland.*

The results for all other classes in the supervised classification are of very low quality. Producer's Accuracies range from 8.6 percent to 53.1 percent, and User's Accuracies from 15.2 percent to 54.6 percent. Some classes are undersampled as a result of the early stage of rural to urban land cover change, which effects the validity of the sampling of the *Wetland, Recreation* and *Water* classes.

The large IFOV of the Landsat MSS sensor results in a large degree of within-pixel averaging that restricts the ability of the classifier to clearly identify targets which have some similar spectral components such as *Residential under construction, Grassland, Pine plantation (low density)* and *Bare ground.* In particular, Table 4.9 confirms the difficulty in separating *Residential under construction* and *Bare ground,* and *Pine plantation (low density)* and *Grassland.*

5.4.2.2 Landsat TM

Results for supervised classification in Tables 5.11 and 5.12 indicate *moderate* to *excellent* agreement at Level I and Level II, with Kappa Coefficients of 62.5 percent and 59.1 percent, respectively. The *Urban* class has high Producer's and User's Accuracies with most confusion occurring with the *Open* class. Only seven percent of the *Urban* reference data are classified as *Open*, however there is a 14 percent commission rate of *Open* pixels in the *Urban* classification. This indicates the *Urban* signatures overlap with the characteristics of *Open* pixels, and the extent of the variability of the *Open* class is not taken into account by the *Open* spectral signatures. The moderate Producer's and User's Accuracies for *Forest* indicate it is subject to large variability as evidenced by the high level of errors of omission and commission with the *Urban* and *Open* classes. Once again, the *Water* class suffers from confusion with *Forest* and inspection of Table 5.12 indicates this is caused specifically with *Wetland* pixels.

At Level II the *Residential* class is well defined with Producer's and User's Accuracies of 83.5 and 77.2 percent, respectively. Patterns of omission and commission errors amongst classes are similar indicating that signature refinement is unlikely to improve the results. *Commercial* areas have moderate levels of misclassification with the *Residential, Pine plantation* and *Woodland* classes as predicted in class significance testing in Table 4.10. High errors of commission and moderate levels of omission for *Pine plantation* are mainly caused by confusion with the *Grassland* class. Low density pine forest often represents a transition stage between *Grassland* and high density pine forest and, with similar scene components but different proportions, confusion may result.

The definition of *Wetland* areas has improved from that achieved with the supervised classification of the Landsat MSS data, however not to an extent that is considered reliable. Confusion is distributed amongst all classes, with errors of commission for *Woodland* representing the only major source. *Woodland* often occurs adjacent to *Wetland*, therefore registration errors and land cover transition effects may be significant. *Woodland* displays relatively low Producer's and User's Accuracies, with the largest errors of omission occurring with *Residential, Pine plantation* and *Wetland*, while most errors of commission occur with *Residential* and *Grassland*. A similar level of misclassification, but with different classes, occurred when an unsupervised approach

Classified Data	Reference Data					User's Acc.
	Urban	Forest	Open	Water	Total	%
Urban	**561**	37	98	0	*696*	*80.6*
Forest	31	**145**	65	6	*247*	*59.7*
Open	48	32	**378**	0	*458*	*82.5*
Water	1	2	0	**4**	*7*	*57.1*
Total	*641*	*216*	*541*	*10*	*1408*	
Prod. Acc.	*86.4*	*66.5*	*69.4*	*40.0*		
					Overall Classification Accuracy (percent)	*77.3*
					Kappa Coefficient ($\hat{K}$) (percent)	*62.5*

Table 5.11 *Error matrix for supervised classification at Level I using Landsat TM data*

Classified Data	Reference Data										User's Acc.
	Res	Com	Pine	Wet	Wod	Gras	Rec	Bare	Wat	Tot.	%
Residential	**444**	19	6	2	16	39	11	38	0	*575*	*77.2*
Commerc	19	**79**	4	2	7	6	1	3	0	*121*	*65.3*
Pine plant	7	5	**44**	3	11	35	8	2	0	*115*	*38.3*
Wetland	1	0	5	**22**	15	2	2	0	5	*52*	*42.3*
Woodland	12	6	5	3	**37**	11	4	1	1	*80*	*46.2*
Grassland	15	3	11	1	8	**190**	3	16	0	*247*	*76.9*
Recreation	7	5	0	3	2	8	**55**	0	0	*80*	*68.8*
Bare grnd	18	0	3	0	4	12	2	**92**	0	*131*	*70.2*
Water	1	0	0	2	0	0	0	0	**4**	*7*	*57.1*
Total	*524*	*117*	*78*	*38*	*100*	*303*	*86*	*152*	*10*	*1408*	
Prod. Acc.	*83.5*	*67.5*	*56.4*	*56.4*	*36.6*	*62.5*	*62.5*	*60.1*	*40.0*		
										Overall Classification Accuracy (percent)	*68.7*
										Kappa Coefficient ($\hat{K}$) (percent)	*59.1*

Table 5.12 *Error matrix for supervised classification at Level II using Landsat TM data*

was used. This suggests the spectral characteristics are variable and difficult to specify, either through clustering or supervised signature extraction, and improvement in the class accuracy may not be achieved through signature refinement.

Producer's and User's Accuracies for *Grassland*, *Recreation* and *Bare ground* are all similar within the range 60 to 76 percent. Table 4.10 indicates *Grassland* and established *Residential* areas are difficult to separate and this is borne out by the classification results. Mutual errors of commission occur with the *Residential* and *Grassland* classes, and signature refinement to decrease omissions would also increase commissions, and *vice versa*. The spectral similarity of residential gardens and *Grassland* may be the cause of this confusion. Large errors of omission also occur with *Pine plantation* and the confusion may result from the similarity of the grasses that form the understorey of pine forest at low tree densities with the signatures of the *Grassland* class. *Recreation* has maximum confusion (omission) with *Residential*, which may be due to the spectral similarity of mature residential areas and recreation areas. *Recreation* areas are also commonly found in isolated patches within the *Residential* class and misregistration may be significant.

Table 4.10 indicates almost all Landsat TM bands are significantly different for *Bare ground* and *Residential*, but classification results in Table 5.12 show confusion of *Bare ground* pixels with the *Residential* signatures. High rates of omission and commission for *Bare ground* occur for *Residential* and lesser levels for *Grassland*. The error rates for *Grassland* may be due to the presence or absence of grasses at the time of sensing, which is the major differentiating factor between *Bare ground* and *Grassland*. Confusion with *Residential* is most likely due to the transitional nature of new residential areas between *Bare ground*, and more established and heavily vegetated established *Residential* areas.

5.4.2.3 SPOT HRV

Supervised classification results for SPOT HRV data are reported in Tables 5.13 and 5.14. The Level I and Level II Kappa Coefficients are 50.8 and 47.1 percent respectively, and indicate a *moderate* level of agreement with the reference data. Previous comments in Section 5.4.1.3 with respect to the spatial complexity and distribution of land covers in the context of the spatial resolution of the SPOT HRV

Classified Data	Reference Data					User's Acc.
	Urban	**Forest**	**Open**	**Water**	**Total**	*%*
Urban	**808**	21	50	0	*879*	*91.9*
Forest	74	**108**	14	0	*196*	*55.1*
Open	152	23	**152**	1	*328*	*46.3*
Water	1	1	0	**8**	*10*	*80.0*
Total	*1035*	*153*	*216*	*9*	*1413*	
Prod. Acc.	*77.4*	*70.1*	*69.4*	*88.9*		
					Overall Classification Accuracy (percent)	*75.6*
					Kappa Coefficient $(\hat{K})$ (percent)	*50.8*

Table 5.13 *Error matrix for supervised classification at Level I using SPOT HRV data*

Classified Data	Reference Data										User's Acc.
	Res	**Com**	**Pine**	**Wet**	**Wod**	**Gras**	**Rec**	**Bare**	**Wat**	**Tot.**	*%*
Residential	**660**	35	6	1	8	23	15	4	0	*752*	*87.8*
Commerc	32	**81**	2	3	1	3	1	4	0	*127*	*65.8*
Pine plant	43	3	**25**	5	34	7	3	0	0	*120*	*20.9*
Wetland	1	0	0	**19**	2	0	0	0	0	*22*	*86.4*
Woodland	20	7	4	4	**15**	2	2	0	0	*54*	*17.8*
Grassland	79	6	4	3	10	**59**	0	6	0	*167*	*35.3*
Recreation	14	5	1	2	2	5	**62**	0	0	*91*	*68.1*
Bare grnd	46	2	1	0	0	5	0	**15**	1	*70*	*21.4*
Water	1	0	0	1	0	0	0	0	**8**	*10*	*80.0*
Total	*896*	*139*	*43*	*38*	*72*	*104*	*83*	*29*	*9*	*1413*	
Prod. Acc.	*73.1*	*57.4*	*58.1*	*50.0*	*20.5*	*56.2*	*73.8*	*50.0*	*88.9*		
										Overall Classification Accuracy (percent)	*66.8*
										Kappa Coefficient $(\hat{K})$ (percent)	*47.1*

Table 5.14 *Error matrix for supervised classification at Level II using SPOT HRV data*

sensor, also apply to the results of the supervised classification. *Urban* dominates the Level I classification with all Producer's Accuracies at least 70 percent and User's Accuracies varying between 46 and 92 percent. Errors of omission of approximately 15 percent for *Urban* are classified as *Open*, and are possibly due to the higher resolution sensor identifying areas of grass, recreation or bare ground within otherwise *Residential* areas. Similar situations arise for the other classes within the classification and it may represent the limit of classification accuracy for the available level of detail contained within the reference data. Such a condition also indicates the spatial resolution of the sensor is approaching the level of the *support* (10 - 15 m) for some of the urban target components such as houses and roads (see Section 4.3).

The User's Accuracy for *Urban* is very high at 91.9 percent. Approximately five percent of the errors of commission for the *Urban* class are pixels from the *Open* class, and are again possibly caused by urban-like targets within the *Grassland*, *Recreation* or *Bare ground* areas. User's Accuracies for *Forest* and *Open* are low at 55.1 and 46.3 percent, respectively. In both cases *Urban* accounts for at least 40 percent of the errors of commission, which are the same pixels discussed above as errors of omission for the *Urban* class. Refinement of the signatures may not improve the quality of the classification, as a resolution of 20 m is now capable of detecting individual features that were of subpixel size for the Landsat MSS data. Signature refinement to make *Urban* more inclusive may also result in increased errors of commission with the Open class because many small *Urban* areas will also have the characteristics of the *Open* class. Alternatively, the detail in the reference data may need to be refined to account for the increased discriminating power of higher resolution sensors such as Landsat TM and SPOT HRV.

Apart from *Residential* and *Water*, the Producer's and User's Accuracies for all classes at Level II are variable, with values from moderate to very low. Most of the classification errors are concentrated as errors of omission from the *Residential* class, and as commissions of most other classes. The overall number of pixels in the *Residential* class is high (63 percent of the study area), therefore even a relatively low rate of omission translates to high rates of commission and low User's Accuracy for classes with comparatively small numbers of reference samples. For example, by ignoring the errors of omission for *Residential*, the User's Accuracy for *Grassland*

changes from 35.3 percent to 67.0 percent. No other classification method applied in this study has produced such consistent errors within a single class, as has *Residential* for the supervised classification of the SPOT HRV data. In situations where there are large discrepancies between the number of pixels in each class the User's and Producer's Accuracies become less relevant as diagnostic measures of classifier performance.

Comparison of the results in Table 5.14 with those in Table 5.8 derived from the same data using the ISODATA algorithm, and based upon automated formation of clusters, indicates the Producer's Accuracy has decreased from 99.0 percent to 73.1 percent, and the User's Accuracy has increased from 71.9 percent to 87.8 percent. Adjustment of spectral signatures to reduce the rate of errors of omission for *Residential* may be undertaken, but this will increase the rate of commission of *Residential* and substantially decrease the classification accuracy of all other classes.

5.4.3 Comparison of Classification Results

Results of individual approaches to classification have been analysed in Sections 5.4.1 and 5.4.2. These analyses consider the level of agreement, measured through the Kappa Coefficient, between the classified data and stratified random sampling of the reference data. The results provide a comparison at Level I and Level II classifications for each of the multiscale datasets. The extent of interclass confusion was evaluated with a view to understanding the patterns of classification error, to assist in the selection of an appropriate level of detail for classification and to determine the algorithm most likely to provide a classification accuracy suitable for change assessment. Consideration of the impact of factors contributing to the outcome of the classification results will be made in Section 5.5.

The relative importance of the Kappa Coefficient and Overall Classification Accuracy for assessing the quality of classification results has been discussed in Section 5.2. Tables 5.15 and 5.16 provide a comparison of the Kappa Coefficient values for each dataset at Level I and Level II, respectively. The standard normal deviate of the Kappa Coefficient is used to evaluate whether a significant difference exists between the results of the supervised and unsupervised classifications and, where none exists, the classification with the highest Overall Classification Accuracy will be utilised in further

change detection analyses. Where a significant difference is detected the classification result with the higher Kappa Coefficient will be used for analysis. Class-based comparisons of classification could also have been utilised in the comparison, however for statistical analysis a single measure such as the Kappa Coefficient is required.

	Landsat MSS		**Landsat TM**		**SPOT HRV**	
Classification	Unsuper	Super	Unsuper	Super	Unsuper	Super
Overall Class. Acc.	79.1	75.4	75.0	77.3	81.6	75.6
$\hat{K}$	57.3	56.2	58.7	62.5	50.4	50.8
$\hat{\sigma}^2_{\hat{K}}$	0.000835	0.000364	0.000928	0.000307	0.001379	0.000495
$Z_{\hat{K}_1 - \hat{K}_2}$	0.32		1.08		0.09	
Significance *	*NS*		*NS*		*NS*	

** NS = not significant, S = significant*

Table 5.15 *Comparison of Level I classification results at the 95 percent confidence level*

	Landsat MSS		**Landsat TM**		**SPOT HRV**	
Classification	Unsuper	Super	Unsuper	Super	Unsuper	Super
Overall Class. Acc.	61.6	58.4	61.2	68.7	70.1	66.8
$\hat{K}$	45.9	46.3	46.9	59.1	36.9	47.1
$\hat{\sigma}^2_{\hat{K}}$	0.000279	0.000241	0.000296	0.000234	0.000517	0.000337
$Z_{\hat{K}_1 - \hat{K}_2}$	0.18		5.23		3.49	
Significance *	*NS*		*S*		*S*	

** NS = not significant, S = significant*

Table 5.16 *Comparison of Level II classification results at the 95 percent confidence level*

Examination of the classification results at Level I in Table 5.15 indicates that at the 95 percent confidence level all computed values of the standard normal deviate of the Kappa Coefficient (Equation 5.3) are less than 1.96. Therefore, no significant differences are detectable between the classification results achieved by the supervised or unsupervised approaches for any of the datasets. Consequently, for the Landsat MSS and SPOT HRV data, results of the unsupervised classification will be utilised, while for the Landsat TM data the supervised classification provides the most favourable outcome as each has recorded the highest Overall Classification Accuracy.

Classification results for Level II are shown in Table 5.16 and indicate at the 95 percent confidence level the standard normal deviate of the Kappa Coefficient is less than 1.96 only for the Landsat MSS data. No significant difference is detectable between the classification results and only the unsupervised classification will be considered further. For both the Landsat TM and SPOT HRV classifications a significant difference was detected between the results of the two algorithms with the supervised classification providing superior results in each case.

The main contributing factor to the observed difference between the supervised and unsupervised classifications of the Landsat TM and SPOT HRV data is the failure of the clustering process to recognise some land cover classes. In Table 5.6 no clusters were allocated to the *Pine plantation* class for Landsat TM data, and in Table 5.8 no clusters were allocated to *Pine plantation, Wetland, Grassland* or *Bare ground* classes during unsupervised classification of the SPOT HRV data. The selection of reference classes is based on an *a priori* determination of land covers present in the area, and failure to identify pixels sampled from these reference classes results in these pixels being recorded as errors of omission and commission for other classes.

In both cases for the SPOT HRV data the Kappa value for the unsupervised classification is less than the value for supervised classification, yet the Overall Classification Accuracy is substantially higher. The apparent discrepancy in the variation of these values is caused by the dominance of the *Residential* class in the classification. The *Residential* class unduly influences the Overall Classification Accuracy, which does not consider class distributions apart from correctly classified pixels located on the main diagonal of the error matrix. The Kappa Coefficient is

specifically designed to provide a more comprehensive assessment of classification accuracy by considering the off-diagonal values and their effect on the classification result (see Section 5.2.2). Consequently, the Kappa values are substantially lower for Landsat TM and SPOT HRV data at Level II compared to the corresponding Overall Classification Accuracy values.

Due to the capability of the supervised classification of Landsat TM and SPOT HRV data at Level II to provide significantly higher quality results than the unsupervised classification, only these data will be considered during subsequent analysis.

5.5 Analysis of Factors Contributing to Classifier Performance

The foregoing analysis provides an assessment of supervised and unsupervised classifier performance for a range of multiscale data at two levels of classification detail. Following statistical analysis of the results the optimum classification algorithm for each dataset at Level I and Level II has been determined. The following section discusses a range of factors that have contributed to the performance of the algorithms and represent important considerations relevant to the outcomes of this research.

5.5.1 Radiometric and Spectral Resolution

Examination of classification results in Tables 5.3 to 5.14 demonstrates that *Urban* and *Residential/Commercial* classes at Level I and Level II respectively, have dominated the classification with large numbers of pixels and relatively high classification accuracies. Other classes generally contain fewer pixels and achieve lower Producer's and User's Accuracies. Discussion of the supervised and unsupervised classifications in Section 5.4 provides a detailed analysis of the patterns of omission and commission relevant to each algorithm. Consideration of the spectral separation of classes is important to understand these classification patterns.

Figure 5.1 illustrates the distribution of pixel values for Landsat TM data in bands 3 and 4, and plots the digital ranges of each land cover class derived from the training data of major land cover types used during supervised classification. Only the major land cover types are shown in order to preserve clarity in the diagram, and the *Residential* class is represented by two sets of training data (*viz Established residential* and *Residential under construction*). Landsat TM data were selected for illustrative purposes as they

provide the broadest range of spectral sensitivities, and bands 3 and 4 are useful for identification of urban and vegetation targets (Jensen, 1996).

Within Landsat TM bands 3 and 4 clear differences between most training data can be observed, except for *Commercial* and *Established residential/Grassland*, where substantial overlap is present. The limits of the training data shown in Figure 5.1 are represented by the intraclass variances from which class membership probabilities are computed. Therefore, for a given distance from any class mean, the probability of belonging to a class is inversely proportional to the class variance. These data must also be considered in the context of the other four bands of Landsat TM data that will also be used to compute class membership probabilities for the maximum likelihood classification.

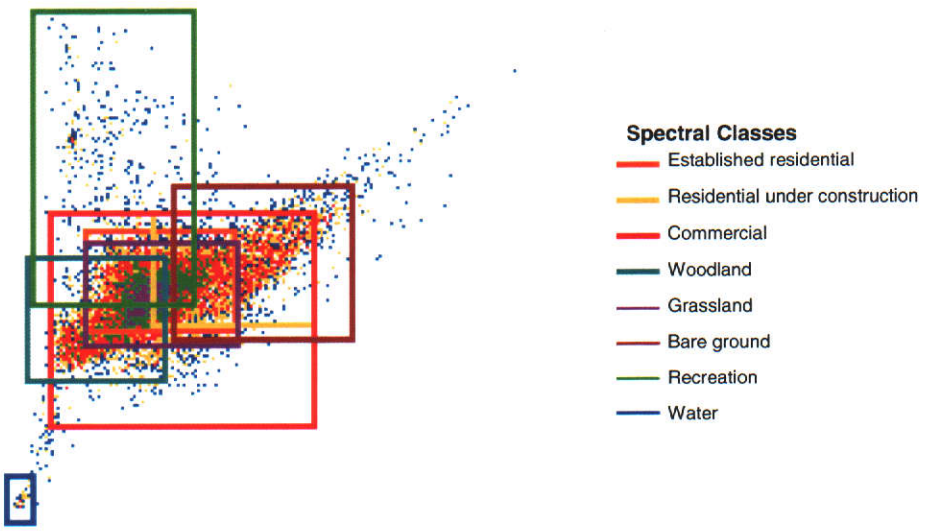


Figure 5.1 *Scatter plot of Landsat TM band 4 (y axis) and band 3 (x axis) showing major class signatures centred on mean value and bounded by two standard deviations (extents of the scatter plot are normalised to the radiance values in each band)*

The rectangle representing the *Commercial* class overlaps a large proportion of most other classes and demonstrates a large class variance. Many pixels in this class are correctly classified and the User's and Producer's Accuracies for supervised classification of the Landsat TM data are both moderately high (65 percent), but due to the relatively large class variance, there are many errors of omission and commission

(35 percent). Even though *Residential* (combination of *Established residential* and *Residential under construction* training data) falls largely within the variance range defined by *Commercial*, it displays the highest classification accuracy of any class because of its low class variance.

Grassland, which has a mean and variance in bands 3 and 4 that is almost coincident with the *Established residential* training data, still achieves a classification accuracy at least as high as the *Commercial* class due to the independence of the spectral characteristics in other bands. Plots of training data for *Bare ground*, *Woodland* and *Recreation* indicate a moderate degree of separation of the classes in these two bands, however large variances contribute to low classification accuracy for *Woodland* and moderate results for *Woodland* and *Recreation*. The other bands in the data set undoubtedly influence these results, however the scatter plot provides a guide to the understanding of class separation.

The data included in Figure 5.1 also demonstrate the degree of band-to-band correlation within the Landsat TM data for these land cover classes. Strong correlation between bands reduces the spectral separability of the classes and suggests that decorrelation of the data may have improved the results of information extraction. Preprocessing through PCA reduces the spectral correlation and improves the spectral separability of targets. These approaches to data preprocessing were not investigated during this research.

Figure 5.2 summarises the classification uncertainty derived from the maximum likelihood classification of the corresponding data represented in Figure 5.1. These results indicate that 33 percent of the classified pixels have a classification uncertainty (Eastman, 1997) of 50 percent or greater. This high level of uncertainty arises from class membership probability for pixels being spread across a number of classes, or low levels of class membership probability in all classes. The result is low classification accuracy and higher than expected levels of confusion between classes, as observed in the Level II classification results (Table 5.12). Decreased uncertainty may be achieved through refinement of training data or structure of the land cover classification system to identify classes with improved spectral separation. Steele *et al* (1998) indicate that misclassification probability estimates (classification uncertainty) may be useful for

analysing the spatial distribution of misclassification errors in thematic land cover maps, however this potential has not been explored further.

Direct comparison of classification accuracies achieved across the datasets analysed in this research is difficult not only due to the range of spatial and spectral sensitivities of the various sensors, but also due to the variations in spatial diversity of the land cover patterns and their corresponding spectral complexity. Examination of the data in Tables 5.15 and 5.16 indicates that, for classifications at Level I and II, Kappa Coefficients for classification of Landsat TM data are greater than those for any of the other datasets. According to Basham May *et al* (1997), this is due to the presence of additional infrared bands, especially compared to the higher spatial resolution SPOT HRV data. Further research by Basham May *et al* (1997) also indicates that the higher spectral resolution of Landsat TM data is more important than higher spatial resolution when compared with Landsat MSS data.

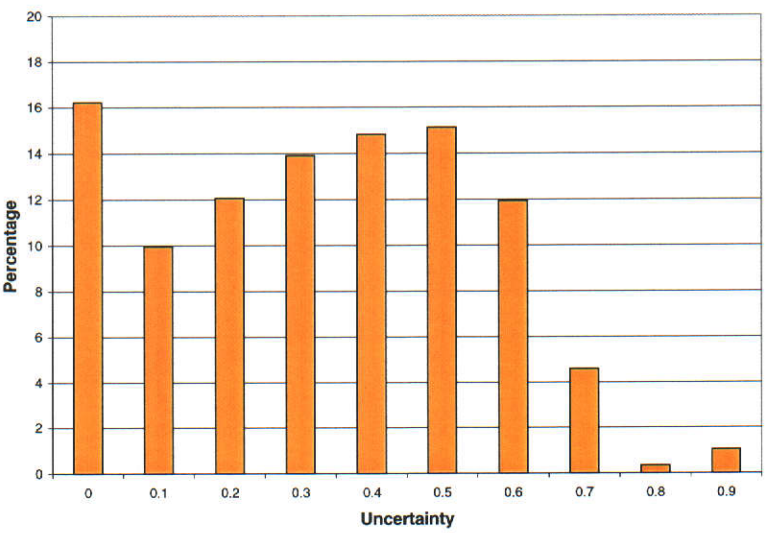


Figure 5.2 *Graph of classification uncertainty at Level II for Landsat TM data using supervised classification*

5.5.2 Spatial Resolution

Contrasting factors associated with the finer spatial resolution of remotely sensed data and its effect on classification are identified by Irons *et al* (1985). The two opposing factors affecting the data are increased within-class spectral (and radiometric) variability and decreased proportion of mixed pixels. The former situation decreases, while the

latter tends to increase classification accuracy. Verification in this study of the results of Irons *et al* (1985) is difficult due to the change in land cover distribution between sensing dates. However, improvements of a similar magnitude (10 percent) were achieved with the Level II supervised classification between Landsat MSS and Landsat TM data, even though a more complex pattern of land covers was present in the Landsat TM data (Table 5.16). The same magnitude of improvement was not observed at Level I or for the unsupervised classification, and no improvement was noted when the SPOT HRV data were classified, possibly due to the reduced number of bands available for analysis (see Section 5.5.1).

5.5.3 Reference Data Issues

The influence of reference data quality on the assessment of classification accuracy of land covers depends on the thematic accuracy of the reference data and the boundary effects which arise due to the spatial complexity of the land covers contained in the reference data. Reference data are assumed to be error free, however even for data collected directly in the field, this is often not the case (Congalton, 1991; Kalkhan *et al*, 1998). Reference data derived from aerial photographs are subject to even greater errors and may lead to erroneous evaluations of classification accuracy.

Reference data were collected for this study from medium scale panchromatic (1972 and 1986) and colour (1992) aerial photographs, which provide good discrimination of most land covers. Confusion occurred between *Bare ground* and *Grassland* because the presence/absence of herbaceous vegetation on panchromatic photographs was difficult to detect. Evaluation of the influence of photointerpretation errors on reference data quality was not possible as the satellite data were collected as early as 1972, and the most recent data were recorded in 1991. Aerial photographs represent the most reliable source of data available for evaluation of the classification results.

Boundary effects in remotely sensed data are mainly related to the interpretation of mixed pixels, determination of class boundary locations and reference data verification errors. Specialist image enhancement algorithms that deal with pixel unmixing and fuzzy classification approaches for pixels that demonstrate multiple class membership are necessary for resolution of issues related to mixed pixels, and are outside the scope of this study. Fuzzy classification may also be applied to the determination of class

boundaries, or it may be achieved through refinement of training data and classification approaches to provide an optimum level of thematic data extraction and therefore boundary determination. Uncertainties in the location of boundaries will remain, but their effect will be measured as part of the classification accuracy assessment process. This section considers the effect of spatial complexity of reference data boundaries and its effect on thematic classification assessment. Apparent misclassification may be recorded where displacements in reference data occur due to misregistration. The effect is most pronounced when land cover classes are fragmented and the boundaries exhibit a high degree of complexity (Kalkhan *et al*, 1998).

Reference data boundary complexity was studied using a shape index known as the *Compactness Ratio*, which relates the area of a feature to its perimeter (Haggett *et al*, 1977). The compactness ratio used in this research compares the area of the polygon and the area of a circle having the same perimeter as the polygon under examination, and is defined as follows (Eastman, 1997):

$$Compactness\ Ratio = \sqrt{Polygon\ area / Circle\ area\ with\ equivalent\ perimeter} \quad (5.6)$$

The compactness ratio was computed from the reference data for each sensing epoch as the classified data contained too many small polygons that produced an artificially low result (Table 5.17). According to comparisons made by Lo and Watson (1998), the average compactness ratio derived from the reference data images are all relatively low, which indicates the land cover patterns are fragmented rather than compact. Values for 1986 and 1991 are lower than those for 1972 and correspond to the increasing spatial complexity of the land covers in the area. Decreased classification accuracy usually results from increased fragmentation, however the improved spatial resolution and increased data quantisation of the Landsat TM and SPOT HRV data counteract these effects. Substantial improvement in the number and distribution of spectral bands within the Landsat TM data also offsets against the negative influence of scene fragmentation on classification accuracy. It is likely that with a structure of classes similar to those present in 1972, the Landsat TM and SPOT HRV data could produce significantly higher rates of classification accuracy than achieved in this study.

Date	Classified Data	Compactness Ratio
1972	Landsat MSS	0.2361
1986	Landsat TM	0.1803
1991	SPOT HRV	0.1767

Table 5.17 *Average compactness ratio of land cover classes for reference data in 1972, 1986 and 1991*

5.5.4 Spatial Autocorrelation of Error Patterns

The pattern of error in classified multispectral images is dependent upon a range of factors including topographic effects, target spatial, spectral and radiometric variability, processing approaches, and the spatial, spectral and radiometric resolution of the sensor. Analysis of error patterns is useful in identifying the source of classification errors and in the development of strategies to increase the quality of information extraction. Errors that occur in blocks are likely to represent gross misclassification of pixels related to errors in spectral class definition or reference data extraction, while fragmented or linear errors are related to boundary errors caused by misregistration of the remotely sensed or reference data (Congalton, 1988b).

Error patterns may be measured using spatial autocorrelation analysis. Techniques of spatial autocorrelation analysis similar to those applied to the assessment of intraclass variability in Section 4.3 will be applied to difference images derived from a comparison of the classified image and the reference data. *Moran's I* will again be used to assess the degree of spatial autocorrelation (see Section 4.3). Figure 5.3 contains difference images for each of the datasets classified at Level II using an unsupervised classification algorithm. The remotely sensed data were collected at a range of spatial resolutions, but were resampled and analysed at a consistent resolution of 20 m. Values of *Moran's I* were computed for each difference image at 20 m intervals to a total of 200 m, and are plotted in Figure 5.4. Spatial autocorrelation values may be converted to equivalent spatial lags (in units of original pixels) by dividing by 60 m for Landsat MSS, 30 m for Landsat TM and 20 m for SPOT HRV. The resulting maximum lags at 200 m are then three (Landsat MSS), six (Landsat TM) and ten (SPOT HRV).

The full extent (total number of lags) of statistically significant spatial autocorrelation was not assessed for these data and all lags examined were found to be significant at the 99 percent confidence level. Given the large proportion of classification errors in the data, and that the errors were not categorised by cover class, this result is not surprising. Comparison of the degree of autocorrelation between the datasets and in conjunction with observed patterns of errors in Figure 5.3 is more useful.

(a) Landsat MSS (b) Landsat TM (c) SPOT HRV

Figure 5.3 *Image showing differences between classified and reference data (all data are resampled to 20 m; White = correct, Black = error) (Scale 1:80 000)*

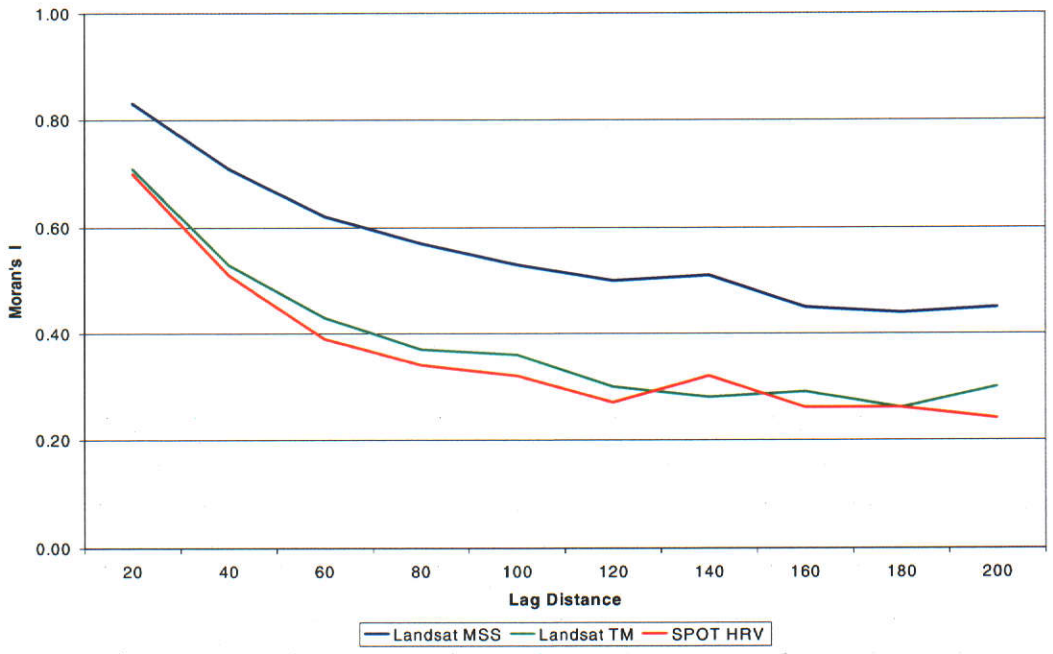


Figure 5.4 *Spatial autocorrelation of classification errors for multiscale satellite data (lag distances are in metres)*

Examination of Figure 5.4 indicates at equivalent distances, but at lower spatial lags, the coarser spatial resolution Landsat MSS data have substantially higher levels of spatial autocorrelation than the other two datasets. This trend is maintained over the complete range of spatial autocorrelation measurements for these data. This pattern of response is expected due to the generalisation produced through pixel averaging by the large IFOV of the Landsat MSS, but also reflects the nature of the classification errors observed in Figure 5.3a. The classification errors occur in large blocks that represent misclassification of complete regions, due mainly to the large IFOV of the sensor. Some of these regions represent classes that occupy areas too small or too narrow to be identified at this resolution, while some regions represent omission of complete classes. An investigation of the class composition of these errors has not been undertaken.

The spatial autocorrelation values for the Landsat TM and SPOT HRV data are very similar despite the IFOV of the Landsat TM being 50 percent larger than the SPOT HRV sensor. This may be caused by the spatial resolution of both of these sensors being close to the level of *support* of the targets (see Section 4.3), especially for urban targets which dominate the 1986 and 1991 scenes recorded by these two sensors. The IFOV of the sensors provides sufficient generalisation to enable reasonably consistent responses, but is not so fine as to increase class variance by identifying individual objects such as trees and house roofs. The fragmented nature of errors identified in Figures 5.3b and 5.3c also indicates that most errors occur along boundaries or along linear features, which are likely to produce lower rates of spatial autocorrelation (Congalton, 1988b), as evident in Figure 5.4. This pattern of errors indicates occurrence of mixed pixels along boundaries or the presence of data misregistration are major factors affecting classification accuracy.

5.6 Summary

Investigations in this chapter have been directed towards the evaluation of supervised and unsupervised classification approaches for the identification of various land cover classes that have evolved during the period 1972 to 1991. The impact of changing land cover patterns and evaluation of information extraction approaches relevant to the available multiscale data have been important considerations in determining the direction of the research.

Descriptive and analytical accuracy assessment methods are available for evaluation of classified data. Descriptive methods have been applied over a long period, however simple statements of Overall Classification Accuracy and User's and Producer's Accuracy are insufficient for comprehensive analysis of classification results. Analytical approaches such as the Kappa Coefficient provide statistically sound techniques, which summarise all elements of the error matrix, and compute an accuracy value and its variance, taking into consideration the probability of chance agreement. Important in this study is the ability for the Kappa Coefficient to statistically compare the results of a range of classification algorithms to determine an optimum approach.

Sampling schemes for error assessment are considered an important component of classification analysis. Most studies indicate that stratified random sampling is the most robust approach, however research by Stehman (1999) determined that systematic sampling theoretically provides an optimum solution marginally better than stratified random sampling, but is contingent upon the absence of periodicity in the data. In the light of strong empirical evidence by Lo and Watson (1998), and the inherent difficulties in identifying periodic patterns in the classified data, a stratified random sampling scheme was adopted for this research.

Overall image sampling rates are determined based on the binomial theorem, but determination of the level of class sampling rates are subject to conjecture and are important where class sizes and spatial distribution are variable within the scene. Research indicates minimum samples between 30 and 75 samples per class are acceptable, with the exact size determined largely by the logistics of reference data collection. A value of 75 samples per class was used in this research because a reference map of the complete study area was compiled using aerial photography, and sampling was undertaken using automated sampling routines.

Classification of the multiscale satellite data was performed at modified Level I and Level II (Anderson *et al*, 1976) using both supervised and unsupervised algorithms. At Level I no significant difference between the two algorithms was detectable for any of the datasets. At Level II no significant difference was observed between the supervised and unsupervised approaches for the Landsat MSS data, however for Landsat TM and SPOT HRV data, the results of the supervised classification were assessed as

significantly higher compared to the unsupervised classification. The difference was caused because not all land cover classes were identified by the unsupervised classifier, which resulted in the lower classification accuracy for these data. Kappa Coefficient values for Level I showed *moderate* to *substantial* (50.4 to 62.5 percent) levels of agreement, while for Level II *moderate* (46.3 to 59.1 percent) levels of agreement were achieved.

Classification of the spectrally diverse class of *Urban* was consistently high at Level I, with the *Forest* and *Open* classes moderately high but variable. In all cases the *Water* class was too small to be reliably assessed. At Level II the *Residential* class was most reliably determined in all datasets, with all other classes showing highly variable results. On an area basis, *Pine plantation* dominated the Landsat MSS data, while the *Residential* class dominated the Landsat TM and SPOT HRV data. Both classification approaches accommodated the spectral variability of the *Residential* class, however further research may be required to refine the classification of other information classes in the images.

The characteristics of land cover within the study area underwent substantial transformation from predominantly rural to predominantly urban during the period under study, and limited the ability of the higher resolution satellite data to improve the classification accuracy. Landsat TM and SPOT HRV data normally produce higher classification accuracies, however no substantial improvements were observed. Landsat TM produced the highest Kappa Coefficient for both Level I and Level II classifications, but this is due more to the superior spectral resolution of the data than to effects of finer spatial resolution.

Evaluation of the patterns of classification errors through spatial autocorrelation and analysis of the compactness of the land cover classes indicates that spatial complexity and associated boundary identification are major factors affecting classification in the Landsat TM and SPOT HRV images. The substantial increase in spatial complexity of the land cover patterns mitigated against any improvement in classification accuracy resulting from finer spatial resolution.

Analysis of classification uncertainty of the supervised classification established up to one third of the pixels in the Landsat TM data being classified with an uncertainty of 50 percent or more, indicating low confidence in many class allocations and potential for misclassification of these pixels. Inability of the unsupervised classification algorithm to recognise some land cover classes reinforces this finding.

Chapter 6

MULTISCALE CHANGE ASSESSMENT

Detection of changes in land cover patterns using remote sensing approaches may be achieved in a number of ways depending on the characteristics of the source data and the targets, and the processing facilities available. Image processing approaches to change detection will be discussed to highlight significant considerations in the process and detail the change detection approach. Evaluation of the quality of land cover change detection processes will be undertaken through verification against reference data. Graphical and numerical reporting methods for communicating the results of land cover change will be investigated and used to develop summarisation techniques relevant to the representation of rural-urban change. Quantitative assessment of the accuracy of land cover change provides an important metric of the significance and reliability of the measured transitions between classes.

6.1 Introduction

Singh (1989) defines change detection as *the process of identifying differences in the state of an object or phenomenon by observing it at different times.* Change detection may extend over global or local scales, and be concerned with short or long term environmental variations. Each spatial resolution provides a different perspective, and the factors influencing surface patterns as measured at each resolution may reflect different underlying processes. A challenge exists in remote sensing analysis to develop relevant nesting of observations at multiple space and time scales in order to satisfy the multiple objectives of short-term, fine scale measurements and long-term, broad scale measurements (Davis *et al*, 1991).

Detection of land cover change in satellite imagery is complicated by many adverse temporal factors. These include differences in bandpasses and spatial resolutions, spatial misregistration, variations in the radiometric responses of the sensors, differences in the distribution of cloud and cloud shadow, variations in solar irradiance and solar angles, and differences in phenology (Yuan and Elvidge, 1998). The basic premise in using these data for change detection is that changes in land cover result in variations in radiance values, and that variations in radiance due to land cover changes are large with respect to variations caused by other system and environmental factors not related to change (Mas, 1999).

Many change evaluation studies suggest imagery be acquired at or near anniversary dates and similar times of the day in order to minimise variations caused by illumination and seasonal factors (Jensen *et al*, 1997). Vegetation targets that undergo significant phenological changes, or targets subject to other seasonal variations, such as snowfall, are especially susceptible. Such control is not always possible and selection of processing algorithms and data characteristics must attempt to take this into account.

Implementation of change detection techniques using data acquired from different sensors is more problematic. Hill and Aifadopoulou (1990) indicate the use of data from different operational satellites may provide a solution to overcome data availability problems, however the combination of data sources may also complicate or limit the analysis. Tokola *et al* (1999) investigated analysis of multisensor Landsat MSS and Landsat TM data for change detection using image classification followed by

image differencing, and found the quality of data varied over time. Landsat MSS data were subject to degradation of the optics, problems with altitude control and coherent noise in the data, while Landsat TM suffered from uncertainties in radiometric within-scene calibration on some bands. Many of these parameters may be controlled through calibration procedures, but careful assessment of all scene dependent and acquisition and processing variables is required to optimise the capabilities for change detection in these circumstances.

Image normalisation in the data preprocessing stage usually improves the results of change detection (Hall *et al*, 1991). However, even use of sophisticated calibration approaches as an attempt to achieve image standardisation may not always result in an improvement in the radiometric quality of the data (Collins and Woodcock, 1996a). On the other hand, reliance on a single sensing system may also lead to compromise. Specification of a single dataset for the development of the North American Landcover Characterisation (NALC) program for the period 1972 – 1992 (Lunetta *et al*, 1998) has constrained the program to using Landsat MSS data. This occurred even though for half of this period 30 m resolution Landsat TM data were available, and for six years 20 m resolution SPOT HRV data were also recorded.

A wide range of approaches to change detection analysis have been reported (Fung, 1990; Green *et al*, 1994; Jeanjean *et al*, 1996; Jensen *et al*, 1993; Singh, 1989; and others), but four aspects of change detection are considered particularly important when monitoring the environment (Macleod and Congalton, 1998):

(i) Detecting that changes have occurred,

(ii) Identifying the nature of the change,

(iii) Measuring the areal extent of the change,

(iv) Assessing the spatial pattern of the change.

Evaluation of change detection approaches in Sections 6.2 and 6.3 will mainly consider detection and identification of the nature of the change, while the extent and representation of change will be considered in Sections 6.4 and 6.5.

6.2 Approaches and Processing Techniques

Image processing approaches for the purpose of change detection may be classified in a variety of ways, however no single taxonomy comprehensively describes all techniques. The approach used in this section focuses on the main techniques of change detection and excludes lesser known methods such as change vector analysis and image regression (Singh, 1989), and combined classification (Muchoney and Haack, 1994). Processing approaches for change detection are considered in the categories of image algebra, spectral transformation and postclassification analysis.

6.2.1 Image Algebra

This group of change detection techniques relies upon evaluation of pixel by pixel comparisons of coregistered multidate images to reveal changes between sensing epochs. Methods included in this range of techniques are image differencing and image ratios.

Image differencing is performed by subtracting the digital value of one image date for a particular band from the digital value of the corresponding pixel in the same band for a second image date. To eliminate the possibility of negative values a constant is usually added to the resultant difference image. This algorithm produces a series of new images with a difference distribution for each band (Singh, 1989). Pixels representing changes in land cover are located in the tails of the histogram, while unchanged pixels tend to cluster around the mean.

Image ratios are produced by dividing the digital value of one image date for a particular band by the digital value of the corresponding pixel in the same band for a second image date (Singh, 1989). A new ratio image is formed with pixels representing changes in land cover located in the tails of the histogram, while unchanged pixels tend to cluster around the mean value of unity.

A critical component of change assessment using image algebra techniques is the determination of threshold values for the resultant change images that separate change from no-change, or in some cases identify a range of change classes. Thresholds are specified on the basis of empirical assessments or deterministic measures of distances from the mean, generally quantified in units of standard deviation. Change classes

identified by density slicing the change image at standard deviation intervals as small as 0.1 have been applied Fung (1990), however the range of values must be empirically verified against reference data for the image area. Evaluation may be undertaken using standard thematic accuracy assessment approaches such as the Kappa Coefficient or Overall Classification Accuracy.

Image data utilised as input are generally normalised to reduce temporal atmospheric and seasonal illumination effects (Lillesand and Kiefer, 1994), and may be masked to eliminate known areas where change has not occurred, such as upland areas in coastal wetland studies. Source images may comprise single bands, ratios of bands, or vegetation indices developed from combinations of image bands. In some cases methods such as PCA (see Section 6.2.2) may also be utilised to derive images for input to the differencing or ratioing process (Singh, 1989).

Approaches using image arithmetic are straightforward in concept, easy to implement and allow for flexible specification of thresholds. Results for image differencing are reasonable and are often comparable with those achieved from the brightness component of PCA (Fung, 1990; Muchoney and Haack, 1994). Difficulties arise when comparing data from different sensing systems, and the imperfect nature of scene normalisation algorithms is exacerbated by the differences between sensors in terms of spatial, spectral and radiometric resolution.

6.2.2 Spectral Transformation

Spectral transformations are performed in image processing in order to generate new bands of data that represent an alternative description of the scene, and may therefore be more amenable to interpretation (Richards, 1993). The original and derived datasets are related via a linear operation and generally provide substantially the same information content in a reduced number of image dimensions.

Spectral transformations for change detection purposes may be implemented in two ways. Firstly, separate derivation of the spectral transformation of each image then comparison using either of the methods described in Sections 6.2.1 and 6.2.3 may be applied. Differencing or ratioing to highlight changes, or classification then comparison of each transformed image, may all be used to determine areas of change (Richards,

1984). Secondly, simultaneous analysis of bands from multiple dates may be undertaken to highlight change on the premise that multitemporal image data are highly correlated and that spectral transformations can be used to highlight differences due to change (Byrne *et al*, 1980).

The two most common approaches are PCA and the Kauth-Thomas tasselled cap transformation. Detailed descriptions of PCA are available in Gonzales and Woods (1993) and Richards (1993), and details of the Kauth-Thomas transformation are given in Richards (1993). The advantages of PCA are derived mainly from the ability to decorrelate the data and provide the majority of the spectral information in only two or three new spectral bands. This form of data summarisation is particularly important when separate PCA is undertaken using multidate imagery.

When simultaneous analysis of multidate imagery is undertaken, the first principal component is related to changes in overall scene brightness with the second component providing a measure of changes in overall scene greenness (Muchoney and Haack, 1994; Richards, 1984). Subsequent components may be interpreted as representing changes in features relevant to the particular scene. In a study of forest defoliation using Landsat TM data Muchoney and Haack (1994) indicate Components 3 and 4 are related to changes in vegetation cover, and clouds and cloud shadows, respectively. In an examination of bushfire scars on the landscape using multidate Landsat MSS data, Richards (1984) states that Components 3 and 4 both highlight recent bushfire scars. Component 3 also details scars up to one year old which are not evident in the fourth component, but revegetation since the earlier bushfire is obvious in the fourth component which is not found in the third.

The Kauth-Thomas transformation was developed initially for Landsat MSS data and is used to derive four new bands representing soil brightness, green vegetation, senescent vegetation and atmospheric effects (Jensen, 1996). Identical interpretations cannot be made for other sensors, however Hill and Aifadopoulou (1990) indicate that analogous results that are linearly related may be obtained from Landsat TM and SPOT HRV data. The method has been applied most successfully in agricultural applications where interpretations based on the balance between soil and vegetation responses are most informative.

The PCA and Kauth-Thomas approaches provide efficient methods of summarising multidate imagery for analysis as single or separate images, however interpretation by techniques such as classification is still required to extract the meaning of observed characteristics from the derived images, and identification of individual image components may be complex. Integration of combined multisensor and multidate data into a single interpretation process is not possible, and therefore requires separate processing and analysis.

6.2.3 Postclassification Analysis

Postclassification analysis of multidate images is conceptually the simplest approach for assessment of change detection. Each image is independently classified using a supervised or unsupervised algorithm, and then compared through crosstabulation to produce a change image. The approach may also be modified to utilise image enhancement to produce a mask for isolating general areas of change. The mask is used to eliminate areas of invariant land covers from further consideration and has the effect of minimising errors of commission for the change classes (Pilon *et al*, 1988).

The spectral, radiometric and spatial independence of the data processing permits the comparison of classified images derived from different sensors (Jakubauskas *et al*, 1990). The need to normalise the data for atmospheric and sensor variations is optional because each image is treated as an independent dataset (Singh, 1989; Mas, 1999), unless signature extension between sensing dates is utilised. Due to the independent nature of the classification of each image, care must be taken in the analysis to ensure consistency in the classification process in terms of class allocation, signature extraction and classification quality.

Direct classification of images enables formal identification of information classes compared to enhancement-based approaches that rely upon density slicing of image differences, ratio images and forms of spectral transformations, which must then be related to information classes. Postclassification comparison provides the analyst with a significant degree of flexibility through selective grouping of classification results for presentation of customised change detection classes (Singh, 1989). Generalisation of the data during change detection or identification of particular classes enables specific change detection objectives to be demonstrated, such as providing change data only for

woodland or urban classes. All other change detection approaches require a complete new analysis when interpretation objectives are altered, such as the identification of only a subset of the original change classes.

Most other change detection techniques are limited to identification of change versus no-change, or identification of simple changes such as forest defoliation (Muchoney and Haack, 1994) or transitions between water and eelgrass (Macleod and Congalton, 1998). Postclassification comparison is amenable to identification of transitions between any classes identifiable through the classification process. Mas (1999) used postclassification comparison to successfully monitor changes in an area comprising forest, natural (woodland) vegetation, urban, pasture and water.

The most significant issue related to change detection derived from postclassification comparison is concerned with the estimation of the thematic accuracy of the final product. Past research (Quarmby and Cushnie, 1989; Singh, 1989) suggests that because each image is subject to thematic classification errors, the final change detection product contains much larger errors than either of the component images, and may therefore be less accurate than any other change detection method. In common with other change detection approaches is the issue of spatial errors related to image registration. As a result of the influence of spatial factors on the distribution of thematic errors, the propagation of both of these error components must be considered and are examined in Chapter 7.

6.3 Change Detection Approach

Section 6.2 has considered the fundamental characteristics of change detection techniques that may be used to determine the most appropriate processing approach to apply. While Townshend and Justice (1988) consider the ability to detect changes in land cover types over time with remote sensing depends on the spatial, spectral, radiometric and temporal properties of the sensor system, even more fundamental is the availability of the data itself. The desire to limit analysis to data from a single sensing system may constrain change detection assessment to a particular time frame or compromise the potential quality of results by not considering inclusion of higher quality data. The NALC project has been limited in this manner by a desire to use only Landsat MSS data for the period 1972 to 1992 (Lunetta *et al*, 1998).

Ridd and Liu (1998) used multidate Landsat TM data to determine patterns of land cover change in a near urban area by image differencing, image regression, Kauth-Thomas transformation, and a χ^2 transformation developed by the authors. They found none of the algorithms was clearly superior to the others and concluded that algorithm selection should be soundly based on environmental conditions and application objectives.

Davis and Simonett (1991) consider change detection to be considerably more complicated when more than one sensor system is involved. Failure to understand the impact of sensor system parameters and environmental characteristics on the change detection process can lead to inaccurate results (Jensen *et al*, 1997). For multiple sensor image processing approaches to be successful, they require radiometrically calibrated images and a good comprehension of the spectral features that can be computed from corresponding bands, or from a combination of bands from multiple images (Hill and Aifadopoulou, 1990). Where multidate images are combined as part of change detection, such as image algebra or spectral transformations, inclusion of multiple sensors significantly complicates the process.

Davis *et al* (1991) indicate that comparisons based on more than one sensor convolve surface changes with instrument noise, atmospheric influences, varying IFOVs and spectral response characteristics of the sensors. Change detection via postclassification comparison enables these components to be separated and evaluated individually for each sensing epoch.

Mas (1999) found image enhancement methods of change detection performed poorly due to soil moisture and vegetation phenology variations that the image processing algorithms could not accommodate. Their research determined the postclassification comparison method was less sensitive to these variations and enabled the same classes observed in multidate images with different signatures (for example deciduous trees) to be labelled similarly. Temporal comparisons based on independent classifications also provided details of the nature of changes through the enumeration of the class changed *from* and the class changed *to*.

The requirement to accommodate multisensor and multiscale data into the change detection process clearly establishes that postclassification comparison is the only image processing approach capable of providing coherent change information. Image algebra and spectral transformation approaches rely on scene and sensor normalisation techniques that are unable to calibrate the range of data to be processed in this research in a consistent and accurate manner. Consequently, change assessment will be undertaken using the results of independent supervised and unsupervised classification of each dataset derived in Chapters 4 and 5.

Jakubauskas *et al* (1990) have used postclassification comparison of Landsat MSS and Landsat TM data for change detection. The study stipulated that comparison could only be made at a land cover classification level no finer than could be accurately determined by the lowest resolution sensor and that appropriate generalisation of the finer resolution Landsat TM classified data be performed using a 3 x 3 modal filter. This approach was used in order to achieve consistency of the results, however in the process led to the degradation of the Landsat TM classification.

In the context of the current research, the highest quality of interpretation detail avaiable from each dataset must be maintained. Even when Landsat TM data are compared with Landsat MSS data that contain a lower level of detail, the Landsat TM data will subsequently be compared with SPOT HRV data for change analysis at a finer resolution. In a hierarchical change evaluation process such as developed in this research, it is important to preserve the spatial and thematic quality of the data throughout the change detection process.

The Landsat MSS, Landsat TM and SPOT HRV data have all been classified at Level I and Level II (Chapters 4 and 5), and will be compared pairwise to evaluate land cover changes between 1972 and 1991. Comparisons of classified data at Level I are relatively straight forward as all data have been classified at this level with varying, but acceptable, degrees of accuracy. However by constraining change detection to comparison of Level I classes only, the enhanced information content of the finer resolution Landsat TM and SPOT HRV data is ignored.

Effective evaluation of change in multiscale data is contingent on the use of a hierarchical classification scheme. By using a hierarchical approach, a classification at Level I may be derived through generalisation of the Level II classification matrix into the component Level I classes for a particular sensing epoch. Comparisons between sensing epochs may then be undertaken on the basis of the common Level I classes, and subsequent sensing epochs may be compared using Level I classes, or change the classification level and make all subsequent comparisons on the basis of Level II classes. Application of this approach to change assessment results in classifications for each epoch being treated as discrete entities rather than requiring the integration of all data.

Alternatively, a modified change assessment matrix has been developed for integration of multiscale data classified at different levels within the classification hierarchy. Table 6.1 shows a sample modified change matrix for comparison of Landsat MSS data classified at Level I and Landsat TM data classified at Level II. The major diagonal values (shaded) indicate classes of no-change, and the off-diagonal values indicate classes of change. Within the Level I classes, the change matrix illustrates transitions between Level I classes at the first sensing epoch and the resultant Level I and Level II classes at the second sensing epoch. For example, an area of *Forest* in the Landsat MSS (1972) data may change to *Open* in the Landsat TM data (1986) at Level I. The elements of the change assessment matrix also indicate the distribution of change pixels amongst the Level II classes of *Grassland*, *Recreation* and *Bare ground* that comprise the *Open* class.

The accuracy of change analysis will be investigated using the Kappa Coefficient and Overall Change Accuracy in Section 6.4. Patterns of change within the multiscale satellite data will be considered in Section 6.5 and represented using this new change assessment matrix with the objective of maintaining the integrity and detail contained within the original classification of the respective satellite datasets.

From Landsat MSS 1972	To Landsat TM 1986								
	Urban		Forest			Open			Water
	Res	Com	Pine	Wet	Wood	Gras	Rec	Bare	Water
Urban									
Forest									
Open									
Water									

Table 6.1 *Change assessment matrix format representing land cover changes between Level I (1972) and Levels I and II (1986). No-change areas are shaded*

6.4 Change Detection Accuracy Assessment

Methods of classification accuracy assessment have been discussed in Chapter 5 and the evaluation techniques of Overall Classification Accuracy and Kappa Coefficient have been applied to classifications of multiscale satellite data. Investigation of change detection accuracy introduces an additional dimension because it is the accuracy of the comparison of two classifications that is being evaluated, not the accuracy of the individual classifications. Consequently, change detection accuracy will be assessed by the Overall Change Accuracy metric, which is directly analogous to the Overall Classification Accuracy, as well as the Kappa Coefficient. The multiscale nature of change detection includes classifications at Level I and Level II being directly compared, which must also be included in the accuracy assessment process.

6.4.1 Approaches to Accuracy Assessment

Detailed change detection accuracy assessment has been performed by a number of authors (Fung and Le Drew, 1988; Jensen *et al*, 1993; Macleod and Congalton, 1998;

Martin and Howarth, 1989; Mas, 1999). However most studies incorporate evaluation of the interpretation accuracy of the component multidate images, but do not proceed to a full accuracy assessment of the change image. The emphasis in this research is to evaluate the effectiveness of multiscale change detection, therefore a comprehensive evaluation of the change detection error matrices will be undertaken.

The general approach to change detection accuracy assessment involves pixel by pixel comparison of multidate data using a logical cross-classification approach to identify all combinations of spatially coincident classes between the two image acquisition dates. Where each dataset has been allocated into n classes, there is the potential for n^2 change classes to result. For change detection involving multiscale data where m classes are allocated at date 1 and n classes at date 2, there are $m \times n$ possible change classes. The equivalent cross–classification is undertaken with the reference data to produce a change detection reference dataset. The reference dataset includes the complete image or can be represented by samples as determined by the sampling scheme employed for the accuracy assessment.

Accuracy assessment is performed using standard error matrix approaches with the computation of values for the Overall Change Accuracy and Kappa Coefficient. The magnitude of change detection error matrices presents certain logistical problems in the error assessment process. For Level I comparisons with four land cover classes represented at each date there are 16 possible change classes, while for Level II comparisons with nine land cover classes represented at each date, there are 81 possible change classes. A combined Level I and Level II change detection image produces 36 possible change classes. These comparisons produce error matrices of dimension 16 x 16, 36 x 36 and 81 x 81 for Level I, combined Level I and Level II, and Level II change detection assessments, respectively. Consequently, accuracy assessments will be undertaken for all Level I and combined Level I and Level II change detection matrices, but for Level II a complete error matrix will be computed only for comparison of the 1986 and 1991 data. For the Level II change detection error assessment of the 1972 and 1986 data, only the Overall Change Accuracy will be computed.

Generalisation of the change detection error matrix will also be used to analyse error patterns in the complete change detection error matrices (Macleod and Congalton, 1998). A simplified error matrix representing change/no-change classes will be derived by aggregating the change detection results into a single class representing land covers that do not exhibit change between sensing dates, and another class representing land covers that do change. Measures of Overall Change Accuracy and the Kappa Coefficient will also be computed for this generalised error matrix. Where the accuracy in the change/no-change error matrix increases significantly over the full change detection error matrix, it indicates that general change is being detected, but not the exact classes of change. If the two matrices produce similar results, then not even the general changes are being detected by the change assessment approach (Macleod and Congalton, 1998).

As a result of the logical cross-classification to determine temporal changes in the data, there are many change classes that comprise small numbers of pixels. Stratified random sampling procedures discussed in Section 5.3, and applied to the classification accuracy assessment of individual images in Section 5.4, will not be used to evaluate the change detection images due to the very large number of pixels required to adequately sample the large number of small change classes. Complete maps of reference data are available for all sensing epochs, therefore a total enumeration approach to accuracy assessment will be utilised and all pixels within the change images will be evaluated against the equivalent reference data.

6.4.2 Change Statistics

Change statistics for these interpretations are based upon the premise that the reference data interpreted from the aerial photography are error free. The primary satellite data and aerial photography have all been obtained from image archives, and there were no coincident ground data or independent information available to verify the thematic accuracy of the reference data interpreted from the aerial photography. Errors in the detection of change by the reference data may occur, however have been minimised through careful interpretation of the medium scale aerial photography. Discrepancies are most likely caused by changes in land cover between the acquisition of the satellite data and exposure of the aerial photography used for compilation of the reference data.

It is not possible to quantify these errors in this research, and it is assumed the reference data contain minimal thematic error.

Diagonal elements of the change error matrix represent pixels in the classified satellite data that correspond with the same change class present in the reference data, and are used to compute the Overall Change Accuracy value. Off-diagonal values represent errors of commission and omission, and are included in the computation of the Kappa Coefficient. Consequently, the Kappa Coefficient provides a more comprehensive and reliable assessment of change accuracy.

Interpretation of errors of commission and omission may be made on the basis of spectral and spatial properties. Spectral properties are important in the allocation of classes during the classification phase and have been analysed in detail in Chapters 4 and 5 with the objective of defining the optimum classification approach to land cover mapping using multiscale satellite data. In the change detection phase the class labels allocated as a result of classification are considered fixed, and only the spatial distribution of each land cover class is significant because the correspondence of change labels between the classified data and the reference data is being evaluated.

Table 6.2 details the results of change detection accuracy assessment between 1972 and 1986 using Landsat MSS and Landsat TM data, respectively. At Level I the Overall Change Accuracy is 64.0 percent and the Kappa Coefficient is 55.0 percent (Table 6.2a). According to Landis and Koch (1977) this represents *moderate* agreement based on the value of the Kappa Coefficient. This compares with Overall Classification Accuracy and Kappa Coefficient values of 79.1 and 57.3 percent for the classification of Landsat MSS data, and 77.3 and 62.5 percent for the classification of Landsat TM data.

The distribution of change is concentrated in the *Forest-Urban* and *Forest-Open* classes, as indicated by the aerial photographs for each date in Figures 2.6 and 2.7. The *Forest-Urban* change class reflects the increasing urban development of the area from rural land uses, and the *Forest-Open* class represents a transitional stage of cleared land between rural and urban land use. Substantial areas of *Urban* and *Forest*, and small areas of *Open* and *Water* remain unchanged.

Classified data	Reference Data																
Change class	Urban-Urban	Forest-Forest	Open-Open	Water-Water	Urban-Forest	Urban-Open	Urban-Water	Forest-Urban	Forest-Open	Forest-Water	Open-Urban	Open-Forest	Open-Water	Water-Urban	Water-Forest	Water-Open	
Urban-Urban	7732	201	305	0	3	11	0	804	147	0	595	15	0	0	0	0	9813
Forest-Forest	121	4482	454	60	3	0	1	429	1330	44	31	200	0	1	4	0	7140
Open-Open	16	25	183	0	0	0	0	8	24	0	37	5	0	0	0	0	278
Water-Water	0	0	0	27	0	0	0	0	0	0	0	0	0	0	0	0	27
Urban-Forest	612	176	89	3	2	0	1	57	46	0	22	21	0	0	0	0	1029
Urban-Open	267	125	682	0	5	0	0	70	215	0	84	25	0	0	0	0	1473
Urban-Water	0	5	0	2	0	0	1	0	0	0	0	0	0	0	0	0	8
Forest-Urban	1079	628	651	0	0	14	0	6697	1591	3	853	145	0	0	0	0	11661
Forest-Open	54	781	1507	2	0	0	2	824	10397	10	164	87	0	0	0	0	13828
Forest-Water	5	40	0	83	0	0	0	0	0	2	0	0	0	0	1	0	131
Open-Urban	201	32	25	0	1	0	0	135	4	0	160	2	0	0	0	0	560
Open-Forest	4	5	4	0	0	0	0	5	0	0	5	0	0	0	0	0	23
Open-Water	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Water-Urban	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Water-Forest	0	5	0	14	0	0	0	0	0	0	0	0	0	0	0	0	19
Water-Open	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	10091	6485	3880	191	14	25	5	9029	13754	59	1951	500	0	1	5	0	45990
																Overall Change Accuracy	64.0
																Kappa Coefficient (percent)	55.0

(a) Change matrix including all classes

Classified data	Reference data		
	No change	Change	
No change	13 566	3 692	17 258
Change	7 081	21 651	28 732
	20647	25 343	45 990
Overall Change Accuracy (percent)			76.6
Kappa Coefficient (percent)			51.9

(b) Change matrix including change and no-change classes

Table 6.2 *Error matrix for Level I classification between Landsat MSS (1972) and Landsat TM (1986) data. No-change areas are shaded*

Errors of omission and commission are concentrated in change classes that include the *Urban, Forest* or *Open* classes. These errors are the result of incorrect identification of land covers at the classification stage at either date. Inspection of Tables 5.3 and 5.11 shows that *Open* and *Forest* achieved only moderate classification accuracies in the classification of the Landsat MSS and Landsat TM data, respectively. Errors in change class allocation may also be caused as a result of imperfect geometric registration of the images and subsequent errors caused by boundary displacement (Wickware and

Howarth, 1981), even though the original classification for these pixels may have been correct. The effect of boundary displacement will be discussed in Chapter 7.

The *Urban-Urban* class includes substantial commission errors from the *Forest-Forest, Open-Open, Forest-Urban, Forest-Open* and *Open-Urban* classes. Errors of commission indicate classification errors in one or both satellite datasets, which place the pixels in the incorrect change class. Additionally, the errors may be due to discrepancies in the reference data used for change detection comparisons. Discrepancies are caused by changes in land cover between the acquisition of the satellite data and exposure of the aerial photography used for compilation of the reference data, or by errors in the actual interpretation of the reference data.

The explanation of commission errors where the identified change sequence is highly unlikely to exist is much clearer, such as with *Urban-Open.* The presence of *Urban-Open* reference data for the study area indicates misregistration or interpretation errors in one date of the reference data. The fact that no change data are recorded for this transition sequence indicates confidence in the classification of the remotely sensed data.

For higher spatial resolution data, some changes such as *Urban-Forest,* may be a result of the enhanced ability of the sensor to identify small areas of trees that blend with the more expansive urban structures when recorded in data with lower spatial resolution, and are therefore not detected. Additionally, where these unlikely transitions occur along class boundaries they may be the result of edge effects caused by mixed pixels influencing the classification process.

Commission errors within the no-change classes (*viz Urban-Urban, Forest-Forest, Open-Open*) are relatively low, but are comparatively larger with the change classes of *Forest-Urban, Forest-Open, Open-Urban* and *Open-Forest.* For other change classes such as *Urban-Forest, Urban-Open, Forest-Urban* and *Forest-Open* a large proportion of the commission errors fall within the no-change classes of the reference data. The cause of these errors is due to a mixture of classification and reference data errors discussed above, but more importantly, these errors have a detrimental effect on the accuracy of the generalised change/no-change classes considered below.

Omission errors follow a similar pattern to the commission errors. Relatively few omission errors fall within the no-change classes, but substantial omission errors for the no-change classes are distributed in the change classes, in particular *Urban-Forest, Urban-Open, Forest-Urban* and *Forest-Open.* Omission errors result from change pixels identified in the reference data being allocated to a different class by the classified change data. The primary source of these errors is due to classification errors in each satellite dataset and to a lesser extent the interpretation of the reference data or registration errors. Many of the omission errors recorded in Table 6.2a fall in the *Urban-Forest* and *Urban-Open* change classes, which are not logical change sequences, and clearly indicate errors in either or both classifications.

At the generalised level of change/no-change the Overall Change Accuracy increases to 76.7 percent, but the Kappa Coefficient decreases to 51.9 percent (Table 6.2b). An increased Overall Change Accuracy value is a logical consequence of generalising the change classes, however the decreased Kappa Coefficient is due to a proportionally larger increase in errors compared to the increase in correctly identified change pixels discussed above. This same pattern is reflected in all change comparisons undertaken in this research and is contributed by classification errors in each dataset, reference data interpretation errors, and registration errors within the satellite and reference datasets.

Results of change detection accuracy assessment at Level I between the Landsat TM and SPOT HRV data are indicated in Table 6.3. The Overall Change Accuracy has increased to 69.6 percent and the Kappa Coefficient has decreased marginally to 54.0 percent (*moderate* agreement), compared to the results in Table 6.2 for the Landsat MSS and Landsat TM data. These changes must be considered in the context of the improved spatial resolution of the data, but also considering the significantly more complex spatial arrangement of land covers present in the 1986 and 1991 epochs, compared to the 1972 satellite data.

The change pattern has shifted substantially compared to the 1972 to 1986 time period. In the earlier period *Forest-Urban* and *Forest-Open* formed the major change classes, however in the 1986 to 1991 period the change is concentrated in the *Open-Urban* transition, which reflects the pattern of development in the region.

Classified data	Reference data																
Change Class	Urban-Urban	Forest-Forest	Open-Open	Water-Water	Urban-Forest	Urban-Open	Urban-Water	Forest-Urban	Forest-Open	Forest-Water	Open-Urban	Open-Forest	Open-Water	Water-Urban	Water-Forest	Water-Open	
Urban-Urban	18140	506	1227	1	20	54	0	333	55	0	1401	52	19	0	0	0	21807
Forest-Forest	33	2641	191	3	3	0	0	3	22	0	0	39	0	0	0	0	2875
Open-Open	39	46	1358	1	0	8	0	1	15	0	5	17	0	0	0	0	1490
Water-Water	2	23	0	115	0	0	0	0	0	22	0	0	0	0	0	0	162
Urban-Forest	33	117	12	0	0	0	0	0	0	0	2	9	0	0	0	0	173
Urban-Open	5	11	28	0	0	0	0	0	0	0	0	0	0	0	0	0	44
Urban-Water	2	7	0	2	0	0	0	0	0	0	0	0	0	0	0	0	11
Forest-Urban	1214	896	985	5	20	7	0	886	142	0	555	92	0	0	0	1	4803
Forest-Open	2	30	101	0	0	0	0	0	11	0	0	8	0	0	0	0	152
Forest-Water	8	208	6	113	0	0	0	0	6	30	0	0	6	0	0	1	378
Open-Urban	1443	428	3473	8	11	19	0	337	102	0	7923	157	18	0	0	0	13919
Open-Forest	3	111	26	0	0	0	0	0	3	0	0	3	0	0	0	0	146
Open-Water	0	1	3	5	0	0	0	0	4	2	5	0	5	0	0	0	25
Water-Urban	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
Water-Forest	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Water-Open	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	20927	5027	7348	253	54	88	0	1560	360	54	9891	377	48	0	0	2	45989
																Overall Change Accuracy (percent)	69.6
																Kappa Coefficient (percent)	54.0

(a) Change matrix including all classes

Classified data	Reference data		
	No change	Change	
No change	24 265	2 069	26 334
Change	9 290	10 365	19 655
	33 555	12 434	45 989
Overall Change Accuracy (percent)			75.3
Kappa Coefficient (percent)			47.1

(b) Change matrix including change and no-change classes

Table 6.3 *Error matrix for Level I change detection between Landsat TM (1986) and SPOT HRV (1991) data. No-change areas are shaded*

The overall pattern of commission and omission error remains similar to that in Table 6.2 with most error occurring amongst the change classes compared to the no-change classes. This again results in an increased Overall Change Accuracy and decreased Kappa Coefficient when the data are aggregated into the generalised change/no-change classes (Table 6.3b).

Detailed evaluation of change assessment at Level II between the Landsat MSS (1972) and Landsat TM (1986) data produced an Overall Change Accuracy of 33.0 percent and due to this relatively low value a corresponding Kappa Coefficient was not computed. The Overall Change Accuracy was lower than expected, especially considering that the individual classification accuracies for the Landsat MSS and Landsat TM data were 61.6 and 68.7 percent, respectively. At an Overall Change Accuracy of 33 percent these data cannot make a useful contribution to change assessment.

The major contributing factor to the low accuracy relates to the large difference in spatial resolution between the datasets. The magnitude of change detection errors due to boundary definition uncertainty when comparing 79 x 56 m resolution Landsat MSS data and 30 m resolution Landsat TM data is substantial. Compounding this effect is the generalised nature of the land cover patterns present in the 1972 data compared to the substantially more complex patterns in the 1986 data. Both factors contribute to discrepancies in the change data and result in the very low Overall Change Accuracy.

Table 6.4a provides a detailed evaluation of change assessment results for the Landsat TM and SPOT HRV change detection data at Level II. An Overall Change Accuracy of 51.4 percent and Kappa Coefficient of 43.9 were achieved indicating a *moderate* level of agreement. Inspection of Table 6.4a confirms the cause of the lower Kappa Coefficient value due to considerable errors of commission and omission for most change classes. This compares to the Level I change assessments where the errors were restricted to a few classes. As was the case for the change assessments at Level I, between-class errors amongst the no-change classes are relatively few in number, however errors between change and no-change classes are considerable. The increasing spatial complexity of the classes at the 1986 and 1991 sensing epochs have considerable influence on this outcome, as well as the increase in within-class variance as pixel dimensions decrease.

Table 6.4b summarises generalised change at Level II into change/no-change classes. The Overall Change Accuracy improves to 67.2 percent, but the Kappa Coefficient decreases to 34.2 percent. This result reinforces the conclusion that high levels of commission and omission are present in the change assessment, but are not measured by the Overall Change Accuracy parameter.

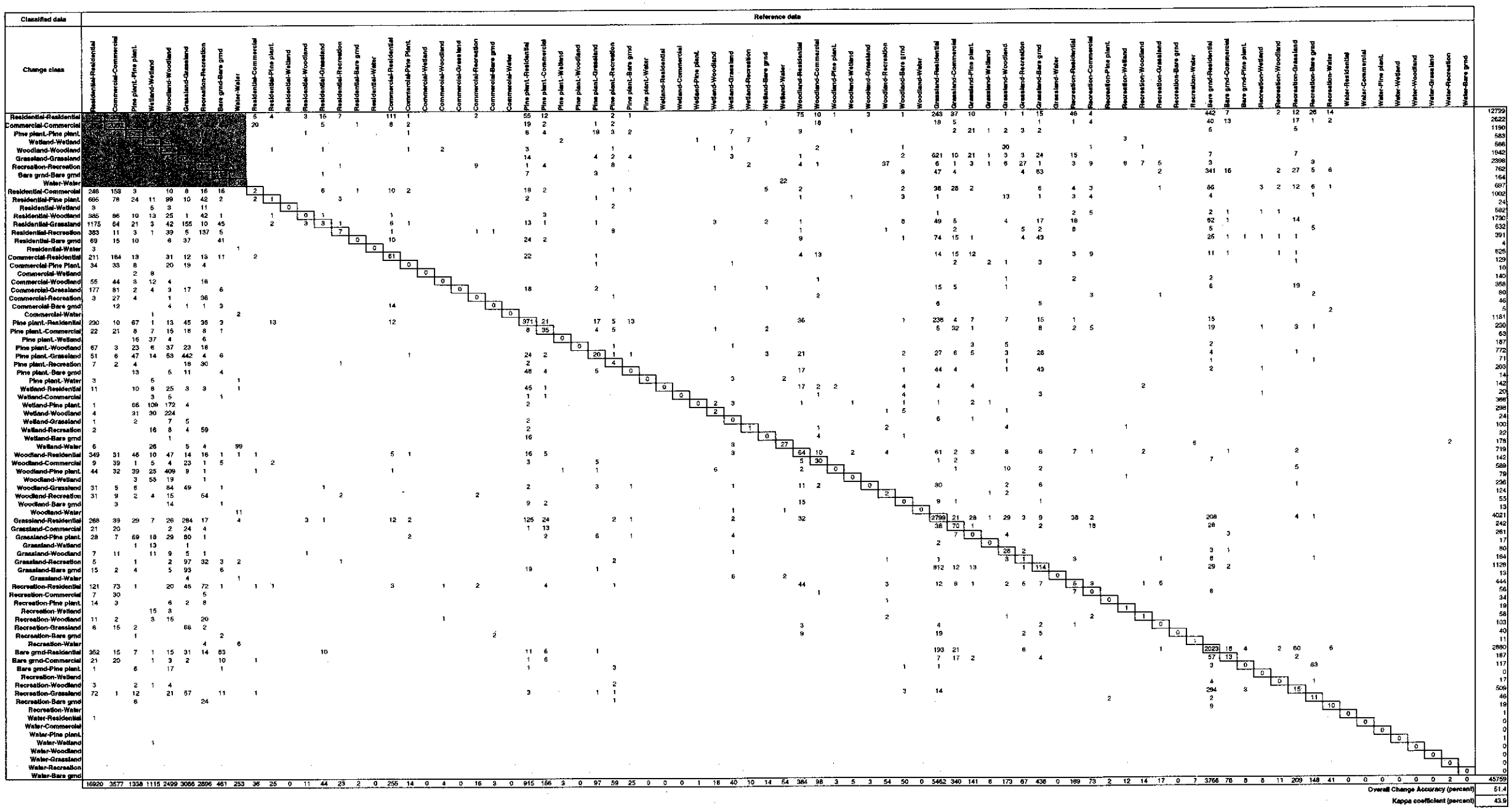


Table 6.4 *(a) Change matrix including all classes (no-change areas are shaded)*

Classified data	Reference data		
	No change	Change	
No change	20 075	2 953	23 028
Change	12 070	10 661	22 731
	32 145	13 614	45 759
Overall Change Accuracy (percent)			67.2
Kappa Coefficient (percent)			34.2

(b) Change matrix including change and no-change classes

Table 6.4 *Error matrix for Level II change detection between Landsat TM (1986) and SPOT HRV (1991) data. No-change areas are shaded*

Comparative change statistics at Level I, Level II and the generalised change/no-change classes are summarised for both change sequences in Tables 6.5 and 6.6. Accuracy results at Level I are consistently higher than Level II due to the lower levels of spatial and thematic detail derived at Level I. The Overall Change Accuracy and Kappa Coefficient vary consistently for these change assessments at both levels of generalisation, which indicates that the specific patterns of change identified at Level II are consistent with the general patterns identified at Level I.

Change Period	Level I		Level II	
	Over. Change Accuracy	Kappa Coefficient	Over. Change Accuracy	Kappa Coefficient
1972 - 1986	64.0	55.0	33.0	NA
1986 - 1991	69.6	54.0	51.4	43.9

Table 6.5 *Change statistics (percent) at Level I and Level II using four and nine land cover classes, respectively (Kappa Coefficient was not computed at Level II for the 1972 to 1986 data due to low Overall Change Accuracy results)*

Aggregation into change/no-change classes has produced conflicting results for the Overall Change Accuracy and Kappa Coefficient. In all cases the Overall Change Accuracy has increased, however the Kappa Coefficient has decreased compared to the full Level I and Level II error matrices. These results are consistent with a study by Macleod and Congalton (1998), but indicate the contrasting effect of the two types of

generalisation. Generalisation of classes in a hierarchical manner from Level II to Level I results in an increase in the Overall Change Accuracy and Kappa Coefficient (Table 6.5). However, when a more comprehensive generalisation from Level I or Level II change assessments to the equivalent change/no-change classes is made without consideration of the structural relationships between change classes, the Overall Change Accuracy increases, however the Kappa Coefficient decreases (Table 6.6). Results from the hierarchical generalisation of the data reinforce the need to structure classification schemes according to the established relationships between the targets. Arbitrary generalisation may not provide the desired quality of results compared to robust evaluation techniques such as the Kappa Coefficient.

Change Period	Level I		Level II	
	Over. Change Accuracy	Kappa Coefficient	Over. Change Accuracy	Kappa Coefficient
1972 – 1986	76.6	51.9	NA	NA
1986 – 1991	75.3	47.1	67.2	34.2

Table 6.6 *Change statistics (percent) at Level I and Level II using change and no-change classes (change statistics were not computed for 1972 to 1986 Level II data due to low Overall Change Accuracy)*

Macleod and Congalton (1998) suggest the comparison of the responses between the full change assessment and the change/no-change assessment provides a measure of the effectiveness of change class identification. Where the generalised comparisons produce a significantly higher classification accuracy they indicate that general change is being detected but not the exact change, and if the two matrices produce similar results, then not even general changes are being detected. Results in this research, and those by Macleod and Congalton (1998) and Mas (1999), indicate this conclusion is only relevant when using the Overall Change Accuracy for comparison because the Kappa Coefficient produces inconclusive results. Substantially increased values of Overall Change Accuracy following generalisation to change/no-change classes confirms that general change in the study area is being identified. Moderate values of the Kappa Coefficient in the full Level I and Level II classifications (Table 6.5) shows there is still scope for improvements in change detection accuracy and reduction in errors of omission and commission.

Classification results and change statistics at Level I and Level II (Tables 5.15, 5.16 and 6.5) indicate that useful information is available from all datasets analysed at both levels within this research. Integration of multiscale data for change assessment requires a combined analysis of the data in order to utilise effectively the information contained within each dataset. Table 6.1 shows a modified change matrix designed to permit comparison of multidate classifications. The table compares each Level I class at the first epoch with the corresponding Level I and Level II class (or classes) of the second epoch, to determine the distribution of pixels amongst the change classes.

Table 6.7 shows details of the resulting error matrix derived for the assessment of change detection accuracy of Landsat MSS data at Level I and Landsat TM data at Level II. The Overall Change Accuracy is 56.3 percent and the Kappa Coefficient is 51.1 percent. This represents a substantial improvement over the equivalent Level II Overall Change Accuracy for these data (33.0 percent), and is only marginally lower than the value of the Kappa Coefficient derived from the Level I change assessment for the same period (55.0 percent).

Unchanged classes are shaded (Table 6.7) and, because the comparisons are made across levels of the classification hierarchy, the first mentioned class is from Level I and the second class is from Level II. For example, *Urban* at Level I may be classified as *Residential* or *Commercial* in Level II, therefore the equivalent unchanged classes are *Urban-Residential* and *Urban-Commercial.*

The individual classifications are derived from the unsupervised Level I classification of the Landsat MSS data (Table 5.3) and the supervised Level II classification of the Landsat TM data (Table 5.12). Table 6.7 provides an expanded form of Table 6.2a, where aggregation of the classes in Table 6.7 yields the equivalent classes and pixel counts included in Table 6.2a. This occurs because the same signatures were used for supervised classifications at both Level I and Level II, and the classification results were aggregated postclassification in order to produce data at the required classification level. The same approach could not be applied to the unsupervised classification because the allocation of clusters to classes was undertaken independently at each classification level.

Classified data	Reference data																																			
Change Class	Urban-Residential	Urban-Commercial	Forest-Pine plant.	Forest-Wetland	Forest-Woodland	Open-Grassland	Open-Recreation	Open-Bare ground	Water-Water	Urban-Pine plant.	Urban-Wetland	Urban-Woodland	Urban-Grassland	Urban-Recreation	Urban-Bare ground	Urban-Water	Forest-Residential	Forest-Commercial	Forest-Grassland	Forest-Recreation	Forest-Bare ground	Forest-Water	Open-Residential	Open-Commercial	Open-Pine plant.	Open-Wetland	Open-Woodland	Open-Water	Water-Residential	Water-Commercial	Water-Pine plant.	Water-Wetland	Water-Woodland	Water-Grassland	Water-Recreation	Water-Bare ground	
Urban-Residential	6330	213	37	2	119	[illegible]	[illegible]	[illegible]	[illegible]	0	1	2	1	3	0	0	625	40	27	71	31	0	432	42	4	2	5	0	0	0	0	0	0	0	0	0	8045
Urban-Commercial	155	1224	12	0	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0	1	0	6	0	40	99	0	3	15	0	22	99	1	2	1	0	0	0	0	0	0	0	0	0	1768
Forest-Pine plant.	[illegible]	2	1251	184	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0	0	0	0	0	266	25	832	95	19	0	10	13	110	1	35	0	0	0	0	0	0	0	0	0	3569
Forest-Wetland	[illegible]	0	366	210	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	1	0	0	0	1	4	0	40	70	0	43	0	0	1	2	12	0	0	0	0	4	0	0	0	0	1681
Forest-Woodland	[illegible]	3	152	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	2	0	0	0	0	42	92	211	62	1	1	6	2	10	0	29	0	0	0	0	0	0	0	0	0	1869
Open-Grassland	[illegible]	[illegible]	0	0	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0	0	0	0	0	3	0	1	0	0	0	2	7	0	0	0	0	0	0	0	0	0	0	0	0	72
Open-Recreation	[illegible]	0	0	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0	0	0	0	0	0	0	0	3	0	0	2	21	0	0	0	0	0	0	0	0	0	0	0	0	118
Open-Bare ground	[illegible]	0	0	0	2	97	[illegible]	[illegible]	[illegible]	0	0	0	0	0	0	0	3	2	0	0	20	0	4	1	0	0	5	0	0	0	0	0	0	0	0	0	88
Water-Water	0	0	0	0	[illegible]	0	[illegible]	0	47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	27
Urban-Pine plant.	155	11	16	10	26	42	8	14	0	0	0	0	0	0	0	0	18	12	20	3	1	0	2	0	0	1	4	0	0	0	0	0	0	0	0	0	343
Urban-Wetland	16	0	0	6	12	0	0	2	3	0	0	0	0	0	0	1	0	0	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	46
Urban-Woodland	411	19	16	9	79	3	15	5	0	0	0	2	0	0	0	0	7	20	0	18	0	0	8	12	0	0	16	0	0	0	0	0	0	0	0	0	640
Urban-Grassland	52	11	23	2	14	142	2	19	0	0	0	2	0	0	0	0	22	0	74	0	5	0	20	5	0	0	15	0	0	0	0	0	0	0	0	0	408
Urban-Recreation	123	44	0	0	61	7	372	6	0	0	0	3	0	0	0	0	27	8	0	65	3	0	26	16	0	0	2	0	0	0	0	0	0	0	0	0	763
Urban-Bare ground	23	14	8	0	17	38	21	75	0	0	0	0	0	0	0	0	10	3	36	0	32	0	16	1	1	0	7	0	0	0	0	0	0	0	0	0	302
Urban-Water	0	0	0	5	0	0	0	0	2	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8
Forest-Residential	494	58	185	53	234	237	23	232	0	0	0	0	6	3	2	0	5580	273	532	326	542	1	618	83	49	0	41	0	0	0	0	0	0	0	0	0	9572
Forest-Commercial	155	372	50	35	71	87	17	55	0	0	0	0	1	0	2	0	358	486	76	91	24	2	40	112	40	0	15	0	0	0	0	0	0	0	0	0	2089
Forest-Grassland	9	9	307	66	77	512	11	33	0	0	0	0	0	0	0	0	252	40	5588	114	265	7	41	19	24	0	24	0	0	0	0	0	0	0	0	0	7398
Forest-Recreation	21	9	21	35	153	34	304	6	2	0	0	0	0	0	0	2	50	100	254	1244	9	3	6	20	14	0	5	0	0	0	0	0	0	0	0	0	2282
Forest-Bare ground	6	0	74	3	45	40	0	567	0	0	0	0	0	0	0	0	365	17	345	27	2551	0	73	5	6	0	14	0	0	0	0	0	0	0	0	0	4138
Forest-Water	5	0	0	40	0	0	0	0	83	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	1	0	0	0	0	131
Open-Residential	66	36	0	0	25	14	3	2	0	0	0	1	0	0	0	0	79	6	0	0	2	0	133	18	0	0	2	0	0	0	0	0	0	0	0	0	387
Open-Commercial	14	85	0	0	7	0	0	6	0	0	0	0	0	0	0	0	7	43	0	0	2	0	4	5	0	0	0	0	0	0	0	0	0	0	0	0	173
Open-Pine plant.	1	2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	12
Open-Wetland	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Open-Woodland	1	0	0	0	4	0	4	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11
Open-Water	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Water-Residential	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Water-Commercial	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Water-Pine plant.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Water-Wetland	0	0	0	5	0	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	19
Water-Woodland	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Water-Grassland	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Water-Recreation	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Water-Bare ground	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	7955	2136	2338	1242	2905	1677	1004	1199	191	0	1	13	9	6	10	5	7758	1271	8038	2184	3522	59	1465	486	260	8	232	0	0	0	0	5	0	0	0	0	45989
Overall Change Accuracy (percent)																																					56.3
Kappa coefficient (percent)																																					51.1

Table 6.7 *Error matrix for evaluation of change between a Level I classification of Landsat MSS data and a Level II classification of Landsat TM data.*

No-change areas are shaded

The patterns of change error are similar in Tables 6.2a and 6.7, however the linking of the Level I classification from 1972 and the Level II classification from 1986 enables the continuity of change assessment to be maintained across classification levels. Assessment of change within the study area may now be undertaken using data at all spatial scales at comparable levels of detail using the combined Level I and Level II approach described above for the 1972 to 1986 period, followed by the comparison of Level II classifications for the 1986 to 1991 period. This approach minimises the effects of discontinuities in data acquisition systems and provides for the combination of multiscale data for temporal assessment.

The change assessments shown in Table 6.7 are necessarily undertaken at different levels of detail and the quality of information extracted from the lower resolution Landsat MSS data has not increased as a result. The Kappa Coefficient provides a valid assessment of the quality of the changes observed between multiscale data by making a direct comparison with the reference data derived from the aerial photography.

6.5 Change Reporting

A variety of approaches is available for reporting and assessment of land cover change derived from remotely sensed data. These approaches can be categorised either as area-based or pixel-based methods, and utilise image or statistical techniques of representation. Area-based methods rely upon extraction of area statistics for land covers derived at each sensing epoch and comparisons are made regarding the change in the area of each land cover. Pixel-based methods provide data derived from a pixel by pixel comparison of the multidate images and summarisation of the observed change. Change data from both of these approaches are discussed in the following section and utilise the representation methods of change summaries, change maps and change matrices for change reporting.

6.5.1 Change Summaries

Change summaries measure the variation in area occupied by each land cover class between sensing epochs and provide class by class reports of land cover change by area and/or percentage (Wickware and Howarth, 1981). Table 6.8 provides details of the magnitude of land cover change between 1972 and 1991 using the highest quality data available for each period. The period 1972 to 1986 is assessed at Level I while the

Land cover class	Area 1972	Change 1972-86		Area 1986	Change 1986-91	
	Ha	*Ha*	*Percent*	*Ha*	*Ha*	*Percent*
Urban	497.4	384.2	77.2			
Residential				720.4	197.1	27.4
Commercial				161.2	5.9	3.7
Forest	1316.4	-988.0	-75.0			
Pine Plantation				157.0	-12.3	-7.8
Wetland				69.8	-38.0	-54.4
Woodland				101.6	-24.2	-23.8
Open	34.6	588.7	1701.4			
Grassland				315.1	-86.5	-27.4
Recreation				126.9	18.3	14.4
Bare ground				181.2	-75.2	-41.5
Water	1.8	4.8	266.7	6.6	10.2	153.6

Table 6.8 *Change assessment statistics at Level I and Level II representing area-based changes between 1972 and 1991*

Land Cover Class	Area 1972	Gain		Loss		Unchanged	
	Ha	*Ha*	*Percent*	*Ha*	*Percent*	*Ha*	*Percent*
Urban	497.4	488.8	99.3	100.4	20.2	392.5	78.9
Forest	1 316.4	42.9	3.2	1 024.7	78.8	285.6	21.7
Open	34.6	612.0	17 688.0	23.3	67.3	11.1	32.1
Water	1.8	5.5	305.6	0.8	44.4	1.1	61.1
Total	**1 850.2**	**1 149.2**	**62.1**	**1 149.2**	**62.1**	**690.3**	**37.3**

Table 6.9 *Change assessment statistics at Level I representing pixel by pixel land cover changes between 1972 and 1986*

Land Cover Class	Area 1986	Gain		Loss		Unchanged	
	Ha	*Ha*	*Percent*	*Ha*	*Percent*	*Ha*	*Percent*
Urban	881.6	749.0	85.0	9.1	1.0	872.3	98.9
Forest	328.4	12.7	3.9	213.3	65.0	115.0	35.0
Open	623.3	7.9	1.3	563.6	90.4	59.6	9.6
Water	6.6	16.5	250.0	0.1	1.5	6.5	98.5
Total	**1 839.9**	**786.1**	**42.7**	**786.1**	**42.7**	**1 053.4**	**57.2**

Table 6.10 *Change assessment statistics at Level I representing pixel by pixel land cover changes between 1986 and 1991*

Land Cover Class	Area 1986	Gain		Loss		Unchanged	
	Ha	*Ha*	*Percent*	*Ha*	*Percent*	*Ha*	*Percent*
Residential	720.4	404.2	56.1	202.6	28.1	511.9	71
Commercial	161.2	62.1	38.5	55.6	34.5	104.9	65.1
Pine plantation	157.0	96.3	61.3	108.8	69.3	47.6	30.3
Wetland	69.8	8.6	12.3	46	65.9	23.3	33.4
Woodland	101.6	54.5	53.6	78.3	77.1	22.7	22.3
Grassland	315.1	149.3	47.4	237.2	75.3	77.7	24.7
Recreation	126.9	48.7	38.4	30.6	24.1	95.9	75.6
Bare ground	181.2	75.3	41.6	150.2	82.9	30.5	16.8
Water	6.6	10.3	156.1	0	0	6.6	100
Total	**1 839.8**	**909.3**	**49.4**	**909.2**	**49.4**	**921.1**	**50.1**

Table 6.11 *Change assessment statistics at Level II representing pixel by pixel land cover changes between 1986 and 1991*

period 1986 to 1991 is assessed at Level II. In each case the change in area and change as a percentage of the class present at the commencement of the change period are listed. These data are computed from the classified images at each date and show only the net change in area. In an extreme case no net change may occur, however it is possible between dates for an amount equal to the area of a complete cover class to disappear and an equivalent area to appear at the next epoch. No net change is recorded, yet the location of the whole class has altered between imaging dates.

Where classes experience an increase in area at one location but a decrease in area at another location, these variations are not reflected in these simple figures. For example, for the period 1986 to 91 the area of *Grassland* shows a net decrease of 86.5 ha (Table 6.8), yet from Table 6.10 it increases by 149.3 ha and decreases by 237.2 ha. Minor differences between areas result because change area summaries are computed using two methods. The values in Table 6.8 are computed directly from the differences between areas identified in each image, and the *Gain/Loss* areas are derived from the crosstabulation matrix between dates.

An alterative method of change summarisation by area is shown in Tables 6.9, 6.10 and 6.11. These data are calculated from the change images derived through a pixel by pixel comparison of each classified image (Green *et al*, 1994). Consequently, the data represent the summation of changes identified in individual pixels for each land cover class and also the direction of change, which is recorded as a gain or loss. Table 6.9 includes data interpreted at Level I between 1972 and 1986, and Table 6.10 includes data interpreted for the 1986 to 91 period. Minor discrepancies between area estimates are due to differences in classification techniques used at Level I and Level II.

This approach elaborates on the direction of change in terms of gains and losses, but is limited in the detail provided. While the technique adequately describes the changes in the area of each of the land covers, it does not provide any indication of the specific classes that are affected by the changes, nor does it provide information regarding the spatial location of changes in area. Knowledge of the transition sequence (e.g. from *Pine plantation* to *Open*) is useful for interpretation of the significance of observed changes. Inclusion of the spatial location and distribution of land cover change is also

an important assessment indicator. The knowledge of whether change occurs as many isolated pockets or over a limited number of large areas may be significant.

6.5.2 Change Maps

Change maps are used extensively for representation of change (Laba *et al*, 1997; Riley *et al*, 1997) and provide a convenient summary of the overall magnitude and distribution of change within an area. Figures 6.1 and 6.2 show change maps at Level I covering the two change periods included in this research. Change maps rely on the enunciation of all change classes and representation through legends and variable shading or patterning, which may lead to complex graphical representations. Both maps are dominated by the unchanged *Urban* class, with Figure 6.1 containing substantial unchanged *Forest* areas and large areas of the change classes of *Forest-Urban* and *Forest-Open*. The transition class of *Open-Urban* dominates change in Figure 6.2. Both change maps also include small areas comprising a large number of classes that are difficult to identify. The presence of these classes may be significant (such as *Wetland* or *Water*), but in a graphical representation their existence and change status is difficult to determine. A variation on this approach provides for display of a greyscale scene as a backdrop and superimposition of selected change classes (Jensen *et al*, 1993), however it represents a loss of overall information content as many change classes are excluded from the display.

In the Level I change maps shown, only four primary land cover classes were identified, with a total of sixteen change/no-change classes being represented overall. Interpretation at Level II provides 81 classes overall and, if only those classes where change is actually recorded are represented, 74 categories would be required. Change representation through change maps produces complex graphical representations with many classes. Even at Level II, a map with 81 categories provides a very detailed information product. Establishment of a priority of classes may be possible, however without representation of the total dataset a certain amount of information loss occurs.

Change maps provide useful visual assessments of change distribution and magnitude, however the overall complexity of the representation and lack of specific area estimates are drawbacks of this technique.

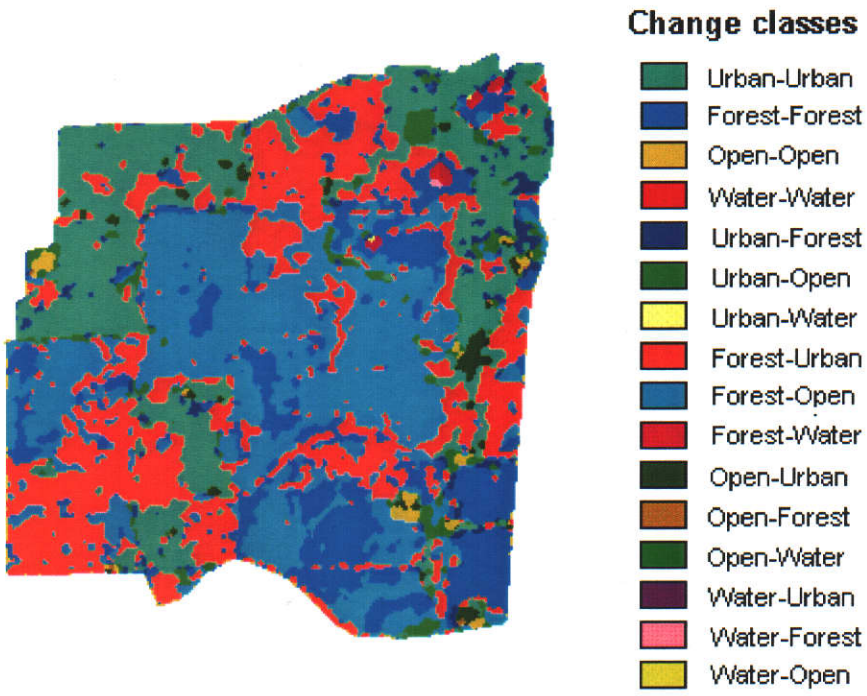


Figure 6.1 *Level I change map for the period 1972 to 1986 (Scale 1:50 000)*

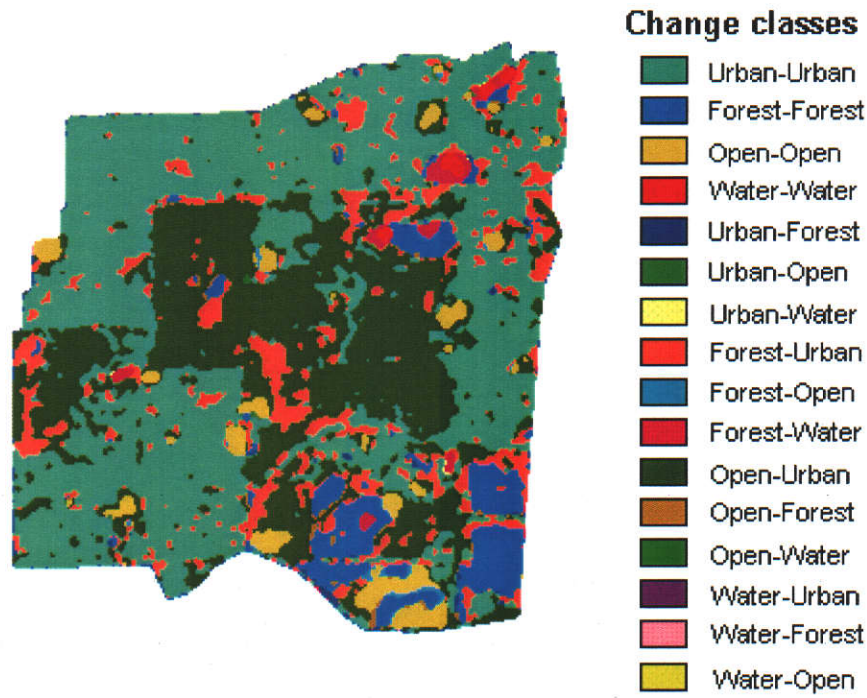


Figure 6.2 *Level I change map for the period 1986 to 1991 (Scale 1:50 000)*

6.5.3 Change Matrix Formulation

Communication of land cover change information is complicated by the large amount of detail available from the image analysis process. Where n classes are interpreted at each sensing epoch, n^2 change classes result and require interpretation and representation. Change assessment implies analysis of temporal transition, therefore it is advantageous to present the outcome of the analysis in a manner that communicates the results of this process. A change detection matrix (Jakubauskas *et al*, 1990; Jensen *et al*, 1993; Martin and Howarth, 1989; Pilon *et al*, 1988) provides a convenient means of summarising all land cover changes between sensing epochs.

The following change detection matrices represent various levels of change evaluated throughout this research. Each matrix lists all classes from the first epoch on the left side and all classes from the second epoch across the top. Each element of the change matrix then represents a transition sequence, the diagonals representing unchanged pixels and the off-diagonals representing pixels in transition. Element contents may contain pixel counts, direct area measurements, percentages or proportions depending on the interpretation objective.

Tables 6.12 and 6.13 represent changes in land cover from 1972 to 1991 using the results of a Level I classification. The magnitude of change in hectares for each cover class is recorded and it is possible to determine the main change trends, which include *Forest* to *Urban* and *Forest* to *Open* in the 1972 to 1986 period and *Open* to *Urban* in the 1986 to 1991 period. The change matrix provides the opportunity to highlight significant groups of classes and explain the relationship using a typical legend structure. The important change transitions from any class to *Urban* and *Open*, and the no-change classes have been highlighted, in this case emphasising the groups of *New Urban*, *New Open* and *No-Change*, respectively.

Table 6.14 provides similar information at both Level I and Level II, and Table 6.15 provides information at Level II. The approach used in Table 6.14 was described in Section 6.3 and was developed in this research to specifically facilitate the integration of multiscale data where the data are interpreted at different levels of detail. The modified change matrix allows comparison of classes at the lower level of detail (Level I) which

is common to both datasets, but also permits comparison of the Level I classes from 1972 with the more detailed interpretation at Level II from the 1986 data. Identification of the major change trends is also possible, with *Forest* to *Residential* (as opposed to *Commercial*) and *Forest* to *Grassland* and *Bare ground* (but not *Recreation*), representing the main transitions. The same legend as was used in Tables 6.11 and 6.12 can also be applied to this matrix to emphasise transitions where major changes are taking place.

From 1972	To 1986			
	Urban	**Forest**	**Open**	**Water**
Urban	392.5	41.2	58.9	0.3
Forest	466.4	2855.6	553.1	5.2
Open	22.4	0.9	11.6	0
Water	0	0.8	0	1.1

New Urban New Open No-Change

Table 6.12 *Change assessment matrix at Level I representing land cover changes between 1972 and 1986 (measurements are in hectares)*

From 1986	To 1991			
	Urban	**Forest**	**Open**	**Water**
Urban	872.3	6.9	1.8	0.4
Forest	192.1	15.0	6.1	15.1
Open	556.8	5.8	59.6	1.0
Water	0.1	0	0	6.3

New Urban New Open No-Change

Table 6.13 *Change assessment matrix at Level I representing land cover changes between 1986 and 1991 (measurements are in hectares)*

From 1972	To 1986								
	Urban		Forest			Open			Water
	Res	Com	Pine	Wet	Wood	Grass	Rec	Bare	Water
Urban	321.3	20.3	13.7	1.8	25.6	16.3	30.5	12.1	0.3
	341.6		41.2			58.9			
Forest	382.9	83.6	142.8	67.2	75.6	295.9	91.7	165.5	5.2
	466.4		285.6			553.1			
Open	15.5	6.9	0.5	0	0.4	2.9	4.7	3.8	0
	22.4		0.9			11.4			
Water	0	0	0	0.8	0	0	0	0	0.8
	0		0.8			0			

New Urban New Open No-Change

Table 6.14 *Change assessment matrix representing land cover changes between Level I (1972) and Levels I and II (1986) (measurements are in hectares)*

Table 6.15 provides full details of all nine land cover classes derived from the 1986 and 1991 classifications. Major land cover change transitions are identified using the legend and shading techniques described above, and inspection of the area values permits identification of change trends. Conversion from *Pine plantation*, *Grassland* and *Bare ground* to *Residential,* and from *Grassland* to *Bare ground* represent the major trends. Apparent changes from *Residential* to *Pine plantation* and *Grassland* do not represent logical transitions and are probably due to classification and registration errors.

Change matrices provide an efficient way of summarising a large quantity of information derived from the analysis of multidate images. For Level II there are 81 separate change classes represented in the change matrix, and the presentation of the data means it can effectively be understood. The matrix arrangement permits assessment of change patterns amongst classes, and related change transitions are grouped via the legend, which facilitates interpretation of the matrix. The change matrix does not link the change transition classes to the spatial distribution of pixels,

however the analysis is based upon pixel by pixel analysis rather than area-based comparison so that recorded changes represent actual transitions from one class to another.

From 1986	To 1991								
	Res	**Com**	**Pine**	**Wet**	**Wood**	**Grass**	**Rec**	**Bare**	**Water**
Residential	511.9	27.9	40.1	1.0	23.3	69.2	25.3	15.6	0.2
Commercial	25.1	104.9	5.0	0.4	5.6	14.3	3.2	1.8	0.2
Pine Plant.	47.2	9.2	47.6	2.5	7.5	30.9	2.8	8.1	0.6
Wetland	5.7	0.8	14.6	23.3	11.9	1.0	4.0	0.9	7.1
Woodland	28.8	5.6	23.6	3.2	22.7	9.4	5.0	2.2	0.5
Grassland	161.0	9.7	10.4	0.7	3.2	27.5	6.6	45.1	0.5
Recreation	17.8	2.2	1.4	0.8	2.3	4.1	25.9	1.6	0.4
Bare Ground	118.6	6.7	1.2	0	0.7	20.4	1.8	30.5	0.8
Water	0	0	0	0	0	0	0	0	6.6

New Urban New Open No-Change

Table 6.15 *Change assessment matrix at Level II representing land cover changes between 1986 and 1991 (measurements are in hectares)*

Effective an complete interpretation and communication of quantitative change and spatial distribution may only be made using change maps in conjunction with a change matrix. A major shortcoming of all of these techniques is that no reliability measure is built into the change reporting methods, therefore any reported changes must be considered in the context of the change detection accuracy assessments computed in Section 6.4.

6.6 Summary

Research in this chapter has been directed towards establishing an effective approach to change detection through the implementation of postclassification comparison of multiscale satellite data and appropriate evaluation and presentation techniques.

Interpretation of multiscale satellite data was investigated to take advantage of a range of past and current remote sensing satellites capable of producing continuous data streams, and to develop an approach for integration of historical satellite data for change detection. Many change assessment programs are constrained because of the requirement to utilise data from a single source for the interpretation of land cover variation. This research has provided a mechanism for integration of multiscale data for change detection using conventional image processing approaches.

Image algebra, spectral transformation and postclassification comparison approaches to change detection were investigated. Most research utilises image algebra and spectral transformation techniques for change detection, without considering the opportunities provided by multiscale satellite data. This chapter investigated the implications of utilising multiscale data, and determined that postclassification comparison is the most appropriate approach to employ with these data. The multisensor approach results in the incorporation of data with different radiometric response patterns and calibration standards, a disparate range of spectral bands and significant variations in spatial resolution. Postclassification comparison is the only standard image processing approach able to accommodate these variations.

Chapter 5 rigorously examined classification accuracy procedures incorporating Overall Classification Accuracy and Kappa Coefficient evaluation methods. These same approaches were used for evaluation of change detection accuracy, however due to the added complexity of the large number of classes resulting from change assessments, the Kappa analysis was not applied to the Level II comparison of Landsat MSS and Landsat TM data.

Change assessments using multiscale satellite data for two periods encompassing 1972 to 1986 and 1986 to 1991 were performed at Level I and Level II. Based on the evaluation of the Kappa Coefficient, change assessments were regarded as *moderate* for all assessments at Level I and for the 1986 to 1991 change assessment at Level II. Values of the Kappa Coefficient in the range of 43 to 55 percent were achieved for these change detection images, with Overall Change Accuracy values ranging from 51 to 70 percent.

Development of an approach to integrate multiscale datasets was an important consideration in the change analysis process. Understanding the importance of maintaining a hierarchical classification scheme resulted in the development of a method for seamless comparison of multidate Level I and Level II classifications. Integration of the classifications at two levels enabled combined change analysis involving data from both the Landsat MSS and Landsat TM sensors, while maintaining the quality and integrity of the individual datasets. Combined analysis provided an effective means of tracing land cover evolution throughout the period of data collection, and enabled the effective comparison of multiple datasets observed at spatial resolutions ranging from 20 m through to 79 m.

A range of change reporting techniques was investigated to provide informative and effective details of the magnitude and distribution of change identified from the satellite data. Reporting techniques included area-based change summaries, maps of change class distribution and change matrices detailing change areas and transition sequences. No single approach was able to convey all of the required information, and effective change could only be communicated via a combination of techniques including change maps and change matrices. Change matrices also enabled the incorporation of multilevel land cover change assessments derived from the multiscale data.

Chapter 7

ERROR ASSESSMENT AND PROPAGATION

Assessment of error components within the image processing and change detection processes, and evaluation of their contribution to the overall quality of the final assessment represents the final stage of change detection. Benchmarks for the assessment of change detection accuracy are discussed and the significance of the spatial and thematic accuracy of the change data considered. The propagation of errors from the source data through to the final change products is investigated through consideration of error models appropriate to data used in this research. An evaluation of the processing components contributing to the spatial and thematic errors is undertaken and approaches to error propagation are discussed.

7.1 Introduction

Assessment of errors has been made separately at the image rectification, classification and change detection stages of this research, but the combined effects and propagation of the error components have not been considered. Error propagation is important because processing leads to the combination of errors, such that the resultant error is significantly larger than any of the component values. MacDougall (1975) claims that due to the propagation of errors the results of some overlay operations produce maps little different from random class allocations. Subsequent research has proven this assertion to be overstated (Newcomer and Szajgin, 1984), however the implications of error propagation warrant further investigation.

All data within a GIS contain a certain amount of error due to measurement, classification, recording, generalisation, interpolation or interpretation errors (Heuvelink, 1998). Walsh *et al* (1987) consider errors in spatial data to arise from *inherent* and *operational* sources. Inherent errors are those present in the source documents while operational errors are due to data capture and manipulation functions within the GIS. Heuvelink *et al* (1989) describe operational errors in terms of processing and modelling errors, and inherent errors as source errors. From the results of image processing, Hord and Brooner (1976) suggest errors arise mainly from boundary location, map geometry and data classification. Where GIS analyses are also involved, this may be extended to include the concepts of error propagation between these sources.

Many error metrics presently in use provide global characterisation of error within the data and apply them uniformly to an entire region or image, and do not take into account localised variability (Edwards and Lowell, 1996). These include RMSE assessments of image rectification and values of Kappa Coefficients for thematic accuracy assessment. Local metrics are more common in applications of fuzzy classifications where individual uncertainties associated with data elements, such as boundaries and contextual components of class membership, are expressed. The importance of local characteristics is indicated by Kiiveri (1997) where decision making that ignores uncertainty may neglect potentially useful areas, and also mistakenly select areas not having the attributes of interest. Site-specific assessments of error (and

therefore uncertainty) provide enhanced discriminating capabilities (Openshaw *et al*, 1991).

Understanding the nature of error in spatial data is necessary to ensure the development of relevant analysis techniques and provide confidence in the quality of outcomes (Chrisman, 1991). Openshaw (1989) describes a GIS as an accurate and virtually error-free environment for manipulating map data, however the data are often of variable precision and different combinations of operations may produce results of unknown quality. While error does not need to be eliminated from GIS manipulations, an understanding of the magnitude and the implications of its presence are essential.

Error is considered to comprise the multiple dimensions of positional accuracy, attribute accuracy, lineage, logical consistency and completeness (Lanter and Veregin, 1992). These dimensions are designed to provide *labels* indicating data characteristics, but it is for the user to interpret the significance of the information, especially with respect to error propagation and data quality. This chapter considers aspects of error related to positional and attribute accuracy.

Veregin (1989) has developed a hierarchy of needs for modelling error in GIS operations. The hierarchy includes error source identification, error detection and measurement, error propagation modelling, strategies for error management and, finally, strategies for error reduction. Error reduction strategies are considered last and have not received a high priority in the modelling process. Identification, detection and measurement of errors continue to maintain a high priority and will be considered in this research. Errors are generally considered to comprise positional or cartographic errors, and identification, thematic or attribute errors. Within this research the terms *spatial* and *thematic* errors will be used to describe errors in position and attribute values, respectively.

The process of dealing with error in spatial data may be described in three basic steps (Forier and Canters, 1996). Identification of the type of error information available for each data layer is the first step, and will be described in Section 7.3. Whether the error is local or global in nature affects the method of analysis. The second step is concerned with defining a conceptual model of error compatible to the type of error identified in

the first stage (see Section 7.4). A global error model relies on a single statistic, such as the Overall Classification Accuracy, which assumes the error levels are uniform for all classes. A local error model considers each class to have specific error characteristics and includes this information within the error model. The final step requires derivation of an error propagation function that describes the manner in which the errors are propagated through the data analysis system. This phase will be considered in Section 7.5.

7.2 Benchmarks of Spatial and Thematic Errors

Comparative assessments of error in mapping are usually restricted to generalised but unambiguous statements of positional accuracy or presentation of reliability diagrams. Topographic mapping standards include statements such as *90 percent of all elevations determined from topographic maps shall have an error not exceeding one third of the contour interval for well defined points* and that, *planimetric detail shall be within 0.25 millimetres of true position at map scale* (Merchant, 1987). Similar explicit standards relative to the accuracy of boundaries of discontinuous thematic data such as soil classes, land use or land cover information are not generally available (Drummond, 1987). Anderson *et al* (1976) specify that land use and land cover maps derived from remotely sensed data should achieve a minimum accuracy level of 85 percent in identifying categories, and that interpretation accuracy across categories should be about equal. They also indicate that consistent results should be obtainable between interpreters and from the results of sensing at different times. These standards were initially regarded as advisory rather than mandatory, and were not subject to rigorous testing (Fitzpatrick-Lins, 1978).

Continued development of data acquisition and interpretation systems has focused on the geometric fidelity of the sensing system, and the enhanced capabilities of analogue and digital interpretation techniques. Enunciation of a coherent spatial and thematic accuracy standard has not been forthcoming, although several authors have stated separate guidelines for assessment of the spatial and thematic accuracy of products derived from remotely sensed data.

Welch (1985) provides specifications of ± 0.3 pixel for image-image rectification and ± 0.5 pixel for geodetic rectification. This value for geodetic rectification is also

reported to represent the NASA standard (Labovitz and Marvin, 1986). Hill and Aifadopoulou (1990) achieved similar results, but indicate that local misregistration may still reach 1.0 - 1.5 pixel. Objectives for the NALC program provide for geodetic rectification with RMSE of 1.0 pixel within and 1.5 pixel outside the United States, and image-image registration of 0.5 pixel (Lunetta *et al*, 1998). Riley *et al* (1997) report that some scenes supplied for the NALC program have geodetic rectification errors exceeding 1.0 pixel and Townshend *et al* (1992) indicate that geodetic rectification within 0.5 - 1.0 pixel is normally regarded as acceptable, especially for change assessment.

While no specific rectification standards apply to these images, it is clear that geodetic rectification within 0.5 pixel is achievable in most cases and is suitable for identification of thematic classes and change assessment analyses. Where these data are used as input for topographic mapping programs the relevant mapping standard must be applied to the data.

Standards for mapping accuracy of thematic classes have also received little attention. Most guidelines are based upon those proposed by Anderson *et al* (1976) and vary between 80 and 90 percent accuracy. The CoastWatch Change Analysis Project (C-CAP) established guidelines of 90 percent for thematic accuracy of all categories, however investigation by Jensen *et al* (1993) recommends it be relaxed to 85 percent because the higher value was not achievable using Landsat TM data. Guidelines for thematic accuracy within the NALC were set at 80 percent accuracy with a 75 percent confidence interval using Landsat MSS data.

Thematic accuracy standards must be considered within the context of the level of detail of the classification scheme and the spectral, spatial and radiometric resolution of the data. The NALC program is interpreted at Level I and is able to achieve the guideline with medium resolution Landsat MSS data. The C-CAP project uses a Level II classification and is able to achieve a marginally higher accuracy, but requires the improved spectral, spatial and radiometric qualities of Landsat TM data.

Table 7.1 summarises results from a range of postclassification comparison change detection studies extracted from recent literature. Some authors have not reported

complete details of accuracy assessments for rectification, single date image classification or change assessment. Jakubauskas *et al* (1990) did not provide any quality assessment details and simply compared proportions of land cover change, while Laba *et al* (1997) reported classification accuracies only for data from individual images. Star *et al* (1997) suggest that spatial and thematic reliability diagrams need to be developed for remote sensing and GIS that convey the reliability of information through the spatial characteristics of uncertainty in information products.

Author	Data	Classification Level	Accuracy	
			Rectification (pixels)	Thematic (percent)
Jakubauskas *et al* (1990)	Landsat MSS Landsat TM	Level II 9 classes	-	-
Macleod and Congalton (1998)	Landsat TM	Level II 2 classes	±0.49	39 - 43 27*
Jensen *et al* (1993)	Landsat TM	Level II 9 classes	±1.0	85 - 90 82*
Laba *et al* (1997)	Landsat MSS Landsat TM	Level II 6 classes	±0.7 - 1.1	73 - 77
Martin and Howarth (1989)	Spot HRV	Level I 5 classes	<1.0	50*
Mas (1999)	Landsat MSS	Level I 7 classes	<1.0	71*
Muchoney and Haack (1994)	SPOT HRV	Level I 2 classes	-	31*

* Kappa Coefficient for change detection product

Table 7.1 *Summary of spatial and thematic accuracies achieved during satellite-based land cover change detection projects (thematic accuracy represented by Kappa Coefficient (percent))*

Rectification accuracies generally less than one pixel were achieved in these studies, however it is not clear in all cases whether the accuracy results from *a priori* or *a posteriori* evaluation of the data. Table 3.2 shows the rectification results obtained in this research with *a posteriori* RMSE values less than one pixel for all datasets, and those for the higher resolution Landsat TM, SPOT HRV and IRS1-D data in the order

of 0.5 pixel. These results are consistent with the general rectification standards indicated in this section and provide a sound basis on which to evaluate the propagation of spatial and thematic errors in the change detection process. The finer spatial resolution IRS1-D data provide an indication of the geometric potential of new generation satellite sensors, which is consistent with the 0.5 pixel level achieved previously. This is largely due to the geometric integrity of the satellite data and the high accuracy of the reference data used for control point values. A complete evaluation of the image rectification process for these data is available in Chapter 3.

The variation in the thematic accuracy achieved for the projects in Table 7.1 represents a combination of the sensor characteristics and classification scheme discussed above. Laba *et al* (1997) and Macleod and Congalton (1998) achieved Kappa Coefficient values for single date images in the order of 75 and 40 percent, respectively. Both images were classified at Level II and the higher Kappa value was achieved when interpreting six classes, with the lower value from interpretation of only two classes. Similarly, other examples of accuracy assessments for postclassification change images listed in Table 7.1 indicate Kappa values ranging from 27 percent (two classes) to 82 percent (nine classes) at Level II, and from 31 percent (2 classes) to 71 percent (7 classes) at Level I. The inverse relationship between Kappa value and the number of classes is coincidental and an indication of the compound effect of spatial, spectral and radiometric properties on target separability.

Tables 5.15 and 5.16 detail the accuracies for data classified during this research. Kappa values between 50.4 percent and 62.5 were achieved at Level I, and at Level II Kappa values between 36.9 percent and 59.1 were obtained. Change detection Kappa values are 55.0 percent and 54.0 percent for Level I, and 43.9 percent for Level II. These results conform to the range of values included in Table 7.1, however the same extreme high and low values have not been achieved in this study. The decreased accuracy at Level II and the accuracy of the single date classifications, compared with the change images, demonstrates that the comparative level of detail included within the Level I and Level II classification schemes is consistent with the classification techniques employed.

No clear benchmarks for thematic accuracy may be universally determined due to variations in the separability of different combinations of targets, even though they may be located on the same level of the classification scheme. Specific labelling of the multiple dimensions of error as suggested by Lanter and Veregin (1992) may enable users to interpret the significance of the error data and the suitability of the product for a specific application. In the absence of a single convenient measure of spatial and thematic accuracy, additional statements of lineage, logical consistency and completeness provide supporting evidence of data quality.

7.3 Error Sources

Analysis of the image rectification process in Chapter 3, and thematic classification and change assessment processes in Chapters 5 and 6, confirms the accuracy of change assessment is subject to a range of interrelated errors. These errors are due to such factors as image geometry, detail of thematic classification, scale of observation, boundary definition, target spectral and spatial properties, and sensor characteristics. Aspinall and Hill (1997) regard these errors to be mainly related to misidentification of classes, positional inaccuracy in boundary location and failure to recognise internal polygon heterogeneity. Chrisman (1987) describes these factors as resulting in errors of *identification* (errors in assigning the correct attribute) and *discrimination* (errors in separating adjacent types). In particular, Chrisman (1987) notes that discrimination errors are another term for describing the fuzzy boundary problem and those errors are impossible to distinguish from spatial error. The effect of these errors is evaluated through analysis of their spatial and thematic components. The correlation between the two types of errors is also important (Star *et al*, 1997), and will be considered. A taxonomy of errors and their relationship to error modelling is discussed in Section 7.4.2, however in this section errors specific to this research are evaluated.

Propagation of errors will be undertaken to analyse the effects of component variables on the classification and change assessment estimates. Standard approaches to propagation of quantitative random errors are applied, assuming the correlation between arguments z_i is zero, according to the following model (Heuvelink *et al*, 1989):

$$\sigma^2 = \sum_{i=1}^{n} \left\{ \sigma_i^2 \left(\frac{\delta g}{\delta z_i}(\mu) \right)^2 \right\} \qquad (7.1)$$

where

$y = g(z_1, z_2 \ldots\ldots z_n)$ *is a continuously differentiable function with arguments* z_i.

Spatially-related errors in both the reference data and primary remote sensing data contribute substantially towards the total spatial error within change assessment systems. Reference data are subject to errors of land cover identification, however these may be minimised through careful ground sampling. The largest errors in photointerpretation are derived from boundary identification and are usually of greater magnitude than the errors induced through the digitising process (Edwards and Lowell, 1996). Table 7.2 details estimated errors for derivation of reference and image data.

Sensor	Reference Data				Image Rectification Error	Image-Ref Error
	Control Error	Photo Int Error	Digitising Error	Total Error		
Landsat MSS	1.25	12.5	5.0	13.5	65	66.4
Landsat TM	1.25	10.0	5.0	11.25	13	17.2
SPOT HRV	1.25	10.0	5.0	11.25	8	13.8
IRS1-D	1.25	1.8	1.25	2.5	3.2	4.1

Table 7.2 *Assessment of boundary error accumulation between reference data and single date satellite images (all values are in metres)*

Additional sources of spatial error in the reference data include error in the digitising transformation (*Control Error)* and *Digitising Error.* The digitising transformation has a maximum RMSE of 1.25 m (0.25 mm) at a map scale of 1:5 000, and represents the effects of control point accuracy and differential media distortion. The *Photointerpretation Error* is based upon an error of 0.5 mm at the scale of the aerial photography. This only represents a geometric consideration of error and does not estimate the error in determination of boundaries caused by separability of class characteristics. This error is additional but not quantifiable at this stage.

The *Digitising Error* is estimated at 1.0 mm at map scale (5.0 m) and represents a combination of the actual digitising accuracy and the accuracy of boundary transfer from the aerial photographs at 1:25 000 (1972) or 1:20 000 (1986 and 1992), to the orthophotmaps at 1:5 000. Some additional uncertainty was introduced in the transfer of detail between the aerial photographs and the orthophotomaps because the aerial photographs were available for each image date (1972, 1986 and 1992), whereas the orthophotmaps were only available for 1984. The total error from these sources is estimated at 13.5 m for the 1972 reference data, and 11.25 m for the 1986 and 1992 data.

Evaluation of image rectification errors has been undertaken in Chapter 3 with the main contributing factors being GCP location error, GCP digitising error, relief error and displacement due to image resampling. The resultant *Image Rectification Error* is shown in Table 7.2, with an estimate of the combined spatial error *(Image-Ref Error)* between the satellite image data and the reference data indicated. The rectification RMSE for Landsat MSS data is comparatively large at 65 m, and dominates the outcome of the analysis. The value is approximately the same magnitude as the original image pixel (0.875 pixel). Comparative results for the Landsat TM, SPOT HRV and IRS1-D data provide *Image-Ref error* values of approximately 0.6 pixel in each case.

Extension of change detection to incorporate high resolution satellite images such as IRS1-D data requires further evaluation of potential error sources. The scale of aerial photography is a limiting factor for the collection of reference data and currently comprises the major source of spatial error. Availability of larger scale aerial photography for reference data collection will lead to lower combined photointerpretation boundary transfer error and digitising error because the differential scale between the aerial photography and orthophotmaps will be substantially reduced. The relationship between component spatial errors in reference data collection is given by:

$$(Total\ ref.\ data\ error)^2 = (Control)^2 + (Photointerpretation)^2 + (Digitising)^2 \qquad (7.2)$$

The total spatial error in the reference data should not exceed the equivalent of 0.5 image pixel, and the RMSE of the *Control Error* remains at 0.25 mm. The *Digitising Error* can be reduced to 0.25 mm at map scale by using larger scale aerial photographs as discussed above. This has the effect of reducing the transfer error between the photography and the orthophotomaps caused by the large scale differential. Applying Equation 7.2, the maximum permissible *Photointerpretation Error* for the IRS1-D data is given by:

$$\begin{aligned} Photointerpretation &= \left((Total\ ref.\ data\ error)^2 - (Control)^2 - (Digitising)^2\right)^{0.5} \quad (7.3) \\ &= \left(2.5^2 - 1.25^2 - 1.25^2\right)^{0.5} \\ &= 1.8\ m \end{aligned}$$

With a maximum spatial error from photointerpretation of 1.8 m, which represents 0.5 mm at photoscale, the aerial photography for reference data collection needs to be at least 1:3 600 scale or, alternatively, utilisation of GPS data of equivalent standard for reference sample location.

Similar analysis of errors for change detection purposes requires separate consideration of errors in the reference data and image data, then combined analysis of the errors for each change period. Table 7.3 provides an analysis of the errors for the two change detection periods actually investigated, and hypothetical values for inclusion of IRS1-D data. Values for *Reference Data Error* and *Image-Image Error* are derived from the application of Equation 7.1 to the corresponding reference data and image data from Table 7.2. The spatial errors within Landsat MSS data dominate the error estimate for the change detection assessment incorporating Landsat MSS and Landsat TM data (68.6 m). Errors inherent in the Landsat TM and SPOT HRV data, and the corresponding reference data, produce an overall change error estimate of 22.0 m. The corresponding estimate for the SPOT HRV and IRS1-D comparison is 14.4 m.

Each of the RMSE values in Tables 7.2 and 7.3 represents an estimate of the total error included when comparisons between corresponding satellite and reference data are made for the purposes of classification accuracy analysis (Table 7.2) or change assessment (Table 7.3). These data are important when pixel by pixel comparisons are

made or, more importantly, when comparisons in the vicinity of boundaries are involved. Pixel by pixel comparisons within areas of uniform cover type are not affected by spatial errors between the reference data and primary image data because spatial displacement within invariant land covers does not alter the outcome. When pixels close to land cover boundaries are subjected to comparison, the displacement error may cause erroneous comparisons from opposite sides of the boundary to be made, with the result that thematic classification or change assessment errors are recorded.

Change Detection Pair	Ref. Data Error	Image-Image Error	Total Error
Landsat MSS – TM	17.6	66.3	68.6
Landsat TM – SPOT HRV	15.9	15.3	22.0
SPOT HRV – IRS1-D	11.5	8.6	14.4

Table 7.3 *Total boundary error between reference data and multidate change images (all values are in metres)*

It is well documented that errors in thematic feature extraction are most prevalent near boundaries (Dunn *et al*, 1990; Edwards and Lowell, 1996; MacDougall, 1975; Middelkoop, 1990; Skidmore and Turner, 1992), and that positional uncertainty implies the prevalence of attribute errors near boundaries when comparing overlays (Kiiveri, 1997).

The most common approach to boundary error assessment is using the *epsilon band*, which is frequently utilised to represent the error in boundaries resulting from the digitising process (Blakemore, 1984). In the current study the definition of the epsilon band has been extended to represent the region of error surrounding polygon boundaries. The dimensions of the epsilon band are derived not only from the errors inherent in the digitising process, but also include estimates of errors in boundary location resulting from multiscale digital image rectification (Table 7.2). The concept of the epsilon band utilised here includes only the spatial considerations in the location of boundaries and does not include errors caused by boundary pixels that are a mixture of cover types.

The definition of the parameters of the epsilon distance may be cast in deterministic or probabilistic terms (Dunn *et al*, 1990). The parameters may be represented by a single value or by the application of fuzzy set theory to define the uncertainty in the width and shape of fuzzy boundaries (Edwards and Lowell, 1996). Estimates of the width of the epsilon band made in this research are based on global error assessments of the data, and provide a deterministic estimation of the epsilon distance. The ability to move from dependency on aspatial global error assessments (confusion matrices, digitising errors and rectification errors) will also improve estimation of the shape and width of the epsilon band.

The epsilon distance as a measure of boundary location error is used in this research to assess the degree of potential error within the study area for each image dataset. Table 7.4 shows the length of boundaries at Level I and Level II for each image dataset and the resulting area of the epsilon band when multiplied by the *Image-Reference Error* calculated in Table 7.2. Reference data at Level I and Level II have been used to determine the lengths of boundaries because of the greater reliability and spatial consistency of the data. The external boundary of the study area was not considered. Further generalisation was undertaken with a 7 x 7 modal filter to eliminate narrow and fragmented classes, and to minimise the overlap of epsilon bands that tend to inflate the area due to complexity of classes and their boundaries (Chrisman, 1987).

Sensor	Level I			Level II		
	Bdy Length (metres)	**Bdy Area (hectares)**	**Percentage of Area**	**Bdy Length (metres)**	**Bdy Area (hectares)**	**Percentage of Area**
Landsat MSS	50 920	338	18.3	71 260	473	25.6
Landsat TM	66 480	114	6.2	95 700	165	8.9
SPOT HRV	60 520	83	4.5	80 580	111	6.0

Table 7.4 *Percentage of study area occupied by epsilon band derived from planimetric errors in reference data and multidate images*

The data in Table 7.4 provide an estimate of the area of uncertainty and percentage of the total area included within the epsilon band for each dataset at Level I and Level II. These parameters have been used elsewhere to estimate the area of uncertainty due to digitising (Dunn *et al*, 1990; MacDougal, 1975), and clearly indicate the substantial

uncertainty introduced to the interpretation of the Landsat MSS data due to reference data and image rectification errors. The land cover classes included within the reference data for the Landsat MSS data are relatively general, as evidenced by the comparatively low total boundary lengths compared to the other datasets at both levels of interpretation. Although the land cover boundaries are more complex in the Landsat TM and SPOT HRV data, the comparatively narrow epsilon bands reduce the area of uncertainty from 25 percent for Landsat MSS data to between four and nine percent of the total area for Landsat TM and SPOT HRV data.

Comparison of the boundary areas computed from the epsilon bands with the results of classification assessment in Tables 5.15 and 5.16 indicates that boundary uncertainty is only one of the potential factors influencing classification accuracy. While Landsat MSS data produce lower classification accuracy results, the differences are not entirely due to the boundary uncertainty indicated by the result in Table 7.4. Other variables related to the spectral, spatial and radiometric resolution of the data must also have a considerable influence on classifier performance as the epsilon band area and Kappa Coefficients do not exhibit consistent variation. The other major factor that is considered important relates to the increased complexity of land covers in the Landsat TM and SPOT HRV data

The spatial distribution of classification errors displayed in Figure 5.3 have also been investigated to assess the relationship between boundary uncertainty and error patterns within the land cover classes. Table 7.5 compares the pattern of error cells represented in Figure 5.3 with the land cover boundaries defined by the reference data. The boundaries derived for each reference data set have been rasterised and buffered to a width equivalent to the *Image-Ref error* computed in Table 7.2. The dimensions are restricted to integer pixel values and are approximated for Landsat MSS at 130 m, Landsat TM at 35 m and SPOT HRV at 25 m total width. The percentage of error cells that coincide with the land cover boundary error bands are tabulated against the overall classification error derived from the complement of the Overall Classification Accuracy for a Level II unsupervised classification. A χ^2 Goodness of Fit test of the proportion of classification errors occurring in proximity to the land cover boundaries against the overall proportion of classification errors indicates a significant difference at the 95

percent confidence level. Consequently the hypothesis that classification errors are spread evenly throughout the study area is rejected and therefore, classification errors occur more frequently in the proximity of boundaries compared to elsewhere.

Sensor	Percentage Errors Near Boundary	Overall Percentage Class Error
Landsat MSS	54.8	41.6
Landsat TM	59.1	38.8
SPOT HRV	59.4	29.9

Table 7.5 *Comparison of near-boundary errors and overall classification error rate*

Approximately 60 percent of all classification errors occur within the boundary error band determined for each dataset. These results indicate that classification errors are not uniformly distributed throughout the remotely sensed datasets and that the spatial location of errors is correlated with the location of the land cover boundaries. The spatial autocorrelation of thematic classification errors at distances up to 200 m was examined in Chapter 5 and determined to be significant at the 99 percent confidence level. In conjunction with the results in Table 7.5 it supports the hypothesis that classification errors are more prevalent in the proximity of class boundaries.

The error band examined above is based upon the global propagation of spatial error components determined for each satellite dataset and does not account for local variations on a class by class basis. Consideration of individual thematic class membership probability or uncertainty of boundary location using fuzzy concepts may be useful in further analysing the patterns of classification errors (Zhang and Foody, 1998), however this is beyond the scope of this research.

7.4 Error Modelling

7.4.1 Types of Errors

Error is defined in terms of discrepancies between the interpreted objects and their *real world* counterparts. Conventional error theory categorises observation and measurement errors as systematic, random or gross errors (Mikhail and Gracie, 1981). Systematic errors occur according to some deterministic process that may be expressed

by a functional relationship if it is known. Random errors result from observational errors that have no known functional relationship based on a deterministic model, and are analysed using probabilistic models. Gross errors are caused by mistakes in measurement and analysis. Error modelling approaches are based on the control of systematic errors through careful design of data collection strategies and instrument calibration procedures. The elimination of gross errors relies upon maintenance of data collection and analysis quality standards designed to eliminate extraneous and avoidable error sources. The magnitude of random errors is dependent on the resolution capabilities of the measurement system and subsequent propagation through the error model. Robust observation and processing systems are designed to take account of systematic and gross errors, leaving random errors for consideration through error modelling. However, inappropriate model development or presence of systematic or gross errors of unknown magnitude may lead to invalid error evaluations.

With respect to the interpretation of remotely sensed data, the elements of resolution include the spatial resolution of the sensor, spectral resolution of the multispectral bands, and the radiometric resolution or quantisation of individual bands. These elements are convolved with the interpretation system and random errors contained within the reference data to produce an aggregate of errors in the spatial, thematic and temporal domains (Veregin, 1995).

GIS models may be categorised as mathematical or logical models (Drummond, 1987), and are also termed quantitative and qualitative models, respectively. Mathematical models utilise deterministic processes to evaluate the outcome of component variables as an aid in decision making. Error propagation is based upon the general law of propagation of variances and assumes only random errors are present in the data (Mikhail and Gracie, 1981). Such an approach was applied in Section 7.3 for analysis of propagation of spatial errors within the image and reference data.

Logical models use a heuristic approach to evaluate the outcome of analyses, such as a system of Boolean overlays for land suitability analysis. Logical models do not lend themselves to the analysis of error propagation through the general law of propagation of variances (Burrough, 1986), and alternative techniques, such as simulation, must be investigated (Heuvelink and Burrough, 1993). Within this research multitemporal

comparison of satellite data for land cover change represents the application of a logical model.

Logical models are characterised by a number of spatially coincident overlays with each representing a particular theme or characteristic of the area. In this research separate overlays contain land cover information interpreted from aerial photographs or classified from multidate satellite images, or change assessments derived from logical differences created from corresponding land cover layers. Each layer comprises nominal data representing land cover or land cover change classes. Error is conventionally measured in terms of the frequency of differences between the actual and observed classes over the spatial, temporal and thematic domains (Veregin, 1995).

Within the logical error modelling process, error may be evaluated at the generalised layer level, at the individual class level within each layer or at the specific pixel level, depending on the error assessment data available. At the most general level an error index provides an assessment of an entire layer, however no information concerning variations in error in the spatial, temporal or thematic domains is available. Alternatively, at a more specific level the error in each feature in a layer may be quantified, and variations in error over space, time and theme may be portrayed (Veregin, 1995).

Definition of geographical entities in terms of space, time and theme results in the portrayal of the corresponding error in each dimension. The inability to adequately characterise spatial data with a single index of error arises due to the multidimensional nature of these errors (Lanter and Veregin, 1992). For example, Hord and Brooner (1976) acknowledge this issue and suggest the accuracy of land use maps be quantified in multiple dimensions by adopting separate standards for thematic classification accuracy, and spatial accuracy of boundaries and control points. A more contemporary view is that thematic accuracy can be a function of both class membership and spatial location, and it may not be appropriate to examine spatial and thematic error independently (Chrisman, 1989; Haining and Arbia, 1993; Mark and Csillag, 1989; Veregin, 1995).

Central to the issue of error modelling in spatial data is the correlation of errors within or between thematic classes. Many studies assume random occurrence of errors, however in practice errors are often correlated (Heuvelink and Burrough, 1993). Where a random (spatially independent) distribution of errors is assumed, simple probability may be applied to assess error propagation (Veregin, 1995), however where spatial autocorrelation is present, more sophisticated simulation modelling is required to account for the spatial dependence of errors (Heuvelink and Burrough, 1993). Spatial autocorrelation of errors is manifest in thematic classifications by aggregation of errors in the proximity of boundaries or the propensity for errors to occur in spatially adjacent locations. Veregin (1995) indicates that spatial coincidence between layers may either inflate or deflate the accuracy of derived data, depending on the degree of dependence and the type of overlay operation. Recognition of the degree and spatial dependence of errors within the data is essential to the success of error propagation within spatial data systems. These effects will be considered in conjunction with error propagation in Section 7.5.

7.4.2 Error Models

The major sources of error within the change assessment process developed in this research have been identified and their significance evaluated in Section 7.3. Consideration of suitable error models is required with a view to identifying approaches suitable for evaluation of error propagation during change assessment.

GIS provides an error-free environment for data manipulation, however the input data are usually characterised by differing levels of error and uncertainty which need to be identifiable in the final output (Openshaw, 1989). Error reduction is one part of the process, but does not provide a complete solution. Development of adequate means of representing and modelling uncertainty and error in spatial data, and the availability of GIS techniques for implementation of error models are equally important (Openshaw, 1989).

Generally three types of error models are applicable to categorical data based upon the type and extent of error information available. Forier and Canters (1996) define error models as *global, thematic* or *spatial/thematic*. Global error models assume that errors occur uniformly for all classes and that confusion is equal among each pair of classes.

Statistics such as the Overall Classification Accuracy and Kappa Coefficient are utilised in global error models to describe error characteristics, but provide no information beyond their overall magnitude. The spatial distribution and variation of errors by class are not considered.

Thematic error models rely upon error information that describes the variations in the level and type of error for each class. Veregin (1995) has developed a thematically partitioned error model that utilises the complete confusion matrix as a descriptor of the nature of errors in the data. The advantage of using the confusion matrix is that it is a crosstabulation of the encoded and actual attribute values of the sampled locations, which are representative of each class. Error modelling using the entire matrix maintains important information beyond the percent correctly classified, and includes the level of omissions and commissions for each class. Subsequent analysis of the propagated confusion matrix yields correspondingly detailed error information for the derived layers.

Thematic error models therefore consider the distribution and propagation of errors on a class by class basis, but take no account of within-class variations or *thematic purity* of classes. Additionally, in situations where even a moderate number of classes are subject to change assessment, detailed modelling using the confusion matrix requires consideration of a very complex range of interactions. For example, comparisons of multidate Level II classifications in this research, which each comprise nine land cover classes, results in the derivation of a confusion matrix involving 81 x 81 elements. The computational overhead using the thematic error model is significant and may not be justifiable.

Spatial/thematic error models require more specific information on the spatial distribution of error within each thematic class, and sometimes spatial correlation of error between data layers (Forier and Canters, 1996). Such conditions may arise when digitising and boundary definition effects related to reference data and image classification cause errors to be spatially correlated. Openshaw (1989) indicates that adequate means of representing and modelling the uncertainty and error characteristics of spatial data are required, as well as the development of relevant techniques that can explicitly take error into account during data processing.

Uncertainty in spatial data may be modelled in a number of ways. Gaussian models may be applied to describe errors, however they only provide a general context for description and simulation of errors, instead of a clear model of the processes of error accumulation (Goodchild *et al,* 1992). Where specific details of error formation and distribution are available, individual error models may be applied, such as the effects of manual digitising on the positional quality of data (Bolstad *et al,* 1990). Models considered for the representation of error in digitised line data include rectangular, bell-shaped and bimodal probability density functions (Dunn *et al,* 1990). These models may be applied to representation of boundary uncertainty using fuzzy approaches to describe the quality of the reference data, as well as to the remotely sensed data (Zhang and Foody, 1998).

Between these two extremes Goodchild *et al* (1992) provide a generic model where nothing is known about the specific processes contributing to the error. The model defines category error in terms of class heterogeneity (purity) and boundary transition (sharpness). Heterogeneity is represented by texture, which is defined by the degree of correlation between adjacent cells and determines the amount of aggregation and uniformity of each class. Boundary transition defines the distortion in the location of a boundary between two classes based on spatial dependence. Class heterogeneity and boundary transitions provide the means to model error (and uncertainty) in thematic data under different levels of spatial dependence.

Spatial/thematic error models generally utilise one of the above techniques for managing the magnitude and distribution of errors, but they do not describe processes for the assessment of the effects of spatial distribution of error during GIS operations. Probabilistic approaches to error propagation have been applied by Newcomer and Szajgin (1984), Veregin (1989; 1995) and others, however a new error model must be derived for each operation, and extension beyond straight forward operators such as AND, OR and buffer is limited (Goodchild *et al,* 1992). The lack of theoretical knowledge of generalised error propagation mechanisms for logical data models has led to the investigation of alternatives that can be universally applied to any data transformation function (Veregin, 1994). Monte Carlo simulation has been investigated as a suitable alternative and can be applied whether or not a formal error model has been developed for the specific function or operator (Openshaw, 1989).

The generality and flexibility of Monte Carlo simulation provides significant flexibility in understanding complex error propagation pathways within GIS (Openshaw *et al*, 1991). Monte Carlo simulation relies upon a large number of repetitions of the GIS operations under analysis, for which the input data have been subject to random perturbations based upon known or assumed pattern of errors. The results of the simulation may be analysed to determine error rates, confidence limits and credibility regions following propagation of errors. Although the approach is general and flexible in its application, the results do require interpretation and there is a significant processing overhead involved (Dean *et al*, 1997; Podobnikar, 1999), which may limit its application in real time error propagation analysis (Forier and Canters, 1996).

Error data defined in Section 7.3 may be utilised as input to several of the error models described above. Overall Classification Accuracy and the Kappa Coefficient are suitable for inclusion in global error modelling, while the complete confusion matrix may be used for class specific accuracy assessment and pairwise comparison of classes for establishment of errors of omission and commission, as input to thematic error modelling. Analysis of spatial errors near the boundaries of reference data through consideration of digitising errors, photointerpretation errors and control point determination may be used in spatial/thematic error models. Consideration of the proportion of classification errors occurring within the epsilon band defined by the accuracy of the image rectification and reference data also provide data suitable for spatial/thematic error modelling.

Further development of the error models for the classified data and reference data will enable detailed modelling of spatial and thematic errors within the change assessment process. In this initial stage the available error sources and models have been identified, however only global error modelling will be pursued through analysis of the available error data.

7.5 Error Propagation

Section 7.4 detailed a taxonomy of errors and the parametric measures required for describing error models applicable to thematic data. The specific approach to error propagation varies depending on the error model applied to the specific error distribution. Global error models assume random distribution of errors within each

layer and it is possible to apply a simple probabilistic approach to error propagation. Thematic error models use class-specific error parameters derived from the confusion matrix, but must still assume there is a spatially random distribution of errors. On the other hand, spatial/thematic error models take into account spatial correlation and utilise simulation approaches to incorporate correlation of errors into the propagation analysis. Based on the error data available from Section 7.3 the effects of global error modelling will be evaluated.

The quality of the results of logical operations may be predicted from the error characteristics of the composite data layers through modelling the analysis process (Newcomer and Szajgin, 1984). Overlay using the AND operator involves comparison of two geographically registered datasets with a correct result being achieved where there are correct values at corresponding locations in each data layer. Change detection between spatially registered temporal datasets represents an example of the AND operation. Bayes Theroem may be used to express the probability of a correct result occurring for the AND operation by the following (Newcomer and Szajgin, 1984):

$$Pr(E_2 \mid E_1) = \frac{Pr(E_1 \cap E_2)}{Pr(E_1)} \tag{7.4}$$

or

$$Pr(E_1 \cap E_2) = Pr(E_1)Pr(E_2 \mid E_1) \tag{7.5}$$

where

E_i = *occurrence of a correct value in layer i at a given location for* $i = 1,2,3.....n$

$Pr(E_i)$ = *probability that a correct value occurs in layer i at a given location,* $0 \le Pr(E_i) \le 1$

E_i' = *occurrence of an incorrect value in layer i at a given location* (E_i')

$Pr(E_i')$ = *probability that an incorrect value occurs in layer i at a given location for* $0 \le Pr(E_i') \le 1$ *and* $Pr(E_i) + Pr(E_i') = 1.$

Equation 7.5 indicates the probability of event E_1 (correct in layer 1) and E_2 (correct in layer 2) occurring together equals the probability of E_1 occurring multiplied by the probability of E_2 occurring given that E_1 has occurred. Further derivation provides the following form of Equation 7.5:

$$Pr(E_1 \cap E_2) = 1 - [Pr(E_1')Pr(E_2' \mid E_1') + Pr(E_1')Pr(E_2 \mid E_1') + Pr(E_1)Pr(E_2' \mid E_1)] \quad (7.6)$$

Equation 7.6 illustrates the progressive decrease in accuracy as a result of overlay analysis compared to the accuracy of the original data layers, due to the accumulation of errors. As the number of data layers increases the possible combination of errors increases rapidly, with the actual accuracy dependent on the spatial coincidence of erroneous pixels. The highest resultant accuracy is achieved when errors in all layers coincide, and the lowest when errors in each layer occur at mutually exclusive, disjoint locations (Newcomer and Szajgin, 1984). The upper bound for the accuracy of the composite map derived from the AND operation for n layers can be no greater than the least accurate layer (Veregin, 1989):

$$Pr(E_1 \cap E_2 \cap .....E_n)^{max} = min\,[Pr(E_i)] \qquad for\ i = 1,2.....n \quad (7.7)$$

The lower bound for the accuracy is generally stated as (Newcomer and Szajgin, 1984):

$$Pr(E_1 \cap E_2 \cap .....E_n)^{min} = \left[1 - \sum_{i=1}^{n} Pr(E_i')\right] \quad (7.8)$$

The importance of spatial coincidence of errors is demonstrated in Table 7.6 where the Overall Change Accuracy determined in Chapter 6 for each multitemporal comparison falls between the upper and lower bounds specified above.

If spatial independence of errors is assumed, the expected accuracy may be computed from the product of the probabilities of correct classification for each layer. The value was originally suggested by MacDougall (1975) to represent the maximum mapping accuracy, but was subsequently clarified by Newcomer and Szajgin (1984), Chrisman (1987) and Veregin (1995).

The expected accuracy values are computed as follows (Veregin, 1995) and are listed in Table 7.6 :

$$Pr(E_1 \cap E_2 \cap .....E_n) = \prod_{i=1}^{n} Pr(E_i) \tag{7.9}$$

Class. Level	Change Dates	Overall Classification Accuracy	$Pr(E_1 \cap E_2)^{max}$ / $Pr(E_1 \cap E_2)^{min}$	Expected Accuracy (Veregin, 1995)	Overall Change Accuracy
I	1972 1986	0.79 0.77	0.77 0.56	0.61	0.64
I	1986 1991	0.77 0.82	0.77 0.59	0.63	0.70
II	1972 1986	0.62 0.69	0.62 0.30	0.42	0.33
II	1986 1991	0.69 0.67	0.67 0.36	0.46	0.51
I/II	1972 1986	0.79 0.69	0.69 0.48	0.54	0.56

Table 7.6 *Comparison of predicted and actual change assessment results using global error modelling*

The Overall Change Accuracy for all change comparisons, except for Level II between 1972 and 1986, are above the expected value computed from Equation 7.9. This indicates the presence of spatial correlation of errors because the accuracy is greater than expected for spatial independence of classification errors between data layers. These results show that spatial dependency of errors affects the outcome of the change assessment and, considered in conjunction with the analysis in Section 7.3, indicates that boundary error effects make a major contribution.

The 1972 to 1986 Level II Overall Change Accuracy is close to the lower bound of accuracy for these data, however the results were derived from a comparison of relatively coarse Landsat MSS data and much finer spatial resolution Landsat TM data. This combination resulted in a poor Overall Change Accuracy. The high rate of change assessment errors makes it difficult to extract any meaningful conclusions from this change comparison, and the overall error rate was so high that any spatial dependence in the location of errors between the two classification dates is not apparent.

Global error models have been applied to these data, however the aspatial nature of the analyses limits the specific conclusions that can be derived. While the global error model may be used to predict the Overall Change Accuracy *a priori*, it cannot provide any more specific error data, and it is only through *a posteriori* analysis that an assessment of the existence of spatial dependence can be made. Spatial/thematic error models implemented through Monte Carlo simulation are required to develop a more comprehensive understanding of error distribution.

7.6 Summary

Understanding of error patterns within the classification and change assessment process provides an important diagnostic capability for evaluation of data quality. Error modelling provides the logical framework for analysis of error propagation and determination of aggregated error loads within the analysis system. These important metrics are useful in identifying critical error sources and facilitate the design of data collection and analysis approaches to minimise error accumulation.

Previous studies were examined to establish traditional benchmarks used to assess spatial and thematic accuracy in classified remotely sensed data. Unlike topographic maps, few general standards have been developed specifically for thematic products, and most rely upon comparison with relevant topographic mapping standards. Evaluation of comparable studies of land cover mapping using remotely sensed data revealed that spatial accuracy in the order of $\pm$0.5 - 1.0 pixel is achievable, and compares favourably with the results obtained within this research. Thematic accuracies vary considerably, however for those related to large area coherent mapping programs, thematic accuracies above 80 percent are desirable. Research-based projects rarely achieve this level of accuracy, and often take an evaluative approach to algorithm development and attainment of a set standard is not the primary objective.

Valid comparisons of classification and change detection accuracies between studies is problematic due to the variation within classification schemes and the objects being identified. While a degree of standardisation is provided through the hierarchical classification scheme developed by Anderson *et al* (1976), the contrast between classes within close proximity has a significant effect on the achievable accuracy in the final product. For example, classification of *sand* versus *water* is likely to result in greater

classification accuracy than a classification that compares *deciduous* to *evergreen forests*, even though all are Level II categories of land cover.

Evaluation of the sources of spatial error was undertaken to assess the relative contributions of components in the image interpretation and reference data collection process. Spatial error within the image interpretation process is quantified by the rectification RMSE when independent check points are used for evaluation. Errors were within the expected range of 0.5 - 1.0 pixel, with Landsat MSS relatively high at 0.9 pixel and all other data in the range 0.4 - 0.6 pixel.

The contribution of spatial error from the reference data was mainly determined by the photointerpretation error. For the three epochs evaluated in this study, aerial photographs at scales of 1:20 000 and 1:25 000 were used for reference data collection, and with conventional photointerpretation techniques, were limited in boundary identification to approximately 0.5 mm at photograph scale. For higher resolution data such as IRS1-D, photography at a scale of approximately 1:3 600 is required to provide reference data with a spatial accuracy suitable for reliable verification of information extracted from the satellite images.

When dealing with multiscale and multitemporal data, control of spatial and thematic errors continues to be an issue due to the wide range in data quality. The data in this research have been rectified to a high standard, however error propagation between corresponding reference data and satellite images results in spatial errors for change detection products in the order of one pixel for the coarsest resolution data in the change comparison. For example, the Landsat MSS and Landsat TM comparison produced a total RMSE for the change image of 68.6 m, and the SPOT HRV and IRS1-D comparison resulted in an RMSE of 14.4 m. This is equivalent to one Landsat MSS and one SPOT HRV pixel, or two Landsat TM and three IRS1-D pixels, respectively.

Control of these errors and recognition of their presence is essential for effective application and assessment of the resultant change information. Spatial errors may only be restricted to these levels through stringent control of error sources at all stages of the analysis. When dealing with remote sensing imaging systems elimination of errors is not feasible, but control of their magnitude must be a priority.

Spatial analysis in GIS involves evaluation of mathematical or logical data models, however for change assessment, analysis using a logical (Boolean) model is appropriate. The general law of propagation of errors is not applicable to error analysis of Boolean data models, and a specific error model must be derived for each Boolean operation. Spatial correlation of data between layers influences the propagation of errors and must be considered in the analysis.

Probabilistic and simulation approaches are available for analysis of error propagation in logical models. Probabilistic models consider spatial correlation, but only permit global error modelling without consideration of the spatial distribution of thematic errors or explicit consideration of the magnitude of spatial errors. Monte Carlo simulation provides for spatial and thematic perturbations to be introduced to the data, and enables statistical analysis of the empirical results. Error modelling was undertaken in this research using a probabilistic approach to global error modelling and the actual results achieved are consistent with the values predicted by the modelling technique.

Chapter 8

CONCLUSIONS AND RECOMMENDATIONS

This chapter reports on the conclusions developed as a result of the investigations performed in this research program. Experimental results and analyses are reviewed in the context of the overall research and recommendations for future work are proposed.

8.1 Conclusions

Research in this thesis dealt with the investigation of multiscale digital remote sensing data for the assessment of change in the rural-urban fringe. The principal objective was to evaluate data from a range of sensors and develop a change assessment and representation technique derived from supervised and unsupervised classification approaches.

Digital image data from Landsat MSS, Landsat TM, SPOT HRV and IRS1-D satellites were selected for analysis. Geometric evaluation of all datasets was undertaken, and standard classification techniques were used to extract land cover information from the Landsat MSS, Landsat TM and SPOT HRV data. Comparisons of supervised and unsupervised classifications at two levels of detail were evaluated to establish the optimum classification approach and level of interpretation for each dataset. A method of integrating multiscale change information was developed to provide continuity of change information between the various data sources, and investigations were conducted to establish the extent and distribution of errors in the interpreted data.

The study area for this research comprised parts of the City of Melville municipality, located approximately 11 km southwest of the city centre of Perth, Western Australia. The study area is approximately 1 850 ha of gently undulating coastal plain. In 1972 the area was located on the southern fringe of the Perth metropolitan area and has since been incorporated as portions of the dormitory suburbs of Winthrop and Murdoch as Perth gradually expands.

Raster-based IDRISI software was used for all image processing and geographic data analysis, and vector-based Microstation was required for import of reference data following photointerpretation of land cover boundaries.

The study comprised the following major components:

(i) Review the use of remote sensing for information extraction applied to temporal assessment, focusing on the spectral and spatial resolutions of satellite sensors and how these affect image interpretation. Evaluate the classification accuracy, spatial and thematic error propagation and change reporting;

(ii) Compile relevant Landsat MSS, Landsat TM, SPOT HRV and IRS1-D satellite data for the study area in formats suitable for analysis. Interpret aerial photographs for each satellite image acquisition date for use as ground reference data;

(iii) Geocode all satellite data to the AMG using digital planimetric data provided by DOLA and MRWA, and evaluate rectification accuracy using independent control points for verification;

(iv) Define land cover classes suitable for multiscale assessment based upon a standard classification system and considering the spectral and spatial resolution of the satellite data. Select and refine training sites for image classification;

(v) Assess the accuracy of each classification of the multiscale satellite imagery and evaluate the allocation of classes considering variations in spatial resolution;

(vi) Determine the land cover changes between epochs and evaluate change representation for multiscale satellite data by analysing the change matrices and their accuracy parameters;

(vii) Summarise the changes observed between sensing epochs and develop methods to compare classifications between datasets at multiple levels of abstraction and information content;

(viii) Evaluate the effect of combined spatial and thematic errors on change representation, and investigate appropriate methods for error modelling and assessment.

8.1.1 Multiscale Assessment

Integration of multiscale data relies upon harmonisation of the spatial and spectral properties of the images. Spatial integration is based upon image rectification and requires spatial coincidence of the separate datasets to be established in order for comparisons to be made between images collected at different times. Spectral integration involves the extraction of the required class information from each image for the assessment of change.

Rectification of images ranging in resolution from 79 x 56 m to 5.8 x 5.8 m was required. Potential variations in sensor altitude, attitude, stability and calibration affect the quality of rectification results, however first order polynomial modelling was appropriate to correct for these influences in all data except the high resolution IRS1-D

data, which required a third order polynomial rectification model. *A priori* estimates of GCP RMSE were in the order of 0.3 - 0.4 pixel at the original resolution of each dataset, and *a posteriori* estimates for independent check points varied between 0.4 and 0.9 pixel. These results meet the established standards of 0.5 - 1.0 pixel and provide a sound geometric basis for change assessment.

Multiscale comparison required all images to be resampled to the same pixel dimension. This was implemented using nearest neighbour resampling to the smallest pixel dimension of the multispectral data, which was equivalent to the 20 m SPOT HRV data. Image displacement during resampling was minimised for all datasets by using the 20 m resampling interval. Resampling to the smallest pixel size maintained the overall quality of the data, and minimised the spatial degradation of the coarse resolution Landsat MSS data as it was transformed from the off-meridian imaging geometry to the map grid oriented resampling grid.

The IRS1-D data were not used in the change assessment process and were resampled using nearest neighbour techniques to 5 m pixels. A third order polynomial provided an *a posteriori* RMSE of 0.55 pixel, indicating that standard polynomial rectification approaches are also applicable to this high resolution sensor.

At resolutions ranging from 20 m to 79 m the level of detail evident in the images varied considerably. Resampling of these multiscale images equivalent to the highest resolution spatial data accommodated these variations and enabled extraction of information relevant to the scale of data acquisition for each satellite.

8.1.2 Multitemporal Assessment

Interpretation of multiscale satellite data was investigated to take advantage of a range of current remote sensing satellites capable of producing continuous data streams, and to develop an approach for integration of historical satellite data for change detection. Multitemporal assessment normally relies on the comparative analysis of corresponding images acquired by the same sensor. Maintenance of standard data acquisition parameters and sensor characteristics are generally the main argument for this restriction. This study specifically investigated multitemporal analysis of data from different sensors to develop techniques to overcome these limitations.

Spectral and radiometric variations limit the scope for direct integration and comparison of multitemporal data from different sensors. Most research utilises image algebra and spectral transformation techniques for change detection, without considering the opportunities provided by multiscale satellite data. This study investigated the implications of utilising multiscale data and determined that postclassification comparison is the most appropriate approach to employ with these data. The multisensor approach results in the incorporation of data with different radiometric response patterns and calibration standards, a disparate range of spectral bands and significant variations in spatial resolution. Postclassification comparison is the only standard image processing approach able to accommodate these variations.

8.1.3 Reference Data Collection

Collection of reference data is an important component of remote sensing analysis as it forms the basis on which decisions regarding processing approaches and interpretations are based. Relevant parameters for reference data collection include the temporal relationship with the primary remote sensing data, as well as the comparative level of thematic detail and spatial accuracy.

The primary remote sensing data were collected over a 19 year period from 1972 to 1991. Strong reliance was placed on utilisation of existing sources of reference data. The reference data were mainly panchromatic and colour medium scale aerial photographs acquired primarily for topographic mapping purposes. These photographs have been routinely acquired annually at approximately the same date each year. The satellite data were selected subject to availability from existing archives, and discrepancies between the time of reference data collection and satellite data acquisition of up to six months occurred.

Aerial photographs were acquired from DOLA immediately before and after each sensing epoch, and compared to establish the nature and magnitude of changes during the period that the satellite data were acquired. While it was impossible to provide an interpretation which was correct at the time of satellite overpass, the potential errors in the reference data were recognised. Additional errors in the interpretation of the reference data were possible, but because it involved historical land cover data no independent information was available to verify the accuracy of the reference data. The

extent of error in historical reference data is often indeterminate and may have a significant (but unknown) influence on the evaluation of the thematic accuracy of the interpretation.

Reference data for GCP extraction were obtained from various sources including 1:100 000 topographic maps (Landsat MSS), digital planimetric maps equivalent to 1:25 000 scale (Landsat TM and SPOT HRV) and digital road centreline data equivalent to approximately 1:10 000 scale planimetric maps (IRS1-D). The use of 1:100 000 topographic maps for extraction of GCPs for Landsat MSS data potentially degraded the quality of the image rectification, but for all other images the accuracy of reference data was not a limiting factor. The lower precision of GCPs extracted for the Landsat MSS data may have adversely affected the rectification of these data as reported in Section 8.1.1.

Derivation of error-free reference data over large areas, at high sampling densities and from historical information is difficult to achieve, however techniques adopted in this study were designed to minimise the extent of reference data errors. Where thematic errors in reference data can be quantified, they should be included within the error modelling process.

8.1.4 Image Classification

Supervised and unsupervised classification procedures were evaluated in this research for the purpose of providing an optimum classification of the multitemporal satellite data. Both techniques were applied as a means of investigating a limited range of classification algorithms commonly used in practice, whereas the remote sensor data were derived from a wide range of sensors. Both algorithms provided similar results with Overall Classification Accuracies in the range of 60 – 80 percent.

Investigation of target characteristics indicated that the selected classes showed least spectral separation in the Landsat MSS data, with Landsat TM and SPOT HRV showing progressively greater spectral separation. Supervised classification relied on analyst guidance for identification of training pixels, while unsupervised classification used the ISODATA self organising clustering algorithm to separate the data into discrete

clusters. Subsequent class identification enabled allocation of clusters to each land cover class.

As a result of unsupervised classification there were no clusters allocated to some land cover classes. This indicated that for some classes identified in the reference data and during supervised training of the classifier, the selected classes were not spectrally separable using unsupervised classification. Close attention needs to be paid to the definition of land cover classes so that they have relevant and identifiable spectral characteristics. The relationship between spectral characteristics and information content is a fundamental tenet of image classification systems.

In a similar manner, unsupervised classification of the data revealed that residential and urban classes were the most spectrally diverse land covers requiring the most clusters for representation at both Level I and Level II. In each case the classification of the classes was also consistently high. Careful attention needs to be paid to the derivation of training sites for supervised classification of these complex classes and an adequate level of clustering must be used in unsupervised classification to provide complete description of each class.

8.1.5 Classification Accuracy

A stratified random sampling approach was used to assess the correspondence between the classified satellite image and the land cover map produced from the reference data. The sampling rate was determined by providing for a minimum number of samples for each area and enabled a statistically valid assessment of classification accuracy to be made.

The objective was to compile a reference map at a level of detail relevant to the satellite data under investigation and the classification level employed. This required consideration of the spatial resolution of the sensor when determining the minimum mapping unit for the reference map. Visual comparison of the Landsat MSS classification and the corresponding reference data map indicated the level of spatial detail in the reference data was too high because features such as unmade roads included in the reference data were too small to be resolved in the satellite data. Further spatial generalisation of the reference data may be required to resolve this issue.

Pixel by pixel assessments of classification accuracy were undertaken in preference to generalised area-based estimates. Analytical techniques such as the Kappa Coefficient were used in conjunction with descriptive approaches such as the Overall Classification Accuracy for assessment of classification accuracy. Generally Kappa indicated *moderate* to *substantial* levels of agreement were achieved between the classified satellite data and the reference data. Kappa provided a more comprehensive analysis of classification accuracy and enabled the statistical comparison of results of classification approaches to identify an optimum approach.

Classifications were undertaken on all datasets at both Level I and Level II using supervised and unsupervised classification algorithms. Comparisons between algorithms at Level I indicated that no significant difference existed between the supervised and unsupervised approaches for any of the datasets. At Level II no significant difference existed for classification of Landsat MSS data, however differences between the results achieved from the two algorithms for the Landsat TM and SPOT HRV datasets were due largely to the inability of the unsupervised classifier to allocate any clusters to some land cover classes.

Consideration of classification accuracies between satellite datasets and also between levels of classification for each dataset indicated that differences were less than expected. Evaluation of the pattern of classification errors through spatial autocorrelation and analysis of the compactness of the land cover classes indicated that spatial complexity and associated boundary identification were major factors affecting classification accuracy in the Landsat TM and SPOT HRV images. The substantial increase in spatial complexity of the land cover patterns mitigated against any improvement in classification accuracy expected to result from the finer spatial resolution of these sensors.

Analysis of classification uncertainty for the supervised classification of the Landsat TM data established that up to one third of the pixels were classified with an uncertainty of 50 percent or more, indicating low confidence in many class allocations and high potential for misclassification of these pixels. Inability of the unsupervised classification algorithm to recognise some land cover classes reinforced this finding.

8.1.6 Change Assessment and Change Representation

The accuracy of change assessment was performed in the same manner as the accuracy assessment of individually classified images. Stratified random sampling of the change image and reference change map were compared pixel by pixel and change statistics computed. The change error matrix increased in size according to the product of the number of classes in the component images, with the result that evaluation of the Kappa Coefficient became a significant computational overhead. In the case of the change accuracy assessment of Level II classifications, the Overall Change Accuracy was used in preference to the Kappa Coefficient due to the large dimension of the change matrix. The Overall Change Accuracy provided an inferior, but adequate assessment tool in this situation.

Change representation involves the measurement and communication of changes in land cover classes between sensing epochs. The methods of change representation that were investigated included change summaries involving calculation of total areas of land cover change, change maps that provided a visual assessment of change and change matrices that demonstrated change areas on a class by class basis. No single method was universally appropriate, however the combination of change maps and change matrices provided the most comprehensive statement of temporal change.

Development of an approach to integrate multiscale datasets was an important consideration in the change analysis process. Understanding the importance of maintaining a hierarchical classification scheme resulted in the seamless comparison of multidate Level I and Level II classifications. Integration of the classifications at two levels enabled combined change analysis involving data from both the Landsat MSS and Landsat TM sensors, while maintaining the quality and integrity of the individual datasets. Combined analysis provided an effective means of tracing land cover evolution throughout the period of data collection, and enabled the effective comparison of multiple datasets observed at spatial resolutions ranging from 20 m through to 79 m. Similar approaches could be used to compare classification results obtained from any pair of sensors regardless of spatial resolution, provided a hierarchical classification scheme is utilised.

8.1.7 Spatial and Thematic Error Propagation

Analysis of error modelling and propagation in remotely sensed data involves a complex consideration of spatial and thematic error components. Both of these components may be contributed by the primary remote sensing data or by the reference data. During the change assessment process the propagation of spatial errors within the component land cover images was analysed using the general law of propagation of errors. For comparison of Landsat MSS and Landsat TM data a total spatial error of 68.6 m was estimated, while for the SPOT HRV and IRS1-D comparison the error was estimated to be 14.4 m. At both extremes of the analysis these values represent significant quantities and, through displacement of the images, can cause substantial errors in change estimates.

Reference data errors should not be the limiting factor in change analysis, however the magnitude of photointerpretation errors contributed by the reference data dominated the error budget. Control of these errors and recognition of their presence is essential for effective application and assessment of the resultant change information. Spatial errors may only be restricted to appropriate levels through stringent control of error sources at each stage of the analysis. When dealing with remote sensing imaging systems elimination of errors is not feasible, but knowledge and control of their magnitude must be a priority.

Land cover change assessment was performed using the AND operator, which is a logical (Boolean) operation. The general law of propagation of errors is not applicable to analysis of thematic errors in a Boolean model, and a specific error model must be derived for each Boolean operation. A probabilistic error model was utilised for error propagation analysis and indicated that errors were spatially dependent between the classified layers. This conclusion is in line with expectations and supported by the spatial distribution of errors observed along the class boundaries.

The global approach to error propagation using probabilistic methods was of limited evaluative and diagnostic application due to the absence of class-specific spatial and thematic error assessment. The technique relies upon image-wide classification accuracies such as the Kappa Coefficient, which do not differentiate between the error

components of location or class label. More specific approaches such as Monte Carlo simulation should be investigated.

This research has provided a means for extracting land cover information for the rural-urban fringe using a range of medium resolution multispectral data and integrating the results, despite differences in the level of detail available from each dataset. This approach will be useful as an increasing number of remote sensing satellites are launched and the data are utilised for a wide variety of applications. The analysis of spatial and thematic errors provides a detailed understanding of quality issues associated with spatial data management and will be useful in the assessment of future mapping and change assessment products.

8.2 Recommendations

The research program was designed to thoroughly investigate the objectives of the study. The results from the research concluded that this investigation identified a coherent method of producing change detection information from multiscale satellite data in a manner that permits meaningful comparisons to be made between classifications derived from different remote sensing systems. During the implementation of this research several limitations were identified.

8.2.1 Data Sources

Satellite remote sensing data commonly used for land cover interpretation and assessment range from low resolution (1.1 km) NOAA Advanced Very High Resolution Radiometer (AVHRR) through to medium resolution (20 m) SPOT HRV data. Recent developments in satellite sensing systems have resulted in the commercial availability of remotely sensed data with spatial resolutions in the order of one to three metres. A wide range of data sources is therefore available for interpretation and monitoring, each with its own spectral and spatial qualities.

Data from these new high resolution satellites should be further investigated to establish the information content and relevance to assessment of land cover change in the rural-urban fringe. These satellites are capable of expanding the range of spectral bands, level of detail (scale) and temporal availability of information. Constant evaluation of new data sources will ensure the continuity of monitoring data for earth surface features.

The geometric qualities of these data need to be investigated and assessments made of appropriate approaches to image rectification. Lower orbit paths are subject to increased disturbance, and smaller pixels with greater sensitivity to geometric distortion may require more sophisticated techniques beyond the polynomial modelling approach utilised in this study. Such techniques have been developed for use with aircraft multispectral scanner data and may be useful for minimising spatial errors in high resolution satellite data.

The time of acquisition of satellite data was not considered as a serious issue during this study as independent classification of each sensing epoch was undertaken. Improved classification results may be achieved if the phenological cycles of land covers such as *Grassland,* are considered during data selection. This may enable superior separation of the *Grassland, Recreation* and *Woodland* classes.

The ability to accurately interpret satellite data is dependent on timely and relevant reference data. Reference data must be acquired close to the sensing date such that changes from the date of acquisition are minimised, and at a scale that provides sufficient information to assist interpretation and minimise the propagation of errors, but not so detailed as to distract attention from the primary interpretation purpose. The acquisition of timely and accurate data to specifically provide relevant reference data should be given serious consideration.

8.2.2 Classification Methods

Standard supervised and unsupervised classification algorithms were applied to the interpretation of the remotely sensed data. The classification results achieved in this research were identified as providing *moderate* to *substantial* levels of agreement with the reference data. These represent a level of classification accuracy that is adequate for change assessment, but improved results are highly desirable. Further research should be performed to evaluate the wide range of algorithms available for interpretation of land cover change information.

This research indicated moderate levels of uncertainty in the allocation of pixels to classes when the maximum likelihood decision rule was implemented for supervised classification. Many pixels exhibited at least moderate levels of multiple class

membership due to confusion resulting from the definition of the spectral classes, or because of the spatial complexity of the targets. More sophisticated classification approaches that include fuzzy membership principles or incorporate the evaluation of ancillary data to discriminate targets should be investigated.

Unsupervised classification indicated that some classes selected for interpretation were not spectrally separable in the satellite data, even though they were identifiable in false colour composite images and formed independent classes when evaluated in the reference data. Detailed examination of the spectral separability of land cover classes needs to be undertaken prior to final design of any classification scheme.

8.2.3 Error Modelling

Modelling of errors within remote sensing and GIS analyses is essential for the evaluation of the quality of interpretation products. Consideration of the spatial and thematic components of the analysis is required for a complete understanding of error propagation within the data. This research evaluated the components contributing to propagation of spatial errors using deterministic techniques, and applied a global error (probabalistic) propagation approach to the evaluation of thematic errors produced as a result of logical modelling of the data.

An integrated spatial/thematic approach to error modelling is required that provides a coherent analysis of the interrelationships between these two distinct sources of error. With multiscale data the variable impact of correlated errors derived from image rectification, boundary location and thematic classification, that are convolved through both the interpretation of the primary remote sensing data and the reference information, results in a complex combination of error components. It is clear that standard deterministic and probabilistic approaches, applied separately to spatial and thematic components, are inadequate for the task, and further research is required to evaluate the effectiveness of techniques such as the analysis of errors using fuzzy modelling and Monte Carlo simulation. The models used to introduce the perturbations into the simulation process also require particular scrutiny to establish their relevance to the data.

Investigations should also be made to determine the feasibility of developing spatial/thematic error modelling as an integral part of GIS software systems and the potential for real time error modelling. This facility will assume greater importance as data from disparate sources and with varying levels of reliability are compared as part of change assessment and other modelling processes.

8.2.4 Interpretation Standards

Planimetric standards for compilation of topographic maps are well established and clearly defined using straightforward statements of horizontal and vertical accuracy. Standards for production of thematic products are less well developed. Thematic standards require a thorough statement of planimetric accuracy, as well as the level of thematic interpretation and its corresponding nearness to the *truth* as derived from the reference data.

Research in this study highlighted the lack of definitive statements of quality attached to thematic interpretations. This was indicated by the diverse range of interpretation results derived from similar datasets at corresponding classification levels, but due to the inherent low spectral separability and spatial complexity of the land cover classes present in the area, comparable levels of interpretation quality were not achievable. For example, a Level II classification between *Pine plantation* and *Urban* is much easier to achieve than a Level II classification between *Pine plantation* and *Woodland*. Further research is required to define a thematic classification assessment standard that can provide a meaningful quality measure relating the level of classification to the spatial and spectral complexity of the target.

Evaluation standards do not need to be stringent, however benchmarking is required to establish the credibility of satellite-derived thematic data and to provide a basis for comparison between analyses, especially where regional and global resource surveys are involved.

This study has provided an insight into the potential for the combination of data from various remote sensors in an evolving climate of increased choice of satellite systems, and increasing uncertainty of a long term data stream from any particular sensor. The foundation has been established for integration of multiple datasets for land cover

mapping and change assessment, and the potential for continuous long term monitoring using established and future data acquisition systems has been shown.

REFERENCES

ACRES (1999) *ACRES UPDATE*, Australian Centre for Remote Sensing, No. 19, August, 16 pp.

Anderson, J.R., Hardy, E.E, Roach, J.T and Witmer, R.E. (1976) *A Land Use and Land Cover Classification Scheme for Use With Remote Sensor Data*, Geological Survey Professional Paper 964, US Geological Survey, Washington, DC, 28 pp.

Arai, K. (1992) A Supervised Thematic Mapper Classification With a Purification of Training Samples, *International Journal of Remote Sensing*, Vol. 13, No. 11, pp. 2039-2049.

Aronoff, S. (1982) The Map Accuracy Report: A User's View, *Photogrammetric Engineering and Remote Sensing*, Vol. 48, No. 8, pp. 1299-1307.

Aronoff, S. (1985) The Minimum Accuracy Value as an Index of Classification Accuracy, *Photogrammetric Engineering and Remote Sensing*, Vol. 51, No. 1, pp. 99-111.

Aspinall, R.J. and Hill, M.J. (1997) Land Cover Change: A Method for Assessing the Reliability of Land Cover Changes Measured from Remotely-Sensed Data, *1997 International Geoscience and Remote Sensing Symposium*, Singapore, August, Vol. 1, pp. 269-271.

Atkinson, P.M. (1997) Selecting the Spatial Resolution of Airborne MSS Imagery for Small-Scale Agricultural Mapping, *International Journal of Remote Sensing*, Vol. 18, No. 9, pp. 1903-1917.

Atkinson, P.M. and Curran, P.J. (1997) Choosing an Appropriate Spatial Resolution for Remote Sensing Investigations, *Photogrammetric Engineering and Remote Sensing*, Vol. 63, No. 12, pp. 1345-1351.

Bailey, T.C. and Gatrell, A.C. (1995) *Interactive Spatial Data Analysis*, Longman Scientific and Technical, Harlow, UK, 413 pp.

Barnsley, M.J. and Barr, S.L. (1996) Inferring Urban Land Use from Satellite Sensor Images Using Kernel-Based Spatial Reclassification, *Photogrammetric Engineering and Remote Sensing*, Vol. 62, No. 8, pp. 949-958.

Basham May, A.M., Pinder III, J.E. and Kroh, G.C. (1997) A Comparison of Landsat Thematic Mapper and SPOT Multi-Spectral Imagery for the Classification of Shrub and Meadow Vegetation in Northern California, USA, *International Journal of Remote Sensing*, Vol. 18, No. 18, pp. 3719-3728.

Beaubien, J. (1994) Landsat TM Satellite Images of Forests: From Enhancement to Classification, *Canadian Journal of Remote Sensing*, Vol. 20, No. 1, pp. 17-26.

Bishop, Y.M.M., Fienberg, S.E. and Holland, P.W. (1975) *Discrete Mulitvariate Analysis*, MIT Press, Cambridge, MA, 557 pp.

Blakemore, M. (1984) Generalisation and Error in Spatial Databases, *Cartographica*, Vol. 21, pp. 131-139.

Bolstad, P.V., Gessler, P. and Lillesand, T.M. (1990) Positional Accuracy in Manually Digitised Map Data, *International Journal of Geographical Information Systems*, Vol. 4, No. 4, pp. 399-412.

Buiten, H.J. and van Putten, B. (1997) Quality Assessment of Remote Sensing Image Registration – Analysis and Testing of Control Point Residuals, *Journal of Photogrammetry and Remote Sensing*, Vol. 52, pp. 57-73.

Burrough, P.A. (1986) *Principles of Geographical Information Systems for Land Resources Assessment*, Oxford University Press, Oxford, UK, 194 pp.

Byrne, G.F., Crapper, P.F. and Mayo, K.K. (1980) Monitoring Land-Cover Change by Principal Components Analysis of Multitemporal Landsat Data, *Remote Sensing of Environment*, Vol. 10, pp. 175-184.

Campbell, J.B. (1981) Spatial Correlation Effects Upon Accuracy of Supervised Classification of Land Cover, *Photogrammetric Engineering and Remote Sensing*, Vol. 47, No. 3, pp. 355-363.

Chrisman, N.R. (1987) The Accuracy of Map Overlays: A Reassessment, *Landscape and Urban Planning*, Vol. 14, pp. 427-439.

Chrisman, N.R. (1989) Modelling Error in Overlaid Categorical Maps, In *Accuracy of Spatial Databases*, Goodchild, M. and Gopal, S. (eds), Taylor and Francis Ltd, London, UK, Ch. 2, pp. 21-34.

Chrisman, N.R. (1991) The Error Component in Spatial Data, In *Geographical Information Systems: Principles and Applications*, Maguire, D.J., Goodchild, M.F. and Rhind, D.W. (eds), Longman Scientific and Technical, Halow, UK, pp. 165-175.

Cliff, A.D. and Ord, J.K. (1981) *Spatial Processes: Models and Applications*, Pion Limited, London, UK, 266 pp.

Cohen, J. (1960) A Coefficient of Agreement for Nominal Scales, *Educational and Psychological Measurement*, Vol. XX, No. 1, pp. 37-46.

Collins, J.B. and Woodcock, C.E. (1996a) An Assessment of Several Linear Change Detection Techniques for Mapping Forest Mortality Using Multitemporal Landsat TM Data, *Remote Sensing of Environment*, Vol. 56, pp. 66-77.

Collins, J.B. and Woodcock, C.E. (1996b) Explicit Consideration of Multiple Landscape Scales While Selecting Spatial Resolutions, *Proceedings of Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, Fort Collins, Colorado, May, pp. 121-128.

Congalton, R.G. (1988a) A Comparison of Sampling Schemes Used in Generating Error Matrices for Assessing the Accuracy of Maps Generated from Remotely Sensed Data, *Photogrammetric Engineering and Remote Sensing*, Vol. 54, No. 5, pp. 593-600.

Congalton, R.G. (1988b) Using Spatial Autocorrelation Analysis to Explore the Errors in Maps Generated from Remotely Sensed Data, *Photogrammetric Engineering and Remote Sensing*, Vol. 54, No. 5, pp. 587-592.

Congalton, R.G. (1991) A Review of Assessing the Accuracy of Classification of Remotely Sensed Data, *Remote Sensing of Environment*, Vol. 37, pp. 35-46.

Congalton, R.G., Oderwald, R.G. and Mead, R.A. (1983) Assessing Landsat Classification Accuracy Using Discrete Multivariate Analysis Statistical Techniques, *Photogrammetric Engineering and Remote Sensing*, Vol. 49, No. 12, pp. 1671-1678.

Cook, A.E. and Pinder III, J.E. (1996) Relative Accuracy of Rectifications Using Coordinates Determined from Maps and the Global Positioning System, *Photogrammetric Engineering and Remote Sensing*, Vol. 62, No. 1, pp. 73-77.

Dai, X. and Khorram, S. (1997) Development of a Feature-Based Approach to Automated Image Registration for Multitemporal and Multisensor Remotely Sensed Imagery, *1997 International Geoscience and Remote Sensing Symposium*, Singapore, August, Vol. 1, pp. 243-245.

Davis, F.W. and Simonett, D.S. (1991) GIS and Remote Sensing, In *Geographical Information Systems: Principles and Applications*, Volume 1, Maguire, D.J., Goodchild, M.F. and Rhind, D.W. (eds), Longman Scientific and Technical, Harlow, UK, Ch. 14, pp. 191-213.

Davis, F.W *et al* (1991) Environmental Analysis Using Integrated GIS and Remotely Sensed Data: Some Research Needs and Priorities, *Photogrammetric Engineering and Remote Sensing*, Vol. 57, No. 6, pp. 689-697.

Dean, D.J., Wilson, K.R. and Flather, C.H. (1997) Spatial Error Analysis of Species Richness for a Gap Analysis Map, *Photogrammetric Engineering and Remote Sensing*, Vol. 63, No. 10, pp. 1211-1217.

Deer, P.J. (1998) A Quantitative Comparison of Traditional Classification Approaches, *Proceedings of 9th Australasian Remote Sensing and Photogrammetry Conference*, Sydney, July, CD-ROM.

Dikshit, O. and Roy, D.P. (1996) An Empirical Evaluation of Image Resampling Effects Upon the Spectral and Textural Supervised Classification of a High Spatial Resolution Multispectral Image, *Photogrammetric Engineering and Remote Sensing*, Vol. 62, No. 9, pp. 1085-1092.

Djamdji, J-P., Bijaoui, A. and Maniere, R. (1993) Geometrical Registration of Images: the Multiresolution Approach, *Photogrammetric Engineering and Remote Sensing*, Vol. 59, No. 5, pp. 645-653.

Drummond, J. (1987) A Framework for Handling Error in Geographic Data Manipulation, *ITC Journal*, No.1, pp. 73-82.

Duggin, M.J. and Robinove, C.J. (1990) Assumptions Implicit in Remote Sensing Data Acquisition and Analysis, *International Journal of Remote Sensing*, Vol. 11, No. 10, pp. 1669-1694.

Dunn, R., Harrison, A.R. and White, J.C. (1990) Positional Accuracy and Measurement Error in Digital Databases of Land Use, an Empirical Study, *International Journal of Geographical Information Science*, Vol. 4, No. 4, pp. 385-398.

Eastman, J.R. (1997) *IDRISI for Windows User's Guide*, Revision 2, Clark Labs for Cartographic Technology and Geographic Analysis, Clark University, Worcester, MA.

Eckhardt, D.W., Verdin, J.P. and Lyford, G.R. (1990) Automated Update of an Irrigated Lands GIS Using SPOT HRV Imagery, *Photogrammetric Engineering and Remote Sensing*, Vol. 56, No. 11, pp. 1515-1522.

Edwards, G. and Lowell, K.E. (1996) Modelling Uncertainty in Photointerpreted Boundaries, *Photogrammetric Engineering and Remote Sensing*, Vol. 62, No. 4, 377-391.

Ehlers, M. (1997) Rectification and Registration, In *Integration of Geographic Information Systems and Remote Sensing*, Star, J.L., Estes, J.E. and McGwire, K.C. (eds), Ch. 2, Cambridge University Press, Cambridge, UK, pp. 13-36.

El-Manadili, Y and Novak, K. (1996) Precision Rectification of SPOT Imagery Using the Direct Linear Transformation Model, *Photogrammetric Engineering and Remote Sensing*, Vol. 62, No. 1, pp. 67-72.

Fitzgerald, R.W. and Lees, B.G. (1994) Assessing the Classification Accuracy of Multisource Remote Sensing Data, *Remote Sensing of Environment*, Vol. 47, pp. 362-368.

Fitzpatrick-Lins, K. (1978) Accuracy of Selected Land Use and Land Cover Maps in the Greater Atlanta Region, Georgia, *Journal of Research of the US Geological Survey*, Vol. 6, No. 2, pp. 169-173.

Fitzpatrick-Lins, K. (1981) Comparison of Sampling Procedures and Data Analysis for a Land-Use and Land Cover Map, *Photogrammetric Engineering and Remote Sensing*, Vol. 47, No. 3, pp. 343-351.

Fonseca, L.M.G. and Manjunath, B.S. (1996) Registration Techniques for Multisensor Remotely Sensed Imagery, *Photogrammetric Engineering and Remote Sensing*, Vol. 62, No. 9, pp. 1049-1056.

Foody, G.M. (1992) On the Compensation for Chance Agreement in Image Classification Accuracy Assessment, *Photogrammetric Engineering and Remote Sensing*, Vol. 58, No. 10, pp. 1459-1460.

Foresam, T.W. and Millette, T.L. (1997) Integration of Remote Sensing and GIS Technologies for Planning, In *Integration of Geographic Information Systems and Remote Sensing*, Star, J.L., Estes, J.E. and McGwire, K.C. (eds), Ch. 7, Cambridge University Press, Cambridge, UK, pp. 134-157.

Forghani, A. (1994) A New Technique for Map Revision and Change Detection Using Merged Landsat TM and SPOT Data Sets in an Urban Environment, *Asian-Pacific Remote Sensing Journal*, Vol. 7, No. 1, pp. 119-127.

Forier, F. and Canters, F. (1996) A User-Friendly Tool for Error Modelling and Error Propagation in a GIS Environment, *Proceedings of Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, Fort Collins, Colorado, May, pp. 225-234.

Forster, B.C. (1980a) Urban Control for Landsat Data, *Photogrammetric Engineering and Remote Sensing*, Vol. 46, No. 4, pp. 539-545.

Forster, B.C. (1980b) Urban Residential Land Cover Using Landsat Digital Data, *Photogrammetric Engineering and Remote Sensing*, Vol. 46, No. 4, pp. 547-558.

Forster, B.C. (1985) An Examination of Some Problems and Solutions in Monitoring Urban Areas from Satellite Platforms, *International Journal of Remote Sensing*, Vol. 6, No. 1, pp. 139-151.

Forster, B.C. and Trinder, J.C. (1984) An Examination of the Effects of Resampling on Classification Accuracy, *Proceedings of Landsat 84-Third Australasian Conference on Remote Sensing*, Brisbane, Qld, pp. 106-115.

Fritz, L. (1996) The Era of Commercial Earth Observation Satellites, *Photogrammetric Engineering and Remote Sensing*, Vol. 62, No. 1, pp. 39-45

Fung, T. (1990) An Assessment of TM Imagery for Land-Cover Change Detection, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 28, No. 4, pp. 681-684.

Fung, T. and LeDrew, E. (1988) The Determination of Optimal Threshold Levels for Change Detection Using Various Accuracy Indices, *Photogrammetric Engineering and Remote Sensing*, Vol. 54, No. 10, pp. 1449-1454.

Gastellu-Etchegorry, J.P. (1990) An Assessment of SPOT XS and Landsat MSS Data for Digital Classification of Near-Urban Land Cover, *International Journal of Remote Sensing*, Vol. 11, No. 2, pp. 225-235.

Gates, D.G., Keegan, H.J., Schleter, J.C. and Weidner, V.R. (1965) Spectral Properties of Plants, *Applied Optics*, Vol. 4, pp. 11-20.

Ginevan, M.E. (1979) Testing Land-Use Map Accuracy: Another Look, *Photogrammetric Engineering and Remote Sensing*, Vol. 45, No. 10, pp. 1371-1377.

Gong, P. and Howarth, P.J. (1990) An Assessment of Some Factors Influencing Multispectral Land-Cover Classification, *Photogrammetric Engineering and Remote Sensing*, Vol. 56, No. 5, pp. 597-603.

Gong, P. and Howarth, P.J. (1992) Frequency-Based Contextual Classification and Gray-Level Vector Reduction for Land-Use Identification, *Photogrammetric Engineering and Remote Sensing*, Vol. 58, No. 4, pp. 423-437.

Gonzales, R.C. and Woods, R.E. (1993) *Digital Image Processing*, Addison-Wesley Publishing Company, Reading, MA, 716 pp.

Goodchild, M.F., Guoqing, S. and Shiren, Y. (1992) Development and Test of an Error Model for Categorical Data, *International Journal of Geographical Information Science*, Vol. 6, No. 2, pp. 77-104.

Green, K., Kempka, D. and Lackey, L. (1994) Using Remote Sensing to Detect and Monitor Land-Cover and Land-Use Change, *Photogrammetric Engineering and Remote Sensing*, Vol. 60, No. 3, pp. 331-337.

Haack, B., Bryant, N. and Adams, S. (1987) An Assessment of Landsat MSS and TM Data for Urban and Near-Urban Land-Cover Digital Classification, *Remote Sensing of Environment*, Vol. 21, pp. 201-213.

Haggett, P., Cliff, A.D. and Frey, A. (1977) *Locational Analysis in Human Geography*, Vol. 2, Second Edition, Edward Arnold, London, UK, 605 pp.

Haining, R. and Arbia, G. (1993) Error Propagation Through Map Operations, *Technometrics*, Vol. 35, No. 3, pp. 293-305.

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*, Vol. 35, pp. 11-27.

Haralick, R.M. and Fu, K. (1983) Pattern Recognition and Classification, In *Manual of Remote Sensing*, Colwell, R. (ed.), American Society of Photogrammetry, Falls Church, VA, Second Edition, Vol. 1, Ch. 18, pp. 793-805.

Heuvelink, G.B.M. (1998) *Error Propagation in Environmental Modelling with GIS*, Taylor and Francis Ltd, London, UK, 127 pp.

Heuvelink, G.B.M. and Burrough, P.A. (1993) Error Propagation in Cartographic Modelling Using Boolean Logic and Continuous Classification, *International Journal of Geographical Information Science*, Vol. 7, No. 3, pp. 231-246.

Heuvelink, G.B.M., Burrough, P.A. and Stein, A. (1989) Propagation of Errors in Spatial Modelling with GIS, *International Journal of Geographical Information Systems*, Vol. 3, No. 4, pp. 303-322.

Hill, J. and Aifadopoulou, D. (1990) Comparative Analysis of Landsat-5 TM and SPOT HRV-1 Data for Use in Multiple Sensor Approaches, *Remote Sensing of Environment*, Vol. 35, pp. 55-70.

Hoffer, R.M. and Johannsen, C.J. (1969) Ecological Potential in Spectral Signature Analysis, In *Remote Sensing in Ecology*, Johnson, P.L. (ed.), University of Georgia Press, Athens, GA, pp. 1-16.

Hord, R.M. and Brooner, W. (1976) Land-Use Map Accuracy Criteria, *Photogrammetric Engineering and Remote Sensing*, Vol. 42, No. 5, pp. 671-677.

Igbokwe, J.I. (1999) Geometrical Processing of Multi-Sensoral Multi-Temporal Satellite Images for Change Detection Studies, *International Journal of Remote Sensing*, Vol. 20, No. 6, pp. 1141-1148.

Irons, J.R., Markham, B.L., Nelson, R.F., Till, D.L., Williams, D.L., Latty, R.S. and Stauffer, M.L. (1985) The Effects of Spatial Resolution on the Classification of Thematic Mapper Data, *International Journal of Remote Sensing*, Vol. 6, No. 8, pp. 1385-1403.

Jackson, M.J., Carter, P., Smith, T.F. and Gardner, W.G. (1980) Urban Land Mapping from Remotely Sensed Data, *Photogrammetric Engineering and Remote Sensing*, Vol. 46, No. 8, pp. 1041-1050.

Jain, A.K. (1989) *Fundamentals of Digital Image Processing*, Prentice-Hall, Englewood Cliffs, NJ, 569pp.

Jakubauskas, M.E., Lulla, K.P. and Mausel, P.W. (1990) Assessment of Vegetation Change in a Fire-Altered Forest Landscape, *Photogrammetric Engineering and Remote Sensing*, Vol. 56, No. 3, pp. 371-377.

Janssen, L.L.F. and van der Wel, F.J.M. (1994) Accuracy Assessment of Satellite Derived Land-Cover Data: A Review, *Photogrammetric Engineering and Remote Sensing*, Vol. 60, No. 4, pp. 419-426.

Jeanjean, H., Achard, F. and Malingreau, J-P (1996) Large Scale Tropical Forest Change Monitoring Using Multiple Resolution Satellite Data: From Hot Spot Detection to Global Deforestation Assessment?, *Proceedings of Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, Fort Collins, Colorado, May, pp. 366-374.

Jensen, J.R. (1996) Introductory Image Processing: A Remote Sensing Perspective, Second Edition, Prentice-Hall, Upper Saddle River, NJ, 316 pp.

Jensen, J.R., Cowen, D.J, Althausen, J.D., Narumalani, S. and Weatherbee, O. (1993) An Evaluation of the CoastWatch Change Detection Protocol in South Carolina, *Photogrammetric Engineering and Remote Sensing*, Vol. 59, No. 6, pp. 1039-1046.

Jensen, J.R., Cowen, D.J., Halss, J., Narumalani, S., Schmidt, N.J., Davis, B.A. and Burgess, B. (1994) Improved Urban Infrastructure Mapping and Forecasting for BellSouth Using Remote Sensing and GIS Technology, *Photogrammetric Engineering and Remote Sensing*, Vol. 60, No. 3, pp. 339-346.

Jensen, J.R., Cowen, D., Narumalani, S. and Halls, J. (1997) Principles of Change Detection Using Remote Sensor Data, In *Integration of Geographic Information Systems and Remote Sensing*, Star, J.L., Estes, J.E. and McGwire, K.C. (eds), Ch. 3, Cambridge University Press, Cambridge, UK, pp. 37-54.

Jensen, J.R. and Toll, D.L. (1982) Detecting Residential Land-Use Development at the Urban Fringe, *Photogrammetric Engineering and Remote Sensing*, Vol. 48, No. 4, pp. 629-643.

Jupp, D.L.B., Strahler, A.H. and Woodcock, C.E. (1988) Autocorrelation and Regularization in Digital Images, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 26, No. 4, pp. 463-473.

Kalkhan, M.A., Reich, R.M. and Stohlgren, T.J. (1998) Assessing the Accuracy of Landsat Thematic Mapper Classification Using Double Sampling, *International Journal of Remote Sensing*, Vol. 19, No. 11, pp. 2049-2060.

Kardoulas, N.G., Bird, A.C. and Lawan, A.I. (1996) Geometric Correction of SPOT and Landsat Imagery: A Comparison of Map and GPS-Derived Control Points, *Photogrammetric Engineering and Remote Sensing*, Vol. 62, No. 10, pp. 1173-1177.

Kiiveri, H.T. (1997) Assessing, Representing and Transmitting Positional Uncertainty in Maps, *International Journal of Geographical Information Science*, Vol. 11, No. 1, pp. 33-52.

Knipling, E.B. (1970) Physical and Physiological Basis for the Reflectance of Near Infrared Radiation from Vegetation, *Remote Sensing of Environment*, Vol. 1, pp. 155-159.

Laba, M., Smith, S.D. and Degloria, S.D (1997) Landsat-Based Land Cover Mapping in the Lower Yuna River Watershed in the Dominican Republic, *International Journal of Remote Sensing*, Vol. 18, No. 14, pp. 3011-3025.

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 Control Points: A Simulation, *Remote Sensing of Environment*, Vol. 20, pp. 237-252.

Labovitz, M.L. and Masuoka, E.J. (1984) The Influence of Autocorrelation in Signature Extraction: An Example from a Geobotanical Investigation of Cotter Basin, Montana, *International Journal of Remote Sensing*, Vol. 5, No. 2, pp. 315-332.

Landis, J.R and Koch, G.G. (1977) The Measurement of Observer Agreement for Categorical Data, *Biometrics*, Vol. 33, pp. 159-174.

Lanter, D.P. and Veregin, H. (1992) A Research Paradigm for Propagating Error in Layer-Based GIS, *Photogrammetric Engineering and Remote Sensing*, Vol. 587, No. 6, pp. 825-833.

Lillesand, T.M. and Kiefer, R.W. (1994) *Remote Sensing and Image Interpretation*, Wiley and Sons, New York, NY, 749 pp.

Lo, C.P. and Watson, L.J. (1998) The Influence of Geographic Sampling Methods on Vegetation Map Accuracy Evaluation in a Swampy Environment, *Photogrammetric Engineering and Remote Sensing*, Vol. 64, No. 12, pp. 1189-1200.

Lodwick, G.D. (1980) *A Computer System for Monitoring Ecological Changes in Multitemporal Landsat Data*, Unpublished PhD Dissertation, University of New South Wales, Sydney, NSW, 340pp.

Lunetta, R.S., Congalton, R.G., Jensen, J.R., McGuire, K.C. and Tinney, L.R. (1991) Remote Sensing and Geographic Information System Integration: Error Sources and Research Issues, *Photogrammetric Engineering and Remote Sensing*, Vol. 57, No. 6, pp. 677-687.

Lunetta, R.S., Lyon, J.G., Guindon, B. and Elvidge, C.D. (1998) North American Landscape Characterization Dataset Development and Data Fusion Issues, *Photogrammetric Engineering and Remote Sensing*, Vol. 64, No. 8, pp. 821-829.

Ma, Z. and Redmond, R.L. (1995) Tau Coefficients for Accuracy Assessment of Classification of Remote Sensing Data, *Photogrammetric Engineering and Remote Sensing*, Vol. 61, No. 4, pp. 435-439.

MacDougall, E.B. (1975) The Accuracy of Map Overlays, *Landscape Planning*, Vol. 2, pp. 23-30.

Macleod, R.D. and Congalton, R.G. (1998) A Quantitative Comparison of Change-Detection Algorithms for Monitoring Eelgrass from Remotely Sensed Data, *Photogrammetric Engineering and Remote Sensing*, Vol. 64, No. 3, pp. 207-216.

Maling, D.H. (1988) The Concept of the Accurate Map, In *Measurements from Maps: Principles and Methods of Cartometry*, Pergamon Press, Oxford, UK, pp. 144-177.

Marceau, D.J., Howarth, P.J. and Gratton, D.J. (1994) Remote Sensing and the Measurement of Geographical Entities in a Forested Environment. 1. The Scale and Spatial Aggregation Problem, *Remote Sensing of Environment*, Vol. 49, pp. 93-104.

Mark, D.M. and Csillag, F. (1989) The Nature of Boundaries on 'Area-Class' Maps, *Cartographica*, Vol. 26, No. 1, pp. 65-78.

Martin, L.R.G. (1989) Accuracy Assessment of Landsat-based Visual Change Detection Methods Applied to the Rural-urban Fringe, *Photogrammetric Engineering and Remote Sensing*, Vol. 55, No. 2, pp. 209-215.

Martin, L.R.G and Howarth, P.J. (1989) Change-Detection Accuracy Assessment Using SPOT Multispectral Imagery at the Rural-Urban Fringe, *Remote Sensing of Environment*, Vol. 30, pp. 55-66.

Martin, L.R.G., Howarth, P.J. and Holder, G.H. (1988) Multispectral Classification of Land Use at the Rural-Urban Fringe Using SPOT Data, *Canadian Journal of Remote Sensing*, Vol. 14, No. 2, pp. 72-79.

Mas, J-F. (1999) Monitoring Land-Cover Changes: A Comparison of Change Detection Techniques, *International Journal of Remote Sensing*, Vol. 20, No. 1, pp. 139-152.

McGwire, K. and Goodchild, M. (1997) Accuracy, In *Integration of Geographic Information Systems and Remote Sensing*, Star, J.L., Estes, J.E. and McGwire, K.C. (eds), Ch. 6, Cambridge University Press, Cambridge, UK, pp. 110-133.

Mead, R.A. and Szajgin, J. (1982) Landsat Classification Accuracy Assessment Procedures, *Photogrammetric Engineering and Remote Sensing*, Vol. 48, No. 1, pp. 139-141.

Merchant, D.C. (1987) Spatial Accuracy Specification for Large Scale Topographic Maps, *Photogrammetric Engineering and Remote Sensing*, Vol. 53, No. 7, pp. 958-961.

Middelkoop, H. (1990) Uncertainty in a GIS: A Test for Quantifying Interpretation Output, *ITC Journal*, No. 3, pp. 225-232.

Mikhail, E.M. and Gracie, G. (1981) *Analysis and Adjustment of Survey Measurements*, Van Nostrand, New York, NY, 340 pp.

MRWA (1999) *Metadata for Metropolitan Road Centreline data*, Main Roads Western Australia, Perth, WA, CD-ROM.

Muchoney, D.M. and Haack, B.N. (1994) Change Detection for Monitoring Forest Defoliation, *Photogrammetric Engineering and Remote Sensing*, Vol. 60, No. 10, pp. 1243-1251.

Newcomer, J.A. and Szajgin, J. (1984) Accumulation of Thematic Map Errors in Digital Overlay Analysis, *The American Cartographer*, Vol. 11, No. 1, pp. 58-62.

Norcliffe, G.B. (1977) *Inferential Statistics for Geographers*, Hutchinson and Co Ltd, London, UK, 252 pp.

Novak, K. (1992) Rectification of Digital Imagery, *Photogrammetric Engineering and Remote Sensing*, Vol. 58, No. 3, pp. 339-344.

Openshaw, S. (1989) Learning to Live with Errors in Spatial Databases, In *The Accuracy of Spatial Databases*, Goodchild, M. and Gopal, S. (eds), Ch. 23, Taylor and Francis Ltd, Bristol, PA, pp. 263-276.

Openshaw, S., Charlton, M. and Carver, S. (1991) Error Propagation: A Monte Carlo Simulation, In *Handling Geographical Information: Methodology and Potential Applications*, Masser, I. and Blakemore, M. (eds), Longman Scientific and Technical, Harlow, UK, pp. 78-101.

Pala, V. and Pons, X. (1995) Incorporation of Relief in Polynomial-Based Geometric Corrections, *Photogrammetric Engineering and Remote Sensing*, Vol. 61, No. 7, pp. 935-944.

Pilon, P.G., Howarth, P.J., Bullock, R.A. and Adeniyi, P.O. (1988) An Enhanced Classification Approach to Change Detection in Semi-Arid Environments, *Photogrammetric Engineering and Remote Sensing*, Vol. 54, No. 12, pp. 1709-1716.

Podobnikar, T. (1999) Modelling and Visualisation of Spatial Data Error: Monte Carlo Simulations in Slovenia, *GIM International*, July, pp. 47-49.

Quarmby, N.A. and Cushnie, J.L. (1989) Monitoring Urban Land Cover Changes at the Urban Fringe from SPOT HRV Imagery in South-East England, *International Journal of Remote Sensing*, Vol. 10, No. 6, pp. 953-963.

Richards, J.A. (1984) Thematic Mapping from Multitemporal Image Data Using the Principal Components Transformation, *Remote Sensing of Environment*, Vol. 16, pp. 35-46.

Richards, J.A. (1993) *Remote Sensing Digital Image Analysis: An Introduction*, Springer-Verlag, Berlin, Germany, 340 pp.

Ridd, M.K. and Liu, J. (1998) A Comparison of Four Algorithms for Change Detection in an Urban Environment, *Remote Sensing of Environment*, Vol. 63, pp. 95-100.

Rignot, E.J.M., Kwok, R., Curlander, J.C. and Pang, S.S. (1991) Automated Multisensor Registration: Requirements and Techniques, *Photogrammetric Engineering and Remote Sensing*, Vol. 57, No. 8, pp. 1029-1038.

Riley, R.H, Phillips, D.L, Schuft, M.J. and Garcia, M.C. (1997) Resolution and Error in Measuring Land-Cover Change: Effects on Estimating Net Carbon Release from Mexican Terrestrial Ecosystems, *International Journal of Remote Sensing*, Vol. 18, No. 1, pp. 121-137.

Rosenfield, G.H. and Fitzpatrick-Lins, K. (1986) A Co-Efficient of Agreement as a Measure of Thematic Classification Accuracy, *Photogrammetric Engineering and Remote Sensing*, Vol. 52, No. 2, pp. 223-227.

Saraf, A.K. (1999) IRS-1C-LISS-III and PAN Data Fusion: An Approach to Improve Remote Sensing Based Mapping Techniques, *International Journal of Remote Sensing*, Vol. 20, No. 10, pp. 1929-1934.

Seddon, G. (1972) *Sense of Place*, University of Western Australia Press, Perth, WA, 274 pp.

Singh, A. (1989) Digital Change Detection Techniques Using Remotely Sensed Data, *International Journal of Remote Sensing*, Vol. 10, No. 6, pp. 989-1003.

Skidmore, A.K. and Turner, B.J. (1992) Map Accuracy Assessment Using Line Intersect Sampling, *Photogrammetric Engineering and Remote Sensing*, Vol. 58, No. 2, pp. 1453-1457.

Star, J.L., Estes, J.E. and McGwire, K.C. (1997) Integration of Geographic Information Systems and Remote Sensing: A background to NCGIA Initiative 12, In *Integration of Geographic Information Systems and Remote Sensing*, Star, J.L., Estes, J.E. and McGwire, K.C. (eds), Ch. 1, Cambridge University Press, Cambridge, UK, pp. 1-12.

Steele, B.M, Winne, J.C. and Redmond, R.L. (1998) Estimation and Mapping of Misclassification Probabilities for Thematic Land Cover Maps, *Remote Sensing of Environment*, Vol. 66, pp. 192-202.

Stehman, S.V. (1999) Basic Probability Sampling Designs for Thematic Map Accuracy Assessment, *International Journal of Remote Sensing*, Vol. 20, No. 12, pp. 2423-2441.

Story, M. and Congalton, R.G. (1986) Accuracy Assessment: A User's Perspective, *Photogrammetric Engineering and Remote Sensing*, Vol. 52, No. 3, pp. 397-399.

Swain, P.H. (1978) Fundamentals of Pattern Recognition in Remote Sensing, In *Remote Sensing: The Quantitative Approach*, Swain, P.H. and Davis, S.M. (eds.), McGraw-Hill, New York, NY, pp. 136-189.

Tokola, T., Lofman, S. and Erkkila, A. (1999) Relative Calibration of Multitemporal Landsat Data for Forest Cover Change Detection, *Remote Sensing of Environment*, Vol. 69, pp. 1-11.

Toll, D.L. (1985) Effect of Landsat Thematic Mapper Sensor Parameters on Land Cover Classification, *Remote Sensing of Environment*, Vol. 17, pp. 129-140.

Townshend, J.R.G. and Justice, C.O. (1988) Selecting the Spatial Resolution of Satellite Sensors Required for Global Monitoring of Land Transformations, *International Journal of Remote Sensing*, Vol. 9, No. 2, pp. 187-236.

Townshend, J.R.G. and Justice, C.O. (1995) Spatial Variability of Images and Monitoring of Changes in the Normalised Vegetation Index, *International Journal of Remote Sensing*, Vol. 16, No. 12, pp. 2187-2195.

Townshend, J.R.G., Justice, C.O., Gurney, C. and McManus, J. (1992) The Impact of Misregistration on Change Detection, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 30, No. 5, pp. 1054-1060.

Treitz, P.M., Howarth, P.J. and Gong, P. (1992) Application of Satellite and GIS Technologies for Land-Cover and Land-Use Mapping at the Rural-Urban Fringe: A Case Study, *Photogrammetric Engineering and Remote Sensing*, Vol. 58, No. 4, pp. 439-448.

Trotter, C.M. (1991) Remotely Sensed Data as an Information Source for Geographical Information Systems in Natural Resource Management: A Review, *International Journal of Geographical Information Systems*, Vol. 5, No. 2, pp. 225-239.

Van Genderen, J.L. and Lock, B.F. (1977) Testing Land-Use Map Accuracy, *Photogrammetric Engineering and Remote Sensing*, Vol. 43, No. 9, pp. 1135-1137.

Veregin, H. (1989) Error Modelling for the Map Overlay Operation, In *The Accuracy of Spatial Databases*, Goodchild, M. and Gopal, S. (eds), Ch. 1, Taylor and Francis Ltd, Bristol, PA, pp. 3-18.

Veregin, H. (1994) Integration of Simulation Modelling and Error Propagation for the Buffer Operation in GIS, *Photogrammetric Engineering and Remote Sensing*, Vol. 60, No. 4, pp. 427-435.

Veregin, H. (1995) Developing and Testing of an Error Propagation Model for GIS Overlay Operations, *International Journal of Geographical Information Systems*, Vol. 9, No. 6, pp. 595-619.

Walsh, S.J., Lightfoot, D.R. and Butler, D.R. (1987) Recognition and Assessment of Error in Geographic Information Systems, *Photogrammetric Engineering and Remote Sensing*, Vol. 53, No. 19, pp. 1423-1430.

Welch, R. (1982) Spatial Resolution Requirements for Urban Studies, *International Journal of Remote Sensing*, Vol. 3, No. 2, pp. 139-146.

Welch, R. (1985) Cartographic Potential of SPOT Image Data, *Photogrammetric Engineering and Remote Sensing*, Vol. 51, No. 8, pp. 1085-1091.

Welch, R., Jordan, T.R. and Ehlers, M. (1985) Comparative Evaluations of the Geodetic Accuracy and Cartographic Potential of Landsat-4 and Landsat-5 Thematic Mapper Image Data, *Photogrammetric Engineering and Remote Sensing*, Vol. 51, No. 9, pp. 1249-1262.

Welch, R. and Usery, E.L. (1984) Cartographic Accuracy of Landsat-4 MSS and TM Image Data, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 22, No. 3, pp. 281-288.

Wharton, S.W (1987) A Spectral-Knowledge-Based Approach for Urban Land-Cover Discrimination, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. GE-25, No. 3, pp. 272-282.

Wickware, G.M. and Howarth, P.J. (1981) Change Detection in the Peace-Athabasca Delta Using Digital Landsat data, *Remote Sensing of Environment*, Vol. 11, pp. 9-25.

Witmer, R.E. (1977) The USGS Land Use and Land Cover Classification System, *Remote Sensing of the Electro-Magnetic Spectrum*, Vol. 4, No. 4, pp. 10-19.

Wolter, P.T., Mladenoff, D.J., Host, G.E. and Crow, T.R. (1995) Improved Forest Classification in the Northern Lake States Using Multi-Temporal Landsat Imagery, *Photogrammetric Engineering and Remote Sensing*, Vol. 61, No. 9, pp. 1129-1143.

Woodcock, C.E. and Strahler, A.H. (1987) The Factor of Scale in Remote Sensing, *Remote Sensing of Environment*, Vol. 21, pp. 311-332.

Yuan, D. and Elvidge, C. (1998) NALC Land Cover Change Detection Pilot Study: Washington D.C. Area Experiments, *Remote Sensing of Environment*, Vol. 66, pp. 166-178.

Zhang, J. and Foody, G.M. (1998) A Fuzzy Classification of Sub-Urban Land Cover from Remotely Sensed Imagery, *International Journal of Remote Sensing*, Vol. 19, No. 14, pp. 2721-2738.