LAND USE CHANGES ON THE HAUTERE PLAINS: A STUDY USING DIGITAL IMAGE ANALYSIS AND GEOGRAPHIC INFORMATION SYSTEMS

A thesis presented in partial
fulfilment of the requirements for the degree
of Master of Agricultural Science (M.Agr.Sc)
in Soil Science at Massey University

SUMBANGAN BAJA

1995

ABSTRACT

An empirical study was conducted using digital image analysis and geographic information systems to undertake land use/land cover classification and change detection analysis of the Hautere Plains, near Otaki, North Island. The study area, comprising approximately 2000 ha, was mainly flat land where pastoral farming had been the predominant land use. The area, however, has recently undergone significant diversification into horticulture.

Land use changes between 1968 and 1993 were analysed. Aerial photographs taken at the two dates were scanned, registered and classified before the change detection process was carried out. Satellite imagery- a 1990 SPOT XS image- was also evaluated.

Accurate registration of all the images was essential for any analysis of changes to be performed with confidence. All the images were first rescaled to produce a uniform pixel size of 10 metres. Registration to the NZMS metric grid resulted in total RMS errors of 0.46, 0.41, and 0.42 pixels for the 1968 scanned aerial photograph (SAP), the 1993 SAP and the 1990 SPOT XS image, respectively.

In the image classification, eight relatively static land use/land cover categories were defined: pasture, orchards, market gardens, trees, residential sites,

commercial sites, river gravels, and roads. Due to the spectral confusion among particular categories, the results obtained from applying spectral-based classification were refined by incorporating information derived from visual interpretation which made use of photo-interpretation criteria. Merging of data sets was carried out using a binary mask created from the rectified-digital cadastral data and implementing GIS-based overlay functions facilitated in IDRISI. An assessment of the classification accuracy revealed that such procedures resulted in a significant improvement of all levels of classification accuracy (i.e., overall accuracy, user's accuracy and producer's accuracy).

A post-classification comparison technique of digital change detection was applied using the GIS-based operations to develop a quantitative land use/land cover change assessment and to identify the spatial location of changes on a category-by-category basis. The latter was undertaken by means of binary change masking. A similar procedure was also applied to the rasterised cadastral data sets to identify spatial locations of land parcels which had undergone subdivisions. The analysis confirmed that, in the study region, the most common change in land use was from pastoral land to orchards. Most of larger land parcels had been subdivided into smaller holdings ranging from 4 to 10 ha- i.e., the most favoured size for lifestyle blocks as well as properties where orchards and market gardens are found.

The results obtained suggested that the use of scanned aerial photographs at an appropriate scale, complemented by a wealth of site information, is sufficient for computer-assisted classification of land use/land cover types in the study area, and subsequent change detection analysis. If aerial photographs are unavailable for a desired date, it is possible to use satellite imagery as one of the multi-image pair. The processes of image registration, resampling, and data integration ensure that the spatial analysis of change detection can be performed accurately. The quantitative data generated may also give further insights to the land use changes.

ACKNOWLEDGEMENTS

"In the name of Allah, most gracious, most merciful."

I am particularly indebted to my chief supervisor Mr. M. P. Tuohy and co-supervisor Dr. Alan S. Palmer for their supervision, guidance, valuable discussion, and unstinting support throughout the course of my study at Massey University.

My deep gratitude is also extended to the Head of the Department of Soil Science, Massey University Associate Professor Paul Gregg for his continuous encouragements.

Sincere thanks are also due to the Dean of Faculty of Agriculture, Hasanuddin University, Ujung Pandang Dr. Ir. Muslimin Mustafa, M.Sc. for his valuable advice.

Financial assistance provided by the Ministry of Foreign Affairs and Trade (MFAT), New Zealand is gratefully acknowledged; without this the study would not have been possible.

Very special thanks are extended to Ita Thamrin for her generous companionship and unending moral support during the course of my study.

Lastly, to my parents (Papa and Mama) and all my brothers and sisters, I would like to express many thanks for their prayer and patience during my absence at Massey University. I believe their prayer has made my study possible.

"Dedicated to my parents"

LIST OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

CHAPTER I

INTRODUCTION

1.1 General introduction

Remote sensing technology has very rapidly gone through many stages of development, particularly, since the launching of the first earth resources observation satellite in 1972 (Abiodun *et al.*, 1990); it was subsequently followed by a number of satellite remote sensing systems, notably, the Landsat series and SPOT *(Systeme Pour l'Observation de la Terre)*. The substantial advances in remote sensing technology have been augmented by the ready accessibly and widespread use of computer-based systems, used for digitally processing remotely sensed data, as well as scanning systems for converting pictorial data to digital format.

Remote sensing deals primarily with earth surface phenomena, produces information with valuable characteristics, and provides an opportunity to access more up-to-date and reliable information in a more effective manner. The data can be generated in an unbiased form; acquired at a known point in time; displayed accurately; geographically referenced; prepared in real time (or nearly so); and assembled in a useful, storable format (Hardy, 1982). With these characteristics remotely sensed data are now widely used in a variety of applications.

An important aspect of environmental remote sensing is its potential to map land use/land cover and to monitor changes that occur in boundary surface conditions over extended period of time (Milne and O'Neill, 1990). Such applications as land use inventory and monitoring have to date seen most intense use of remote sensing technology. The reasons include the feasibility of such applications as well as the growing concern in the use of land resources. Remote sensing provides a viable source of data from which updated land use/land cover information can be extracted efficiently and cheaply (Fung and LeDrew, 1988); while information on the rate and kind of change in the use of land resources is essential to the proper planning, management, and regulation of the use of such resources (Anderson, 1977). Thus, the mapping of land use/land cover patterns and detecting changes that take place in the area of interest would be a prime prerequisite for appropriate planning and management of land resources in a region. Moreover, efforts at improving the accuracy and timeliness of the land use/land cover maps, and the efficiency with which they are produced, are justified (Ehlers, 1990).

In essence, there are three basic methods used to map resources using remotely sensed data (Claasen, 1992): (a) simple visual interpretation of standard black and white, colour, or false colour (photographic infrared) photographs; (b) visual interpretation of photographic products which have been enhanced for easier interpretation by photographic or computer-assisted digital techniques; and (c) production of photo-maps or thematic maps, statistical and attribute tables from computer-assisted analysis and classification of computer-compatible data recorded by electronic scanner. The use of a computer system for land use classification has recently become a major focus of remote sensing applications, because the information produced can be readily merged with other sources of geocoded data in geographic information systems (GIS).

Major advances in digital land use/land cover classification have been achieved when ancillary data (such as cadastral boundaries, topography, soil type information) have been used to improve the accuracies of the final map products. The ease by which these data sets, including those derived from visual interpretation, can be integrated with digital remotely sensed data has been a major interest in integrated image processing/GIS (Coleman, 1992; Janssen *et al.*, 1990; Kenk *et al.*, 1988; van der Laan, 1988). Using this approach, the overlaying of two or more maps to identify changed land use/land cover polygons can be carried out with ease (Lo and Shipman, 1990; Westmoreland and Stow, 1992). Because these applications generally involve several sources of information, the data sets used need to be always geographically referenced (to a particular co-ordinate system), and the procedure often requires data transfer routines between the various image processing and GIS packages.

The present study deals with the use of digital image processing systems and GIS to undertake each stage of category identification and for the analyses of changed land use/land cover types in the study area.

1.2 Objectives of the study

In this study, three major objectives were addressed: (a) performing image registration of remotely sensed images derived from different sources (i.e., scanned aerial photographs and SPOT XS images); (b) carrying out land use/land cover classification of the study area employing a semi-automated approach; and (c) performing change detection analyses using a post-classification comparison method.

To achieve these goals, the study focused on utilizing an image processing system and/or GIS-based operations at each step of the following processes:

1. To undertake image registration (to the standard NZMS map projection), and to determine if the multiple-date sub-scene employed could be registered precisely enough for accurate change detection.
2. To demonstrate the procedure of data integration to enable the data sets to be interchanged between one another at any stage of the process.
3. To perform unsupervised and supervised classification to generate classification maps of the study area.
4. To develop techniques of binary masking and GIS-based overlay operations.
5. To perform accuracy assessment of classification maps and to compare the accuracies of maps generated using different methods.

6. Using binary masking strategies and overlay operations to incorporate information derived from photo-interpretation in the post-classification refinement to obtain land use/land cover maps of the study area with high confidence in terms of classification accuracy.
7. To introduce GIS-based overlay methods for change detection analyses.
8. To assess the quantitative and spatial changes of land use/land cover types as well as cadastral parcels in the study area.

CHAPTER II

AN OVERVIEW OF REMOTE SENSING AND GEOGRAPHIC INFORMATION SYSTEMS

2.1 Remote sensing

2.1.1 Definition and scope

Remote sensing can be simply defined as the process of deriving information by means of systems that are not in direct contact with the objects or phenomena of interest (Star and Estes, 1990). This process of deriving information, generally, gives rise to a particular form of imagery which is amenable to a number of processes and analysis procedures in a wide variety of applications. The prime objective of remote sensing is thus to extract environmental and natural resources information related to the earth (Lo, 1986), and to provide fundamental biological and/or physical (biophysical) information directly (Jensen, 1986).

In essence, there are two basic processes involved in the remote sensing of earth resources: data acquisition and data analysis (Lillesand and Kiefer, 1987). Each of these processes embraces a number of elements, as depicted in Figure 2.1. Simply, in the process of data acquisition, remote sensors mounted on particular platforms are employed to record the variation of reflectance of the objects concerned in the way these features reflect and emit electromagnetic energy. The data produced from this process together with reference data are further examined and analyzed using either visual interpretation or digital analysis techniques. The results are then compiled in various forms, such as maps, tables or files that may be eventually conveyed to users for decision-making processes.

2.1.2 Electromagnetic energy

The most important medium for environmental (i.e. earth surface/lower atmosphere) remote sensing is electromagnetic radiation (Barrett and Curtis, 1992), which accomplishes information transfer from objects to a sensor. Changes

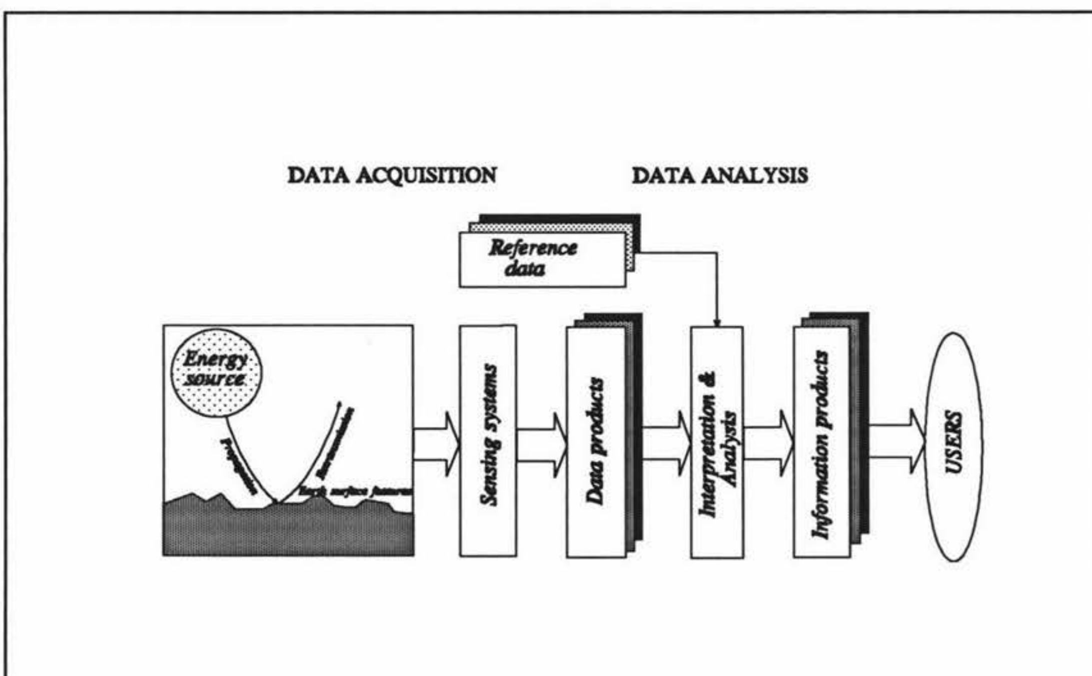


Figure 2.1 Electromagnetic remote sensing of earth resources (modified from Lillesand and Kiefer, 1987).

in the amount and properties of the electromagnetic radiation become, upon detection, a valuable source of data for interpreting important properties of the phenomena with which they interact (Suits, 1983). There are two essential factors encountered: atmospheric interaction and earth's surface interaction. The former refers to both energy propagation from sources and re-transmission from the object of interest as it undergoes scattering, absorption and refraction processes. At the earth's interaction, three main processes occur: reflection, absorption and transmission.

The electromagnetic spectrum occurs as a continuum of wavelengths and frequencies ranging from short wavelength (high frequency), i.e. cosmic rays to long wavelength (low frequency radiowaves). The spectra that are of greatest interest in remote sensing are visible and near-infrared radiation in the waveband 0.4 - 3 μm, infrared in the waveband 3 - 14 μm, and microwave radiation in the 5 - 500 mm wavelength (Curran, 1985). This study will employ imagery from the visible and near-infrared wavelengths of the spectrum.

2.1.3 Spectral signatures

Each object reflects radiation in a unique and identifiable manner (Jensen, 1986). This infers that each type of surface on the earth (e.g. bare ground, water, residential area, pasture, forest, rocks) tends to have discrete spectral characteristics which are referred to as the *spectral signature* of that earth surface features.

A related term is *spectral response pattern*, that is, the response of material as a function of wavelength to incident electromagnetic energy, particularly in terms of the measurable energy reflected from and emitted by the material (Jensen, 1986; Johannsen and Sanders, 1982). Lillesand and Kiefer (1987) used this term in lieu of spectral signature when identifying the condition of various objects, particularly, of the same type. Both terms refer to the distinctive spectral reflectance and/or emittance characteristics of features and thus permit an assessment of the type and/or the condition of that earth surface feature.

Figure 2.2 depicts the idealized spectral response characteristics for wavelengths from 0.4 to 2.6 μm of three basic earth features: green vegetation, dry bare soil (grey-brown loam), and clear water. Lillesand and Kiefer (1987) described the configuration of these curves as representing average reflectance, that is, in general an indicator of the type and condition of the features to which they apply. Reflectance curves above or below the average reveal a particular state associated with the land cover of interest.

It is clearly seen that the spectral reflectance of green vegetation varies considerably with wavelength. In the visible part of the spectrum (0.4 -0.7 μm), reflectance is comparatively low. The interactions of energy-matter with vegetation in this part of spectrum is predominantly controlled by plant pigmentation (Lusch, 1989), which absorbs much radiation. Reflectance and transmittance characteristics within the near-infrared portion of the spectrum (0.7 - 1.35 μm) is high, however. In this region, a relatively small amount of energy is absorbed (Gates, 1970; Thomas and Timmins, 1982), and the physical control is internal leaf structure (Lusch, 1989).

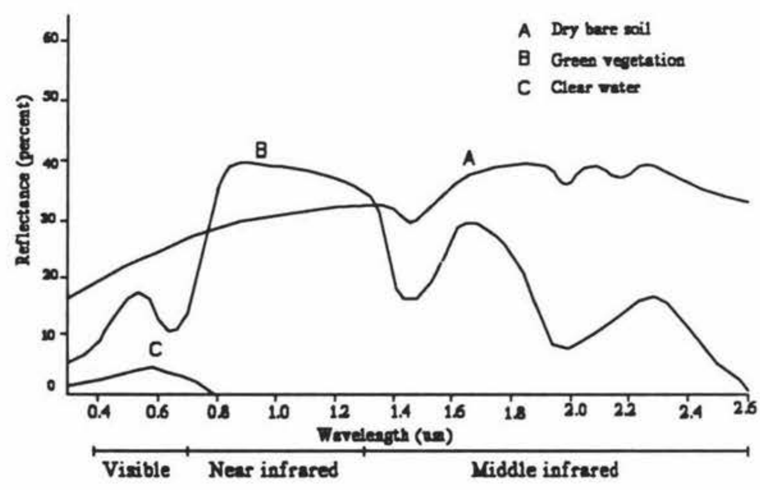


Figure 2.2 Significant spectral response characteristics of soil, vegetation, and water (adapted from Harrison and Jupp, 1990; Hoffer, 1978; and Lillesand and Kiefer, 1987).

The reflectance curve of green vegetation shows there are five wavelength regions where reflectance is relatively low: at or near 0.4, 0.6, 1.4, 1.9, and 2.7 μm. The first two bands are associated with chlorophyll absorption of energy, while the last three correspond to the high energy absorption of water. The dry bare soil curve also drops slightly in these three bands. Reflectance of clear water is very low (almost zero) within the range of 0.7 - 2.7 μm. This is due to the extremely high near-infrared energy absorption of water. As a result, clear water does appear as black tone in this spectral range. This phenomenon can be seen in the band three multispectral mode of SPOT imagery, as used in this study. The distinctive configuration of reflectance curves hence infers the spectral characteristics of land cover types and is the basis of selecting wavelength regions in which remote sensing data are acquired for a particular application.

2.1.4 Data acquisition and analysis

From remote sensing standpoint, information about the earth's surface is acquired by detecting the electromagnetic energy that corresponds to ground features, using either photographic or scanning systems. Both sensing systems are normally mounted on a 'stable' platform, e.g., aircraft and satellites, but cameras and scanners may also be hand-held.

For a particular application the choice of appropriate data from the most suitable platform is crucial for success. Cracknell (1981) noted that relevant factors which need to be taken into account include: (a) the extent of the area to be covered, (b) the speed of development of the phenomenon, and (c) the number and wavelength ranges of the spectral bands of the sensor available on the platforms under consideration. In addition, consideration about the degree of detail (resolution) may also be important.

In the process of data acquisition the information can be generated either photographically or electronically. The process of photography utilizes a chemical reaction on the surface of a light sensitive film to detect energy variations within a scene. With the electronic sensor, on the other hand, electrical signals corresponding to energy variation are generated to produce a image (Lillesand and Kiefer, 1987). The major advantages of photographic imagery are the technical simplicity of its processing and interpretation, and its availability at a lower cost than digitally recorded data (Harrison and Jupp, 1990).

Unlike digital scanner photographic devices can only provide the information on a given scene within a limited range of the electromagnetic spectrum, generally in the visible and near-infrared region. In addition, the interpretation of photographic imagery relies heavily on the skills of interpreters. Nevertheless, photographic data can, for computer-aided analysis, be converted into a digital image using scanning system.

Manual interpretation procedure relies heavily on the basic elements of photo interpretation which are tone/colour, size, texture, shape, pattern, height, shadow, site, association of features in the scene (Caroll *et al.*, 1977; Estes *et al.*, 1983; Lillesand and Kiefer, 1987; Star and Estes, 1990). Because an analyst can discriminate only about 8 to 16 shades of grey (Jensen, 1986), in the image interpretation process one needs to incorporate several crucial techniques and aids. These may include perceptual models, collateral material, and viewing facilities which are further complemented by hypothesis testing and convergence of evidence relating to the features of interest (Estes *et al.*, 1983).

Most of the computer-aided approaches may, in the other hand, employ only a few of these basic elements of image interpretation. Jensen (1986) noted that digital image analysis appears to be dependent primarily on just the tone/colour of objects in the scene, using fundamental statistical pattern recognition techniques. Nonetheless, current research on the decision theoretic approach (expert systems) tends to utilize more parametric components in the analysis procedure. This will allow the incorporation of most of the elements of image interpretation into computer-assisted image analysis (Estes *et al.*, 1983).

2.2 Digital image analysis

2.2.1 General

Remote sensing data in digital format have recently become much more commonly used, since digital analysis facilities have become more easily available and more capable. With the aid of computer systems, a massive amount of digital remotely sensed data can be read and manipulated to produce a wealth of valuable information. Today, even data which were not recorded in digital form initially, are often digitised by means of scanner systems.

Digital image data spatially comprise a number of individual picture elements (pixels) addressed in two dimensional space. Every pixel is radiometrically quantized into discrete brightness levels (Richards, 1986), or digital numbers (DN). The intensity value of a pixel is dependent upon the spectral reflectance characteristics of the object being sensed, and the number of brightness levels used to describe the intensity range of the image (Curran, 1985). Data with 8-bit radiometric resolution will have $2^8 = 256$ levels of brightness. Thus the amount of data to be processed will depend on spectral resolution and the number of spectral bands to be used as well as the physical size of an image, i.e., number of columns and rows.

Digital image processing, in essence, involves the following types of computer-assisted operations: image preprocessing, image enhancement, and image classification. The processing techniques which may be applied vary considerably, but commonly include many forms of mathematical, statistical, geographical, and data management processing (Nichols, 1983).

2.2.2 Image preprocessing

Remotely sensed data acquired from either aircraft or spaceborne platforms are commonly subjected to different deformations due to the earth's curvature, platform movement, orbit variations, and the image projection (Nguyen and Ho, 1988). These adverse factors may bring about subtle changes to the information in an image. The intent of image preprocessing is thus to correct image data for distortions and changes in scene illumination which stem from the data acquisition process. This may lead to obtaining a more faithful representation of the original scene for subsequent processing and analysis procedures.

To correct raw digital images, three types of operation are commonly implemented: geometric correction, radiometric correction, and noise removal. Geometric correction is required to compensate for the distortions introduced by various sources of geometric deformation, so that the corrected image will have the geometric integrity of a map (Lillesand and Kiefer, 1987). In the radiometric correction procedure, the raw image is spectrally manipulated to give the measured brightness values of the pixels in an image radiometrically corresponding to the spectral reflectance of object of interest in the scene. While noise removal is purposed to minimize any unwanted-haze disturbances of image performance. In addition, image preprocessing is also required to restore detector anomalies such as line dropouts, striping, and/or line start problems (Jensen, 1986).

However, when images are purchased they have normally been subjected to these preprocessing operations already, so that the remotely sensed data need only to be rectified to conform with specific georeferencing systems, such as Polyconic Projection, Mercator Projection, UTM (Universal Transverse Mercator) Projection, Polar Stereographic Coordinate System, etc. In this project, all the digital data were rectified to the New Zealand Metric Grid System, which uses seven digits to give coordinates to the nearest metre.

2.2.3 Image enhancement

Image enhancement algorithms are applied to remotely sensed data to increase the information content of an image to be presented for visual analysis, or frequently for subsequent computer-assisted analysis. Images which have undergone these procedures allow the distinction between features in the scene to be more easily interpreted both visually and using a computer. Computer enhancement aims to amplify the slight differences of radiometric and spectral characteristics to make them readily observable (Lillesand and Kiefer, 1987).

Enhancement techniques range from the relatively simple single-band contrast enhancement to two-dimensional filtering in the spatial or Fourier domain, to complex cluster analysis, sometimes coupled with ratioing of spectral bands (Hord, 1986). Lillesand and Kiefer (1987) categorized these techniques into three main groups: contrast manipulation, spatial feature manipulation, and multi-image manipulation. Each group may entail several computer-assisted operations (Table 2.1).

Depending on the purpose of the application, Curran (1985), however, noted that the most popular image enhancements are contrast stretching, band to band ratioing and subtraction, digital filtering, data compression, and colour display.

2.2.4 Image classification

The classification of remotely sensed data is aimed mainly at automatically categorizing the spectral values of pixels in a given image into land use/land cover classes specified by the user. When using multispectral data, this type of algorithm will reduce the number of data bands to a single band with the data presented in groups of user defined themes. Likewise, the data will undergo

Table 2.1 Image enhancement techniques in digital image analysis (adapted from Lillesand and Kiefer, 1987).

Group	Types of operation
1. Contrast manipulation	a) grey-level thresholding b) level slicing c) contrast stretching - linear stretching - histogram equalization
2. Spatial feature manipulation	a) spatial filtering - low pass filtering - high pass filtering b) edge enhancement c) Fourier analysis
3. Multi-image manipulation	a) Spectral ratioing, addition, differencing, and multiplication b) principal components c) canonical components d) vegetation components e) intensity-hue-saturation (IHS) colour space transformation.

transformation from *spectral classes* into *information classes*. The former refers to the classes that are inherent in the remote sensor data, while the latter are those that user defines (Jensen, 1986; Jensen *et al.*, 1983; Campbell, 1983).

In the classification procedures, feature identification can be done by means of *spectral-*, *spatial-*, and *temporal-pattern recognition*. In simple terms, *spectral-pattern recognition* refers to utilizing brightness values of pixels which correspond to the cover classes being identified. Thus the procedure utilizes pixel-by-pixel spectral information as the basis for automated land cover classification (Lillesand and Kiefer, 1987). In *spatial-pattern recognition*, various spatial information of image, such as feature size, texture, repetition, shape, directionality, and context, are used as the basis of feature identification. *Temporal-pattern recognition* considers time as the basis for categorization of feature classes.

There are two broad types of classification procedure which are usually recognized, i.e., supervised and unsupervised classification. In the supervised approach, the assumption being adopted is that each spectral class can be described by a probability distribution in multispectral space (Richards, 1986). Such a distribution will determine the probability of a pixel representing a particular class at any given location in multispectral space. It is thus important, prior to the classification stage, to define representative samples for each category. This is undertaken by collecting numerical data from training areas on spectral response patterns of land cover categories. This step is known as the training stage, while the other two stages of this procedure are classification and presenting output (Estes *et al*., 1983; Lillesand and Kiefer, 1987).

There are three types of classification strategy in common use for the supervised approach. They are: the *minimum distance to mean classifier*, the *parallelepiped classifier*, and the *maximum likelihood classifier*. In the minimum distance to mean strategy, the training data are used to determine class means, and classification is then performed by allocating pixels to the class of the nearest mean. This technique does not make use of covariance information, but instead depends only upon the mean positions of the spectral classes.

In the parallelepiped classification, the decision boundary of each class is determined by the range, i.e., the highest and the lowest, of digital numbers in each band. Hence this will perform better result for the data with high category variance. However, overlap occurs when the category distribution exhibits high correlation and covariance, which may give rise to ambiguity in observing unknown pixels, i.e., in which rectangular decision region they may include. These difficulties could be resolved using maximum likelihood algorithm. This technique quantitatively evaluates both the variance and covariance of the category spectral classes when classifying an unknown pixel (Lillesand and Kiefer, 1987), and modulates its decision with direction based upon the information in the covariance matrix (Richards, 1986).

Unlike supervised classification, the unsupervised technique minimizes attempts from the image analyst to collect data from the training area. The image data pixels require only to be grouped by aggregating them into the natural spectral groupings or clusters present in the scene. Therefore, the procedure is accomplished on the basis of clustering algorithms (Jensen, 1986; Lillesand and Kiefer, 1987; Richards, 1986).

2.3 Geographic information systems

2.3.1 Definition

Geographic information systems (GIS) have been defined in various ways by many authors. In simple terms, a GIS is a computer-assisted system for the capture, storage, retrieval, analysis, and display of spatial data (Clarke, 1986; Parker, 1988), to solve complex research, planning and management problems (Fischer and Nijkamp, 1993). Star and Estes (1990) noted that GIS is an information system that is designed to work with data referenced by spatial or geographic coordinates. Therefore, a GIS may be viewed as data base system in which most of the data are spatially indexed, and upon which a set of procedures operates in order to answer queries about spatial entities in the data base (Smith et al., 1987)

Cowen (1987) described GIS as a management tool and a decision support system. The former relates to the existence of powerful systems in GIS capable of handling large volumes of spatial data derived from a variety of sources. In this context, when complex information is of interest, such systems (i.e., computer-based GIS tools) can be effectively implemented to improve the user's ability to make decisions in research, planning and management.

2.3.2 Basic elements of a GIS

GIS was designed originally for two reasons (Marble and Peuquet, 1983): (a) the volume of data to be processed was so large that manual methods would have been incapable of completing data processing and evaluation problems, and (b) the manipulations were sufficiently complex that, when coupled with large data volumes, the task could not be completed without substantial error. To accomplish the task, a computer-based GIS is in essence facilitated with five component subsystems (Smith et al, 1987): (a) data encoding and input processing, (b) data management, (c) data retrieval, (d) data manipulation and analysis, and (e) data display. For any given application of a GIS, it is important to view these elements as a continuing process (Star and Estes, 1990).

Data encoding and input processing involve collecting and/or processing spatial information derived from existing maps, remote sensors or other forms of data. The information can be input into the systems by disk, tape, keyboard, or digitizer. Scanning existing aerial photographs is a common technique used to convert pictorial information into digital format. Often these data will require manual or automated pre-processing prior to encoding (Smith *et al.*, 1987; Star and Estes, 1990). *Data management functions* govern the creation of, and access to, the database itself, and provide consistent methods for data entry, update, deletion, retrieval (Star and Estes, 1990). In the *data retrieval subsystem* establishing such access procedures is of considerable importance to permit the data to be quickly retrieved by the users for subsequent analysis.

The *data manipulation and analysis subsystem* consists of operational procedures which work with database contents, through user-defined rules, to produce new information. This subsystem is inclusive of format conversion, geometric conversion, generalisation and classification, enhancement, abstraction, spatial analysis, statistical analysis and measurement (Rhind and Green, 1988). The final outputs produced from this algorithm are then created. This phase is known as the *data display/reporting subsystem* and it should be capable of displaying all or part of the original database as well as manipulated data and the output from spatial models in tabular or map form (Marble, 1993).

Of the five functions facilitated in GIS, three are of particular importance in this study, i.e., data encoding and input processing, data manipulation and analysis and display/reporting the final output.

2.3.3 Types of spatial data in a GIS

Every geographical phenomena can in principle be represented by a point, line or area (polygon) to reduce the complexity of data sets (Burrough, 1989). These spatial features are complementary to building the database in GIS.

The standard approach in GIS is to represent *polygon* boundaries as lines and to represent *lines* as a sequence of very short, straight line segments which can, in turn, be represented by an ordered sequence of *points* representing the end points of the short line segments (Marble and Peuquet, 1983). To generate a line network, 'pointers' are built into the data structure with the aid of *nodes* (Burrough, 1989). The latter are used to indicate the junction between lines or the ends of these line segments. A point is a spatial object which consists of only a single XY coordinate pair, while line and polygon entities generally consist of an ordered sequence of XY coordinate pairs (Marble and Peuquet, 1983). Because the area entity is a region enclosed by line features (Aronoff, 1991), in a given database a polygon may have additional topological properties such as area, perimeter and shape.

Points, lines, or polygons that represent desired geographic features should be converted to a format that can be stored, manipulated, and displayed by a computer-based information system (Walsh,1987). There are two basic types of data format used in GIS: the *raster model* and the *vector model*. The raster data model consists of an array of grid cells referenced by column and row number (Burrough, 1989). The location of geographic objects or conditions is defined by the row and column position of the cells they occupy (Aronoff, 1991). Hence, a point is represented by as single grid cell, a line by a number of neighbouring cells strung out in a given direction, and area by an agglomeration of neighbouring cells (Burrough, 1989). An example of this data structure is remotely sensed data generated either by a multispectral scanner from various platforms or by scanning existing aerial photographs.

In the vector model, objects or conditions in the real world are represented by the points and lines that define their boundaries (Aronoff, 1991), and indicated by a series of XY coordinate pairs. This data model therefore provides information on the real positioning of features in the study area. An example of generating a vector data structure, which is in common practice, is the use of digitizer to input pictorial or graphic data, such as maps and photographs, as strings of coordinate values.

The primary theoretical difference between the two data models is that the raster structure stores information on the interior of areal features and implies boundaries; whereas the vector structure stores information about boundaries, and implies interiors (Berry, 1988). In both models the spatial information is, however, represented using homogeneous units (Aronoff, 1991). The homogeneous units in the raster data are the cells, while in the vector data, the homogeneous units are the points, lines and polygons. A comparison of both data models is shown in Table 2.2. Both types of data structure were employed in this study. Several data sets underwent format conversion from vector to raster, or *vice versa*, to make use of the advantages stated in Table 2.2.

Table 2.2 A comparison of the general characteristics and suitable functions of raster and vector models for GIS (after Maguire *et al.*, 1991).

	Raster	**Vector**
General	Simple data model	Complex data model
	Discrete space (field)	Continuous space (feature)
	Uses cheap technology	Uses expensive technology
	Ease of data collection	Difficult and expensive data collection
	Ease of processing	Difficulty of processing
	Area-oriented	Boundary-oriented
	Environmental application	Socio-economic application
	Fixed density	Variable density phenomena
	No topological processing	Topological processing
Appropriate function	Proximity analysis	Network analysis
	Modelling and simulation	Cartographic drawing and annotation
	Intervisibility	Coordinate geometry (COGO)
	Boolean overlay operation	

2.3.4 Basic characteristics of spatial data

Geographic data are commonly characterised as having two fundamental components (Aronoff, 1993; Dangermond, 1993): (a) the actual phenomenon (physical dimension or attribute), such as the variable, class, value, name, etc.; and (b) the spatial location of the phenomenon. These attributes can be both qualitative (e.g., land use at a location), or quantitative (e.g., elevation at the same location). Such phenomena, residing at a particular location, often change either dependent or independent of one another with respect to time. Thus the temporal dimension is also critical and particularly relevant to GIS and is thought of as the third characteristic of spatial data (Langran, 1992). It enables the identification and analysis of change in spatial information.

In a software system, such as ARC/INFO, a vector-based GIS package, these feature attributes are stored in the polygon attribute table (PAT) and arc attribute table (AAT) files for polygon and arc coverages, respectively. A coverage is the basic unit of storage in ARC/INFO, and it contains both the locational data and thematic attributes for map features in a given area (Peuquet and Marble, 1993). Each polygon, arc, or point is topologically represented by one record number which is assigned as a User_ID. Other coverage features in ARC/INFO are: *tics* (i.e., the registration or geographic control points for a coverage); *nodes* which represent end points and the location where line features connect; and *annotation* (text) which is used to label coverage features (Peuquet and Marble, 1993).

In the data base, all of the information about a given feature is stored relative to other features. The relationships used to represent the connectivity or contiguity of these features are referred as *topology.* This provides the basis for a variety of geographic analyses without having to access the absolute locations held in the coordinate files.

CHAPTER III

STUDY AREA AND DATA COLLECTION

3.1 Study area

3.1.1 General description

The study area covers approximately 2000 ha and is situated in the Hautere Plain about 1 km south of Otaki. The area is cut by the National State Highway No.1; the greater part of the study region lies eastward of the Highway. It is bounded by the Otaki River to the north, the hills to the east and Te Horo railway station to the south (Figure 3.1). Part of the western boundary of the area follows Te Waka Road which occurs along an old sea cliff.

In addition to the major river of the area, Otaki River, several small streams also exist in the plains. Paraha Stream flows from the eastern hills, and down across the lower terrace. In the southernmost part of the area, Mangaone Stream and its tributaries flow over some parts of the loess-covered surfaces and dunesands before discharging into the Tasman Sea approximately 3 km west of the study area.

The area is mainly flat land, though there is some variation in relief over the whole area. Elevation rises from 10 to 80 m above sea level.

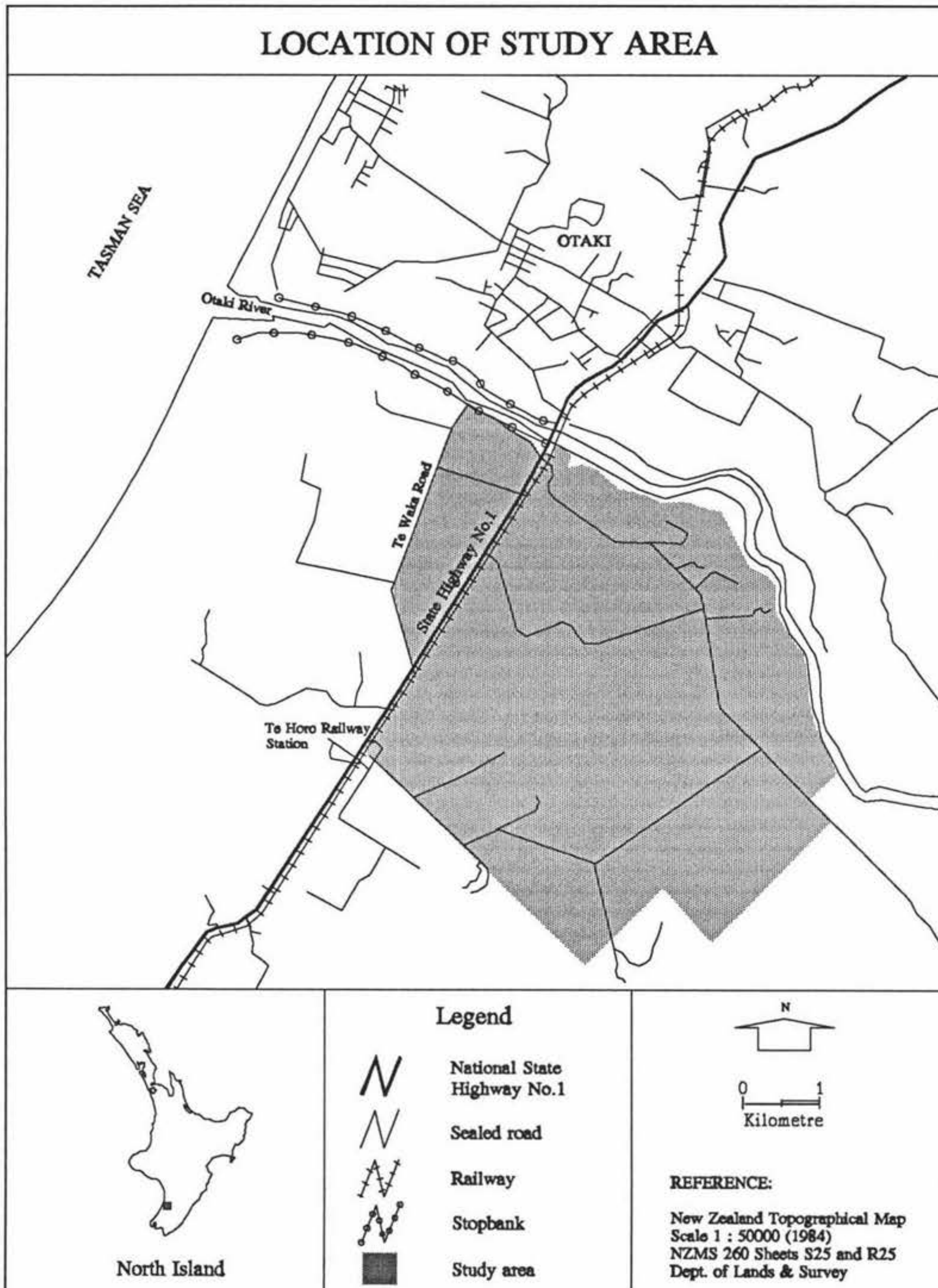


Figure 3.1 Map of the study area.

3.1.2 Physiography

Four major physiographic units are recognized in the study area: river flats, intermediate terraces, high terraces, and dunesands. The intermediate terrace is the dominant landform unit in the study area and is considered as an aggradation terrace of Ohakean age (10,000-25,000 years B.P). The northern part of the intermediate terrace is stony to the surface, but a low 2-3 m cliff east-west across the middle of the plains separates the stony soils from an area in the south that contains soils developed mostly in fine-grained alluvium over gravels (Palmer *et al.*, 1988).

Younger degradation terraces along the Otaki River step down to a narrow band of river flats where soils are less than 1,000 years old. The western part of the study region, near State Highway No.1, is an area of dunesand between 6,000-10,000 years old because it rests on Ohakean gravel and is truncated at Te Waka Road by the 6,000 year old post-glacial sea cliff.

In the south and east, the area comprises small remnants of high loess covered aggradation and marine terraces that are older than the last interglacial (Palmer *et al.*, 1988). The marine terrace is formed in marine sand which occasionally forms the soil.

3.1.3 Soil types

Soils of the Hautere Plains reflect the patterns of parent materials and drainage conditions associated with the landform and physiographic features. Soil forming parent materials consist mainly of quartzo-feldspathic alluvium except over western parts of the coastal terrace where soils have formed from the windblown sands (Cowie and Wilde, 1980), and over the high terraces where soils are mostly derived from quartzo-feldspathic loess of Ohakean age (Palmer and Wilde, 1986). The alluvium, which is derived from greywacke-argillite rocks, comprises both coarse- and fine-textured materials deposited by the Otaki River as well as the more recent and predominantly fine textured alluvium deposited by the Paraha Stream (Cowie and Wilde, 1981).

Descriptions and definition of soils of the Hautere Plains are provided by Palmer and Wilde (1986), and arranged according to the physiographic position. On the River flats (low terraces), three main soil series are found: Rangitikei series, Otaki series, and Manawatu series. The first two series are characterized by weakly weathered stony materials and undifferentiated alluvium due to them being flooded each year (except where protected by stopbanks). Both soil series are correlated to Typic Udorthents of the U.S. Soil Taxonomy. The Manawatu series (or Dystric Eutrochrepts) are mostly formed in moderately well drained soils which are seldom flooded. This area is now developed mostly for horticulture.

On the intermediate terraces, Typic Dystrocrepts occur over approximately two thirds of the study area, comprising mainly Ashhurst series, Kawhatau series, and Hautere series with stony and well drained characteristics. The Hautere series area found to be more developed and are associated with other soils, i.e., Te Horo series, Paraha series, and Ohakea series. Most of these soils are formed in fine grained alluvium with drainage condition ranging from well drained to poorly drained, respectively. Soil units on the dunesands occurring on the intermediate terrace comprise mainly Koputaroa series and Waitawa series with well drained and poorly drained colluvium, respectively. These types of soils also occur on the dunesands in the western part of the study area, and also rarely on the high terraces in the south.

The main soils of the high terraces are characterized by moderately to well weathered loessial materials. Four major soil series are found: Levin series, Waitohu series, Shannon series, and Rahui series. The drainage condition ranges, respectively, from well drained to poorly drained. Most of the well drained soil are correlated to Typic Dystrochrepts, whereas poorly drained soils are various Aquepts.

3.1.4 Land use and vegetation

The Hautere Plain has been subjected to a considerable change in land use over the last few years. A number of areas have been, in recent years, developed into perennial horticulture. Some larger tracts used for dairying, which has been the predominant land use in the plain from the time of bush clearing, have been also intensified and diversified with market gardening, goat farming, and deer farming. In addition, there has recently been more intensive use of agricultural mechanisation in some subdivisions for annual cropping.

At the present day land use and vegetation in the study area can be categorized into four major groups: (1) agriculture, (2) horticulture, (3) bush and forest and other high-stand trees, and (4) other uses. For the agriculture group, the dominant land use is dairy, sheep and cattle grazing on dominantly browntop *(Agrostis capillaris)*, ryegrass *(Lolium spp.)*, and white clover *(Trifolium repens)* pasture. Deer and goat farming also occur in some areas of the pasture grasses. Several parcels of land are being intensively cultivated for a variety of annual vegetable crops (market gardening). Accordingly, these parcels show up as bare land at the time of cultivation, when the land is being prepared for planting or the crops have just been harvested.

For the land in horticulture, four crop types could be identified: kiwifruit, pipfruit, berryfruit, and small areas of stonefruit and citrus. Kiwifruit is the major horticultural land use and covers some 150 ha of the plains; it is followed by berryfruit, and pipfruit. With a great deal of diversification in horticulture, it is hard to differentiate between the various crop types. Each of the horticultural units is bordered by the narrow-spaced shelter belts and is small in area when compared to the pastoral farms.

Despite the area being used predominantly for agriculture or horticulture several blocks of forest and bush can still be found scattered throughout the plains. Pine trees appear as shelter belts both bounding the horticultural blocks and along the metalled roads in the northern part of the study area. Scrubby vegetation occurs along the south bank of the Otaki River, within some forest stands and along the forest block margins. Native forest stands occupy portions of land immediately eastward of the State Highway No.1. These areas of remnant "bush" contain a wide variety of plant species. Three of the most important and dominant species are totara *(Podocarpus totara)*, titoki *(Alectryon excelsus)*, and matai *(Prumnopitys taxifolia)*. These species account for over two thirds of the forest cover; and it is from these species that the forest is named totara-titoki-matai forest, as their crown cover percentages are about 27, 26 and 14.5, respectively (Druce, 1965). The rest mainly consist of a variety of small-leaved shrubs such as melicope *(Melicope Simplex)*, rohutu *(Lophomyrtus obcordata)*, and coprosmas *(Coprosma Crassifolia)*, and of other high-stand trees, e.g., mapou *(Myrsine australis)*, akeake *(Dononaea viscosa)*, mahoe *(Melicytus ramiflorus)*, hinau *(Elaeocarpus dentatus)*, and rewa-rewa *(Knightia excelsa)*.

Residential land tends to exist as single, and widely scattered units throughout the plains. Besides houses, other buildings such as sheds, storage, pack houses, garages, and poultry farming buildings were also recognized in the area. Another important land use identified in the northern edge of the study area is metal extraction from the Otaki River. This activity has been carried out by two organisations- Golden Bay N.Z. Limited and New Zealand Railways- to provide the country's demand for shingle. The materials produced are mainly used for ballast, road metal and concrete aggregate, and smaller amounts for embankment widening. A study by the Manawatu Catchment Board found that the resource management of metal extraction was having some effects on river protection works and could also possibly be contributing to the increase in flood damage (Brougham, 1976).

A more detailed examination of land use/land cover types in the study area and their influence on remotely sensed data classification is presented in Chapter V Section 5.3.

3.2 Data collection

3.2.1 Data sources and their characteristics

The primary data sources used for this study include aerial photographs, SPOT multispectral (XS) imagery, and digital cadastral data. Aerial photographs were acquired on 7 October 1968 and 10 March 1993 at a nominal scale of approximately 1:66,600 and 1:50,000, respectively. These surveys enabled the identification of land use/land cover change in the study area over a 25-year time range. Two photographic enlargements each with a scale of 1:20.000 were also used for this study. The original contact prints were scanned and provided the digital data required for computer-based image processing. Full technical details of these aerial photographs are given in Table 3.1.

The SPOT multispectral (XS) imagery that had been corrected radiometrically and geometrically was purchased from the Landcare Research, Palmerston North. The imagery with a Scene_ID of K443 and J432 was recorded over the study area on 30 January 1990. Look angle and sun elevation angle were 1.9°W and 53°, respectively. From the original full scene image, the subscene used for this study starts at line 1120 and pixel 1900. Table 3.2 presents the general characteristics of SPOT data and the sensor system employed.

Another important data source used was 1993-digital cadastral data purchased from the Department of Survey and Land Information, New Zealand. A 1968-cadastral map was also purchased so that an analysis of property boundary change between the two dates could be carried out. Complementing these data, the 1:50,000 N.Z. Topographical Maps (NZMS 260 sheets S25 and R25) were available to assist the positional identification of features based on the NZMS coordinate system.

Table 3.1 Full descriptions of the Hautere Plain photographs.

	1968 photograph	1993 photograph
Date flown	7 October 1968	10 March 1993
Time of acquisition	11.10 A.M. local time	2.34 P.M. local time
Flying height	25,000 ft	25,000 ft
Camera type	Ag 76	Nr 127754
Focal length	114.41 mm	152.72 mm
Scale	1:66,600	1:50,000
Survey number	3022	9284
Run number	4135/9	A/2
Agency	N.Z. Aerial Mapping Limited	N.Z. Aerial Mapping Limited

3.2.2 Software used in this study

Three primary software packages were used to conduct this study, i.e.,

a. DRAGON *version 4.0*, which is mainly used for digital image analysis, and thus utilizes grid-based format files.

b. The GIS software-IDRISI *version 4.1g* which employs both vector and raster formats. The system is therefore used for both GIS data analysis and some digital image processing.

c. PC ARC/INFO-a vector-based GIS package *version 3.4 D*. This software package is widely used for a variety of GIS applications.

Table 3.2. SPOT sensor system characteristics (after Jensen, 1986).

Characteristics of the HRV[a] sensors	Multispectral (XS) mode	Panchromatic mode
Spectral bands	0.50-0.59 μm	0.51-0.73 μm
	0.61-0.68 μm	
	0.79-0.89 μm	
Instrument IFOV[b]	4.13°	4.13°
Ground sampling interval at nadir	20 x 20 m	10 x 10 m
Number of pixels per line	3000	6000
Ground swath width at nadir	60 km	60 km
Pixel quantization	8 bits	6 bits DPCM[c]
Image data bit rate	25 Mb/s	25 Mb/s

[a] High-resolution visible

[b] Instantaneous-field-of-view

[c] Digital pulse code modulation, i.e., a mode of data compression that does not degrade the radiometric accuracy of the 8-bit or 256-grey-level image data.

3.2.3 Preliminary data processing

The two contact prints from each aerial survey were scanned employing the ADOBE PHOTOSHOP program *version 2.5* operating under WINDOWS. The system was connected to a *Microtek Scan Maker II SP* scanner which read the data and sent the output to the system. The TIFF format was chosen as an intermediate file that could be further converted to the formats compatible with the programs used in the analysis procedures. Spatial resolution specified for both images was 120 dpi (dots per inch). Given the scales of 1:66,600 and 1:50,000, this led to resolutions of 14.10 and 10.58 m, respectively, for the 1968 and 1993 images.

The three bands-SPOT XS data were received in EPIC format and they needed to be converted for use in DRAGON. To do this, the FCONVERT-DRAGON module was employed. This module also allowed the subsampling algorithm to subset the imagery for the area of interest.

As previously mentioned only the 1993-digital cadastral data were available. Hence, the 1968-cadastral information had to be digitized utilizing the ARCEDIT module of ARC/INFO. This procedure was done as follows. Using GENERATE module, the 1993-numeric data sets were first converted to an arc-coverage. A new coverage was then created from the 1993 coverage using COPYCOV function. By reference to the 1968 cadastral map, changes were then made to the new coverage through a manual editing process. This procedure thus gave rise to a new coverage (1968 information) having the same *tics* and map *boundaries* as the original (1993 data) coverage.

CHAPTER IV

IMAGE REGISTRATION AND DATA INTEGRATION

4.1 Image registration

4.1.1 Introduction

Prior to the data analyses, the processes of rectification and registration need to be undertaken to permit accurate comparisons to be made between images acquired at different times. Rectification is applied when the spatial arrangement of features in the data set is to be made conformable with a specific georeferencing system. While in the registration process, ground features do not necessarily have to be on an absolute georeferencing scheme. Rather, they simply need to be in geometric conformance from one to another. Thus the difference is that in image rectification the reference is a map in standard map projection, whereas in image registration the reference is another image (Jensen, 1986).

Haralick (1973) defines registration as the translation-rotation alignment process by which two images of like geometries and of the same set of objects are positioned coincident with respect to one another, so that corresponding elements of the same ground area appear in the same place on the registered image. The process may involve image rotation, translation, and scaling or stretching (Thomas *et al.,* 1978). Therefore, changes in size and number of pixels can be made to allow geometric alignment between two or more images acquired at different dates and by different sensors.

4.1.2 Basic principles

To transform a digital image into a corrected image corresponding either to particular cartographic projections or to other images, one needs to know the mathematical functions which relate the coordinates of the two sets of data, and to select a method of assigning the pixels of the transformed image. This procedure essentially entails three different steps: (a) identification and selection of ground control points (GCPs), (b) choosing a mapping function, and (c) image resampling. The last two steps are straightforward, while the first step relies on manual operation which is labour intensive and time consuming (Chen and Lee, 1992).

A GCP is defined as a point on the surface of the earth where both image coordinates (measured in rows and columns) and map coordinates (measured in degrees of latitude and longitude, feet, or metres) can be identified (Benny, 1983; Jensen, 1986). These points are usually well-defined and spatially small such as road intersections, airport runway intersections, bends in rivers, stream junctions, jetties, prominent coastline features, etc. (Davidson, 1986; Richards, 1986). The main aim of selecting GCPs in both target and reference data is to determine the coefficients of a mapping function which relate the image coordinates to a reference system and bring them into conformance with one another. Mathematical functions that are used may be first, second or third order polynomial mappings. Denoting the reference coordinates by (x,y) and the image coordinates (column and row) by (X,Y), the mapping functions are given by (Ford and Zanelli, 1985):

$$X = f_1(x,y) = \sum_{j=0}^{q} \sum_{k=0}^{q-j} a_{jk} x^j y^k$$

$$Y = f_1(x,y) = \sum_{j=0}^{q} \sum_{k=0}^{q-j} b_{jk} x^j y^k$$

where the q is the order of the polynomial mapping and a_{jk} and b_{jk} are coefficients of transformation. When applying first degree polynomial mapping, the equations are expressed as follows:

$$X = a_0 + a_1 x + a_2 y,$$

$$Y = b_0 + b_1 y + b_2 x,$$

The complexity or order of these polynomials depends on the geometry of the image and the type of map projection to be used (Janssen and van der Wel, 1994). The minimum number of GCPs required for the affine transformation (mapping parallelograms to parallelograms) is three; a minimum of six is required for second order polynomial analysis and third order mapping requires a minimum of ten GCPs (Dymond, 1991; Richards, 1986). The accuracy of the transformation is assigned by the root mean square (RMS) error, which is given by the equation (Jensen, 1986):

$$RMS_{error} = [(X-x)^2 + (Y-y)^2]^{1/2}$$

In the transformation process an image resampling algorithm is applied to the input data to produce new pixel values in a new grid. This involves interpolation of an intensity value from its position in the original input image to the appropriate location in the output image. This also means a transformation of the feature space (Janssen and van der Wel, 1994). Methods of interpolation that are in common use include *nearest-neighbour, bilinear interpolation* and *cubic convolution.*

As the name implies, *nearest-neighbour* resampling simply finds the input-image sample which is closest to the output-pixel location and assigns that intensity value to the output- pixel sample (Bernstein, 1983; Green, 1989; Jensen, 1986). Thus the new pixel value is exactly the same as one of the four spectral values surrounding the interpolated point. This is the preferred technique if the new image is to be classified, as it is simply rearranged in position to give the correct image geometry (Richards, 1986), without changing the original pixel-intensity values. Unlike the nearest-neighbour technique, *bilinear interpolation* employs the four pixels that surround a given point to assign an output pixel value. This new output value is the result of a computation based on the weighted distance to that point. While the technique used in *cubic convolution resampling* is similar to bilinear interpolation, the latter utilizes sixteen surrounding pixels to assign a new output intensity value.

Davidson (1986) pointed out that the geometrically transformed imagery is desirable to: (a) locate common points in different scenes of the same area, (b) perform multi-temporal analysis by overlaying images of the same area taken at different dates, (c) bring adjacent images into register so that they can be mosaicked together, and (d) overlay images of the same area produced by different sensors. This procedure is particularly useful in change detection applications where a difference in ground features of two or more images obtained at different times is of interest.

For any further analyses, it is important that a registered image is labelled with adequate information. The following items are considered desirable: source of reference GCPs, number of GCPs, type of transformation, RMS error or standard deviation, and resampling method (Janssen and van der Wel, 1994).

4.1.3 Registration Procedure

Before performing each stage of the procedure, the images to be used were first subsampled to extract the study area. This task was done using DRAGON *Version*

4.0, which is capable of displaying a 1024 x 700 image. The subsampled-DRAGON format images were then converted to the IDRISI format using the DRAG2GEN module. Subsequent tasks were then undertaken by employing the GIS- IDRISI software package.

4.1.3.1 GCPs identification

The initial step of registration procedure, and the most important, was identifying GCPs on the image to be registered and accurately recording their coordinates from the appropriate positions on the reference map. The reference used to perform this task was the digital cadastral data, purchased from the Department of Survey and Land Information, that had been rectified to the New Zealand Metric System (NZMS). These numeric data were first converted to a line coverage using the GENERATE module of PC ARC/INFO. By displaying this line coverage in the ARCEDIT program, accurate coordinates (eastings and northings) of selected GCPs could be recorded.

To assist the identification and selection of GCPs' locations, several clearly recognisable features on the image were selected. These included road intersections, forest edges, cadastral boundaries, and farm boundaries. Seventeen GCPs were selected (Table 4.1). GCPs numbers 1 to 16 were used for registering the scanned aerial photographs, while for the multispectral SPOT imagery, GCPs

Table 4.1 GCPs coordinates used for image registration (based on the NZMS reference system).

GCPs	Grid reference	GCPs coordinates	
		Easting	Northing
1	R25 897434	2689736	6043487
2	R25 891445	2689160	6044511
3	S25 919429	2691946	6042934
4	S25 931425	2693170	6042581
5	R25 899415	2689920	6041580
6	S25 906454	2690659	6045450
7	S25 923420	2692328	6042014
8	S25 922453	2692240	6045377
9	S25 907425	2690728	6042531
10	R25 896457	2689674	6045785
11	S25 914438	2691405	6043821
12	S25 935437	2693592	6043738
13	R25 889425	2688903	6042540
14	S25 913413	2691340	6041345
15	S25 929457	2692930	6045753
16	S25 915449	2691552	6044967
17	S25 938418	2693830	6041864

numbers 9 to 17 were employed. As far as practicable, these points were distributed evenly over the study area (Figure 4.1). Care was also given to the location and the distribution of the control points around the edges. This was to ensure that the selected mapping polynomials were well-behaved over the image being registered (as fully discussed by Richards, 1986).

Once the control points have been selected on the reference, it is necessary to locate them in the corresponding positions on the image, and to record the image coordinates. To identify the approximate location of a particular control point on the image, the pixels around this point were magnified and displayed on the screen. The cross cursor was then used to indicate the GCPs positions. As the NZMS coordinates increase to the north and east, then the image coordinates were recorded with ascending order from the lower left-hand corner, in lieu of using the number of columns and rows which start from the top left-hand corner of the image. The latter can also be applied, but the output needs to be flipped or transposed to give the correct positions.

Having created these corresponding coordinates (Appendix 1), the IDRISI-RESAMPLE module was then executed. This gave an indication of the accuracy of the transformation, and a first guess of the improvement needed to minimise the error. This could be identified from the total (X and Y) of root mean square (RMS) error and the residuals for each control point. Some time was spent repeating the operation until an acceptable total RMS error was obtained.

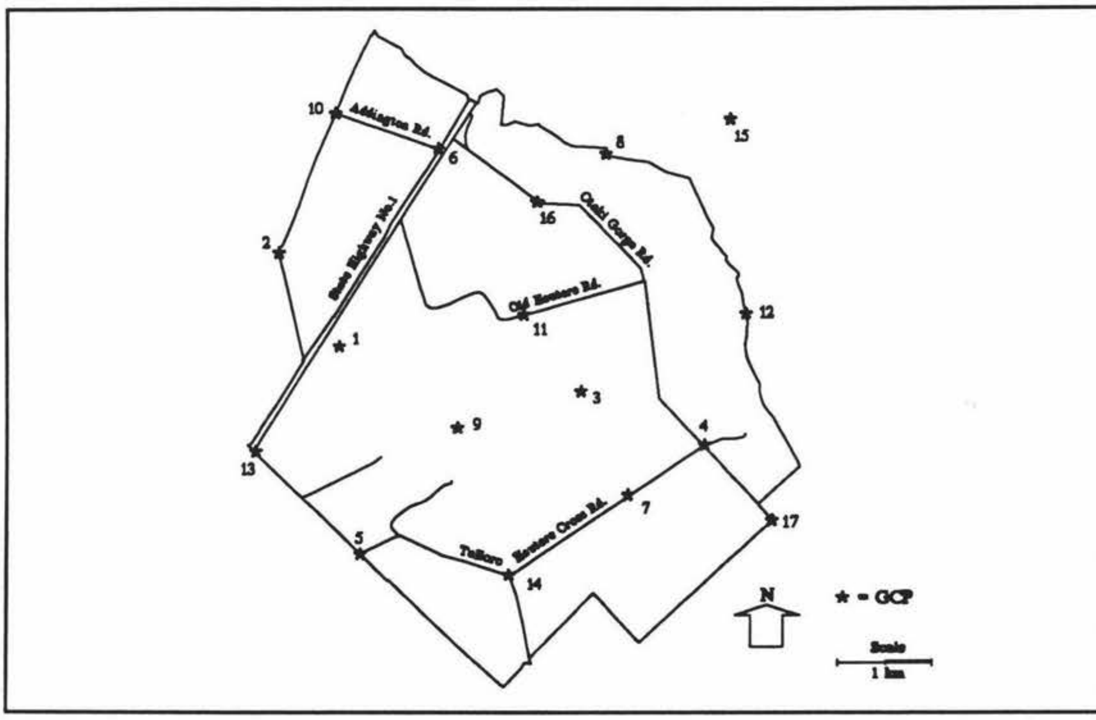


Figure 4.1 Distribution of the GCPs used in the study area

4.1.3.2 Improving the fit

This procedure entailed three different steps: (a) creating vector lines, (b) vector overlaying, and (c) identifying the magnitude and direction of the positional GCPs errors.

The reference-rectified cadastral data were used for vector overlaying. These line coverages (in PC ARC/INFO) were first converted to intermediate formats using UNGEN module. These formats were then imported to IDRISI using ARCIDRIS function. This program is capable of generating line and point vectors compatible with IDRISI which could be either displayed using the vector PLOT function or overlaid onto the raster registered images using the COLOR module. The latter enabled identification of the magnitude and direction of positional GCPs displacements.

To identify such displacements, the pixels around each control point, were magnified. Changes were then made to the coordinates in the corresponding files on the basis of the magnitude and direction of the errors. Each of these trials was followed by executing the registration algorithm from the RESAMPLE module. This procedure was repeated, with new transformation equations being determined every time, until a total RMS error of 0.50 pixel (5 metres) or better was achieved. This level of error is considered acceptable for use in the change detection analysis and is better than that obtained by Lo and Shipman (1990) who reported a total RMS error of slightly over one pixel to perform change detection using digitized aerial photographs.

The coefficients of the linear transformation, total RMS errors, and the residuals are presented in Table 4.2. These figures indicate that even though the three total RMS errors are less than 0.50 pixel, the individual positional errors of several GCPs are still somewhat greater than the threshold. This may be a result of either local distortions or mistakes in GCP interpretation. Unfortunately, this program does not provide information on the direction of each control point displacement,

Table 4.2 Accuracy of image registration by reference to the NZMS using first order polynomial mapping.

Parameter	Imagery		
	Aerial photograph (1968)	Aerial photograph (1993)	SPOT XS mode (1990)
GCPs and Residuals	1. 0.51771 2. 0.66517 3. 0.37548 4. 0.60236 5. 0.50736 6. 0.18546 7. 0.16662 8. 0.18049 9. 0.42399 10. 0.61933 11. 0.55582 12. 0.27701 13. 0.49550 14. 0.14741 15. 0.66388 16. 0.41567	1. 0.55843 2. 0.34638 3. 0.25246 4. 0.63384 5. 0.16803 6. 0.38966 7. 0.60884 8. 0.28463 9. 0.29309 10. 0.35915 11. 0.59011 12. 0.20306 13. 0.42064 14. 0.30516 15. 0.39335 16. 0.29914	9. 0.49066 10. 0.24262 11. 0.37696 12. 0.59910 13. 0.45451 14. 0.21915 15. 0.62215 16. 0.26998 17. 0.25523
RMS Error	0.46072	0.40679	0.41904
Coefficient			
X b0	2688979.93886	2689113.43708	2689056.95001
b1	0.08364	0.10845	0.04867
b2	-0.02459	-0.04619	-0.01116
Y b0	6040199.30262	6040204.33824	6040353.83454
b1	0.02396	0.04635	0.01125
b2	0.08380	0.10776	0.04870

Note: Formula being employed is the back transformation, i.e., new to old (refer to the IDRISI Program module).

as commonly indicated by positive or negative signs to the residuals for X (easting) or Y (northing) axis. Thus the figures presented here illustrate only the magnitude of GCPs errors.

4.1.3.3 Resampling procedure

As the primary concern was change detection analysis, the resampling procedure applied was mainly intended to:

a. Geocode the area of interest by assigning the maximum and minimum coordinates (for easting and northing); hence, features in the imagery are positioned on the NZMS geodetic system.

b. Bring all the data sets into conformance with one another using the same georeferencing scheme.

c. Rescale all the data sets to the same grid cell size.

With reference to the NZMS, the sub-scene of the study area was positioned within the coordinates 2688825 to 2694225 (easting) and 6040225 to 6046625 (northing). Thus the rectangular area comprised 5400 x 6400 square metres. To obtain a grid cell size of 10 x 10 metres, the rectangle area was filled with the corresponding rescaled images which consist of 540 columns and 640 rows. Accordingly, all the registered images had the same cell size, though each of the original images possessed a different ground resolution. Therefore, the SPOT XS imagery, having 20 metres ground resolution, was rescaled to a new cell size of 10 metres. Rescaling was also necessary for both the scanned aerial photographs which had resolutions of 14.10 and 10.58 metres, for the 1968 and 1993 imagery, respectively. This procedure enabled the images to be aligned geometrically to enable multi-temporal analysis.

As the registered images would be subjected to a classification procedure, the *nearest neighbour resampling* technique was employed. This technique simply places the pixels in the new grid with the intensity value of pixel which is closest to them. Hence, the intensity values are not changed by the transformation process.

Given the flat nature of topography in the study area, a first-degree affine transformation was performed. This type of mapping function is often sufficient for satellite images (Janssen and van der Wel, 1994), in which most of the geometric distortions have already been removed (Davidson, 1986).

The best fit vector overlaying produced from this registration procedure is depicted in Figure 4.2. These output-registered images were used for image classification (Chapter V), and further for change detection analysis (Chapter VI).

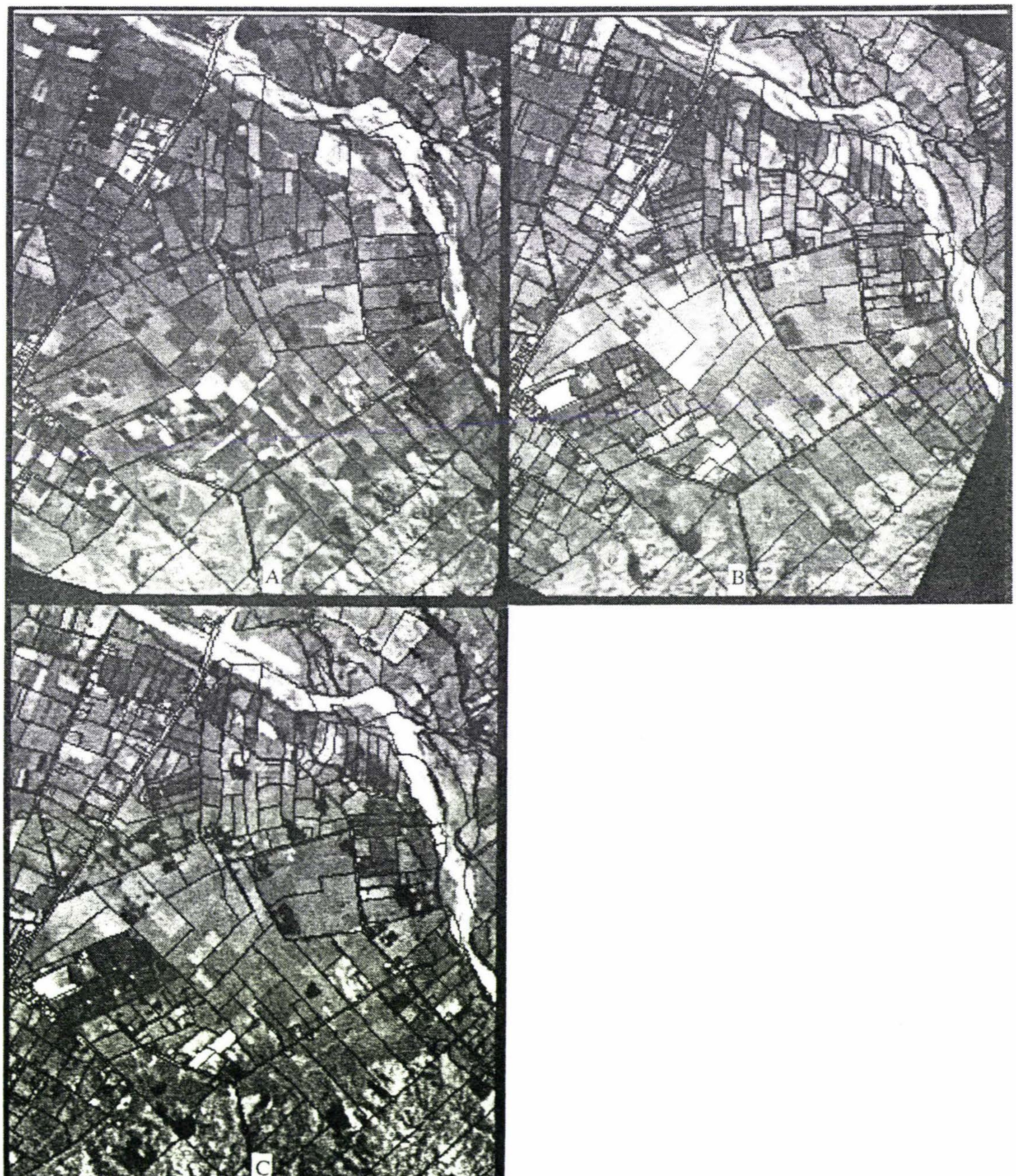


Figure 4.2 Vector-cadastral boundaries overlaid on the 1968 (A) and 1993 (B) scanned aerial photographs, and on the 1990 SPOT XS imagery (band 1) (C), showing the accuracy of registration.

4.2 Data integration

4.2.1 Introduction

The information in a remotely sensed image and that in a GIS, is a representation of the world at two very different levels of generalisation (Ehlers *et al.,* 1989). On the one hand, data generated either from the remote sensor or from image classification are essentially in raster format. On the other hand, data input into the GIS, either manually or automatically, are principally in the form of a vector model. Likewise, data sets used for a particular application could have differences in scale, including spatial and spectral resolution, georeferencing scheme, and boundary represented by the data. When multi-temporal information is of interest, the procedure of *data integration*, i.e., the process by which all 'layers' of the digital data are made to conform in format and in geographic reference (Lo and Shipman, 1990), needs to be employed.

Data integration discussed in this section will focus only on the data conversion procedure and its importance for the following analyses.

4.2.2 Examples of data integration

As previously discussed, most of the analysis procedures in this study were undertaken using DRAGON and IDRISI systems which mainly employ a grid-based format. Hence, rasterising algorithms need to be used to convert any available digitised data, which are in vector format, to the raster structured files.

The rasterising module used in this study was POLYGRID, supplied by the PC ARC/INFO package, which is capable of creating grid files of one of the supported formats from the polygon features of an ARC/INFO coverage. POLYGRID supports six optional grid file types: noncompressed ASCII, compressed ASCII, ERDAS 8-bit, ERDAS 16-bit, EPPL7, and GRID Card Image. In this study, the ERDAS 8-bit format was chosen as an intermediate file. The main reason was that the systems used (i.e., DRAGON and IDRISI) were designed to handle images with 8-bit spectral resolution (2^8 = 256 different colours). In addition, IDRISI provides an ERDIDRIS module that can be used to convert data with ERDAS format to the grid-based structure file compatible with IDRISI programs.

The 8-bit data limitation could, however, be a problem when analysing data sets with more than 256 polygons. This occurred with the 1993 cadastral data which comprised 335 polygons (presented in Chapter VI). To overcome this, each data set, before undergoing format conversion, was divided into three coverages. This task was done by using the DISSOLVE module in ARC/INFO to produce new coverages having the same reference coordinate system. The division procedure is discussed below.

Firstly, from the ARCEDIT module, all the polygons in the original ARC/INFO coverage were labelled. Label numbers 1 to 170 were to be extracted as the first coverage, 171 to 335 for the second coverage. Roads were assigned a discrete number (e.g. 555), as they were to be separated as an individual layer. All labels were automatically stored in the polygon attribute tables (PAT).

Secondly, using the TABLES module, the original PAT file was examined. Access to the PAT file was affected by the SELECT function. Following this, three items (i.e. item1, item2, and roads) were then added to the PAT file. To fill these empty items with the relevant labels, the following algorithm was then applied:

```
SELECT COVER.PAT
RESELECT COVER_ID > = 1 AND COVER_ID < = 170
CALCULATE ITEM1 = COVER_ID
SELECT COVER.PAT
RESELECT COVER_ID > = 171 AND COVER_ID < = 335
CALCULATE ITEM2 = COVER_ID
SELECT COVER.PAT
RESELECT COVER_ID = 555
CALCULATE ROADS = COVER_ID
STOP
```

Finally, the coverages to be separated were extracted using DISSOLVE module with each item specified as above. This gave rise to three separate coverages having the same *tics* and *mapextent*. The POLYGRID module was then used to complete the rasterising process.

The POLYGRID module also supports georeferencing of the data being transformed, so that the output grid files do not have to be resampled to ensure that the data sets are in conformance to one another. Hence, in the rasterising process, the parameters entered were the same as those used in the resampling algorithm; i.e.,

- lower left reference coordinate (x,y): 2688825,6040225
- cell size (width, height): 10,10
- grid size (nrows, ncolumns): 640,540.

Such parameters implied that the program would stop at the coordinate of 2694225,6046625. The resultant grid files contained polygon data that could be overlaid on the registered images with confidence in terms of the correspondence.

4.2.3 The importance of data integration

For the purpose of this study, the data integration process discussed above was required to:

a. Create a binary mask image to define the boundary of the study area (Chapter V).

b. Define, from cadastral boundaries, land use/land cover categories which are difficult to differentiate using automatic techniques (Chapter V).

c. Perform land use/land cover change detection analyses for the two dates (Chapter VI).

d. Allow the analysis of subdivision' sizes for each date using a GIS (Chapter VI).

CHAPTER V

THEMATIC INFORMATION EXTRACTION

5.1 Introduction

One of the crucial stages in the application of remote sensing technology is performing thematic information extraction from remote sensor data using classification techniques. The extracted data are normally segmented and grouped into a number of categories specified by the user. These categorised data may then be used to produce thematic maps of the land use/or land cover present in an image.

Certain methods of classification have been devised that can readily be used to classify remotely sensed images. Effective classification of such data, however, depends upon decomposing ground cover types of interest (*information classes*) into sets of *spectral classes* that represent the data in a form suited to the classifier algorithm used (Swain and Davis, 1978). The algorithm may involve the use of either supervised or unsupervised methods or the integration of both techniques. A manual interpretation technique is generally employed to help identifying subtle information classes which are difficult to differentiate using automatic methods.

The intent of this chapter is to show how thematic information extraction from the remotely sensed data was used, and to produce land use/land cover maps of the study area.

5.2 Basic considerations in the classification process

One of the most critical aspects of land use/land cover mapping is the selection of discrimination levels between information classes. It is essential to properly adjust the specified legend's categories to the data available as well as the local condition of the study region. This is particularly aimed at avoiding assigning pixels to non-representative categories or missing meaningful classes (Chuvieco and Vega, 1990). Accordingly, a particular classification scheme is thus required to designate the levels of category discrimination of the study area.

A number of classification schemes have been developed that can readily incorporate land use/land cover data obtained by interpreting remotely sensed data. The U.S. Geological Survey developed a *resource oriented system* (Anderson *et al.*, 1976) which has to date been widely used as the basis for land use/land cover classification of remote sensor data. The system provides a hierarchical framework for the classification of land information and has been carefully devised to accommodate virtually any classification of land use/land cover (Baker *et al.*, 1979). For this study, the land use/land cover legend defined corresponds to level III of the scheme.

In the classification process, whether using a manual or automatic technique, the problem to be solved is commonly inclusive of (Anderson *et al.*, 1972): (a) placing the decision boundary around each category, (b) heterogeneous mixtures of equally significant land cover, and (c) the area size relative to the minimum size mapping unit. When dealing with automatic methods, the complicity of such factors depends upon the nature of the remote sensor data as well as the availability of the systems used to undertake the task. The former may particularly relate to the source of imagery, spatial and spectral resolution, and the number of image bands; while the latter includes access to efficient hardware and software with which the procedure can be performed.

Thomas *et al.* (1987) outlined basic considerations in the classification of remotely sensed images. Some important principles and postulates paraphrased from this scheme include the following:

a. Decide how much heterogeneity can be permitted in each of the physical sampling units, for user defined classes, whilst maintaining the homogeneity of that class over an acceptable area.
b. Outline representative examples of the required types of land cover to the analysis system.
c. Ensure the system accepts these defined areas, and hence, representative statistics, and thereafter present repeatable results.
d. The class defined by the system should be a little less sensitive to the variations of ground cover types on each resolution element.
e. Know whether or not the specified ground cover type can, in fact, be resolved acceptably by the data.

These reveal that it is important to consider the minimum size of objects that can be represented on a map product of a given scale, and the ability of the system to undertake the task. The minimum size threshold is essentially governed by the ability to consistently recognize, classify, and delineate these objects (Jensen *et al.,* 1983).

5.3 Category selection and definition

After considering the objective of the study and following a site inspection, eight classes of relatively static land use/land cover were selected. They included the following: pastoral farming, orchards, market gardens, trees, commercial sites, residential sites, roads, and gravel pits.

a. Pastoral farming

This land use covers the largest part of the plains, consisting of open areas- dominated by browntop *(Agrotis capillaris)*, ryegrass *(Lolium spp.)*, and white clover *(Trifolium repens)* pasture- generally maintained but lacking evidence of recent tillage. The area is mainly utilized for dairying, sheep and cattle. Deer farming and goat farming were also found in several grazed pasture parcels.

b. Orchards

These instances of perennial horticulture include kiwifruit, pipfruit, stonefruit, and citrus. This category also includes several tracts of subdivision that are likely to be developed into orchards but are presently cultivated for cash crops. Trees and shelter belts, occurring within or bordering between orchard properties were also categorized in this class.

c. Market gardens

These differ from the areas of orchards by the type of management that the land receives. The area is intensively cultivated with annual-vegetable crops to provide for the urban markets and gate sales. The crop varieties grown include onion (*Allium cepa*), cabbage (*Brassica oleracea*), lettuce (*Lactuca sativa)*, tomato (*Lycopersicum escelentum*), and the like. Consequently, this category may exist as recently cultivated land at some times of the year.

d. Trees

This category consists of all types of bush and trees occurring within grazed pasture areas. High-stand shelter belts and native trees bordering orchards and pasture, or market gardening and either pasture or orchards were also categorized in this group. Of the dominant plant species in the plains, some have been discussed in Chapter III Section 3.1.3.

f. Residential sites

Houses or blocks of houses exist as small, and widely scattered units throughout the plains. This category also includes the backyards or sections associated with the houses and obviously belonging to one land parcel.

e. Commercial sites

These include relatively large sized buildings which are used for commercial enterprises such as poultry farming, pack houses, garages, sheds, and storage buildings.

g. Roads

This class is inclusive of only main roads (i.e., surface sealed roads). Narrow-spaced farm roads were grouped into the surrounding land use (pastoral farming, orchards, or market gardens). The reason for this is that they are probably not resolvable at the resolution considered, and also they are not included in the cadastral data sets.

h. Gravel pits

These consist of barren land, with exposed rocks, either from the river or surface extraction activities. This class was mostly found in some land parcels lying along the southern edge of the Otaki River.

5.4 Approaches and techniques adopted in the classification process

The procedure of image classification was carried out using a *semi-automated* approach, i.e., combining a digital technique, which is generally based only upon the spectral information, with a visual analysis method. This, therefore, allows information on the interpretation criteria to be incorporated into the classification procedure. Complementing these, a site inspection was also made to make sure of subtle land cover information which was difficult to differentiate. Accordingly, the procedure of thematic information extraction for each image was undertaken with the following steps:

a. Applying an automated classification using a particular classifier.

b. Using the step (a) above with the addition of binary masking of the most obvious features (roads, gravel pits, residential and commercial sites).

c. Performing a post-classification refinement which involved visual interpretation of the remotely sensed data, with the benefit of ground observation.

This three-stage procedure permitted an assessment of accuracies of thematic maps produced from employing only the first step (a), which relies heavily on the spectral information, and from the addition of masking of the most prominent features (steps a + b). The latter is now considered the most widely-used approach applied to the classification of remote sensing imagery. Post-classification sorting which combines the three steps is then carried out to produce final land use/land cover maps of the study area.

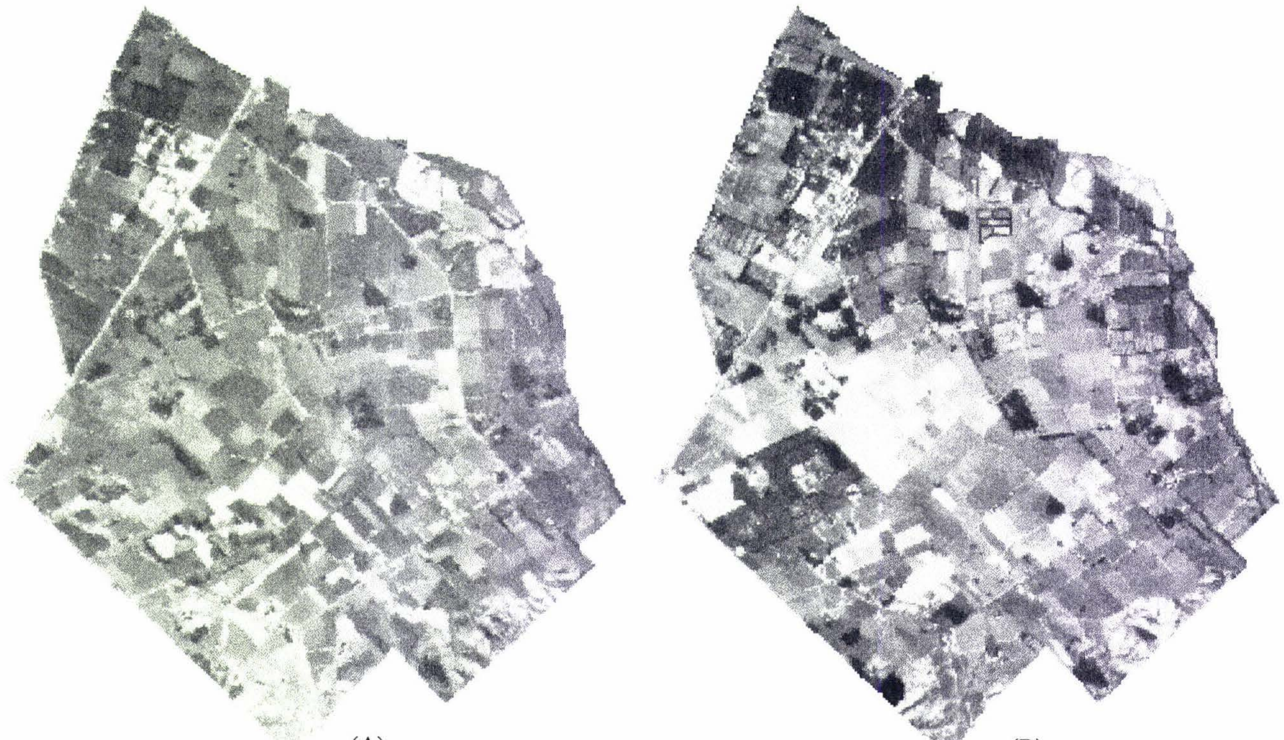
The classification processes were all performed initially using the DRAGON system. The output classified images were then exported to the IDRISI system. This was aimed at enabling GIS-operations to be undertaken, as well as simplifying the input of data sets derived from the ARC/INFO system.

5.4.1 Classification of scanned aerial photographs (SAPs)

As previously stated, the SAPs used were derived from black-and-white photographs so they could be displayed only as a single band. The general appearance of the SAPs used is depicted in Figure 5.1. Using the HISTO function in IDRISI, the histogram values of each of these images were examined. The image statistics are given in Table 5.1.

Classification of SAPs was performed using an unsupervised classifier. With this technique, the system will firstly select a set of starting cluster centres (means) based on the overall distribution of values in the data set; it then classifies each pixel in the image using a minimum distance to mean algorithm (Jensen, 1986; Richards, 1986). In this process, the clusters of target pixels are thus aggregated into the most frequently occurring groups (Thomas *et al.,* 1987).

For this application, the algorithm was firstly asked to produce sixteen clusters (i.e., the maximum number of clusters permitted in the CLUSTERING module of the DRAGON system). Image clustering was then performed using the same image as a mask. This was to avoid the background area (i.e., the area with zero value) being considered as a particular cluster. The procedure of clustering is a repetitive one. The system presents a table- containing class number, pixel count, and means by band- of clustering resulting after each repetition. The clustering



(A) (B)

Figure 5.1 Scanned aerial photographs of the study area taken in 1968 (A) and 1993 (B).

Tabel 5.1 Histogram statistics of the scanned aerial photograph images used for land use/land cover classification.

Statistic parameters	Imagery	
	1968 SAP	1993 SAP
1. Gray levels:		
Lowest	11	7
Highest	207	240
Mean*)	57.68	74.25
2. Standard deviation	23.08	38.36
3. No. pixels		
Range	197	238
Peak (frequency)	5893	2812

*) Based on 201248 values within range.

process is stopped when the algorithm obtains the lowest number of cluster membership changes.

Such a clustering procedure initially produced an output containing clusters that could be associated with the various cover types. As far as practicable, the values of this clustering output were then examined to find their actual identifition on the ground. This task was undertaken by visual interpretation of the aerial photography. Pixels within a given group were then given a symbol (i.e., a discrete value) to indicate that they belong to the same cluster or information class. This procedure was done using the RECODE function in the DRAGON system. The results of this clustering process are given in Figure 5.2.

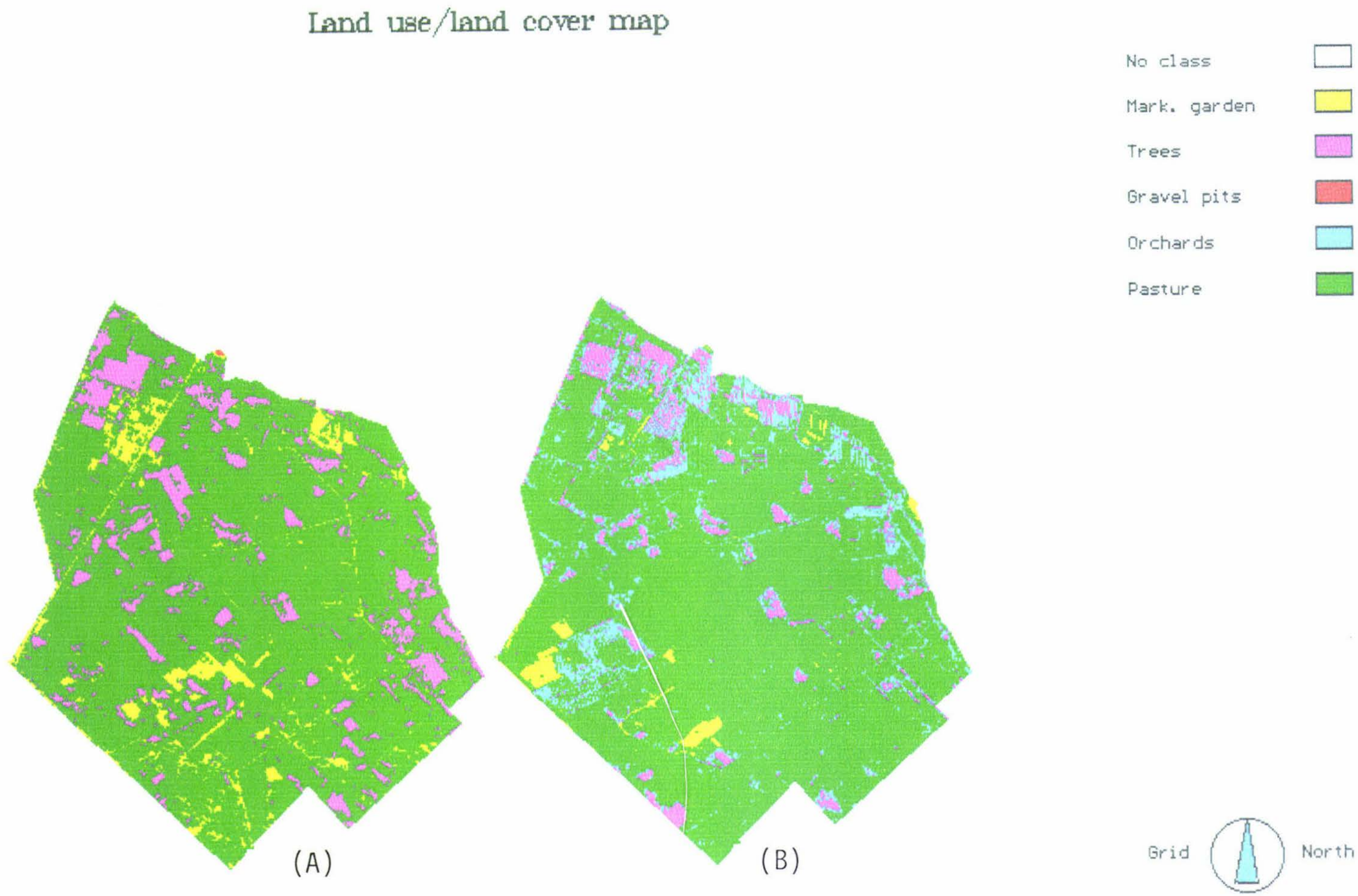
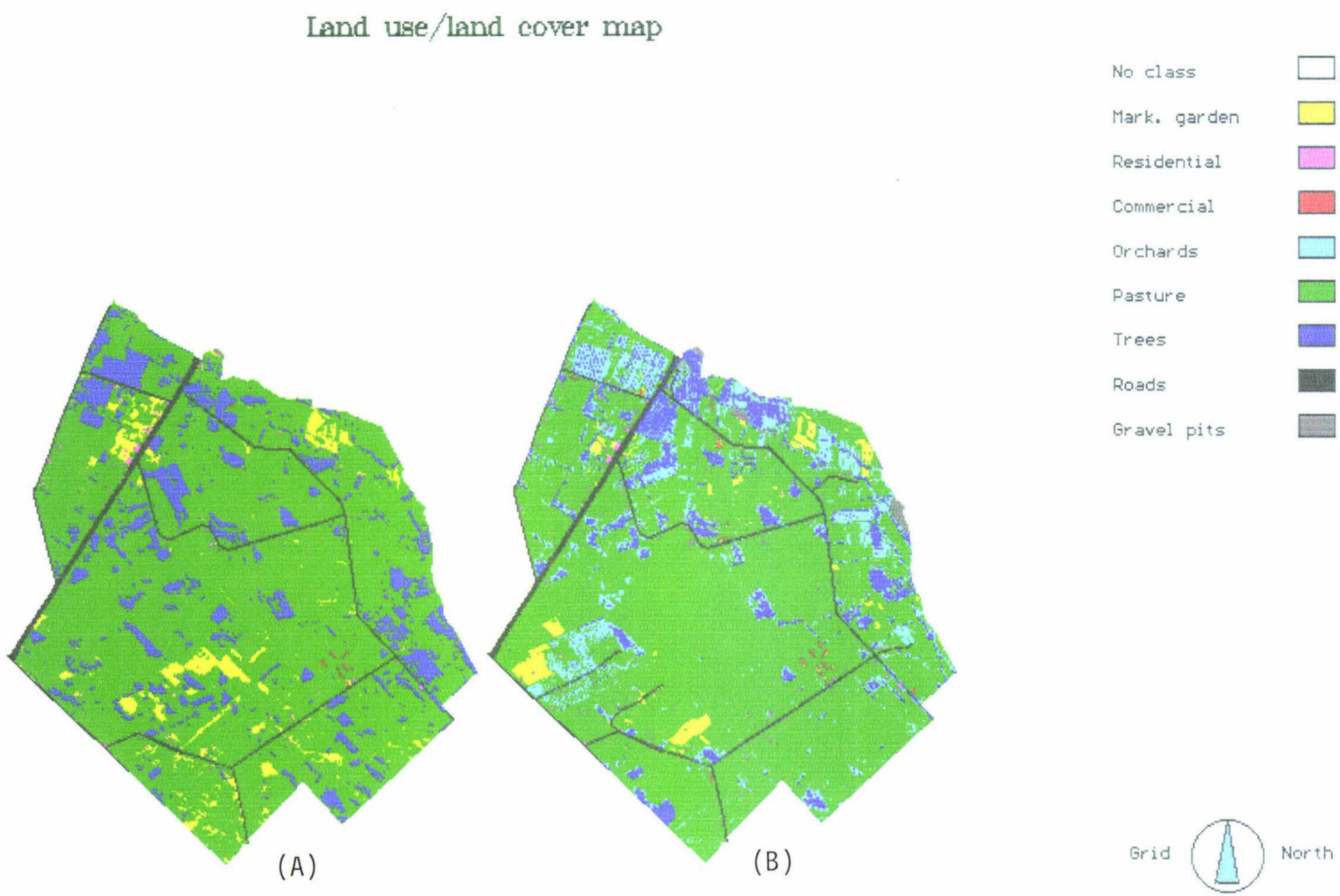


Figure 5.2 Land use/land cover maps of the study area for 1968 (A) and 1993 (B) derived from SAPs by applying an unsupervised classifier.

The next stage involved a technique of image classification which applied an automated approach and image masking. This required the creation of binary masks of features considered to have the most prominent spatial and spectral characteristics, but in some cases, difficult to differentiate automatically. Such features included residential and commercial sites, roads, and gravel pits. The procedure of creating the binary mask is elaborated in Section 5.4.3.

As in the previous stage, the initial procedure produced sixteen clusters, but the areas of features mentioned above were ignored in the clustering process. This was done by first applying binary masking which involved a GIS-based OVERLAY module in the IDRISI system. The sub-scene, which suppressed unwanted features, was then used as an image mask in the clustering process. Such a procedure required the images to be subjected to a format conversion- from DRAGON to IDRISI using ROSETTA-DRAG2GEN module and *vice-versa* employing FCONVERT module. Both modules are facilitated in the DRAGON system.

Once the clustering was completed, the clusters of the output was combined with the binary masks for residential and commercial sites, roads, and gravel pits to create the final map. The results are presented in Figure 5.3.



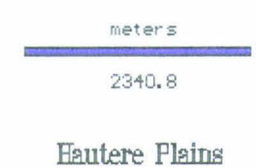


Figure 5.3 Land use/land cover maps of the study area for 1968 (A) and 1993 (B) derived from SAPs, applying an unsupervised classifier and binary masking.

5.4.2 Classification of SPOT XS imagery

Prior to image classification each band of SPOT XS data was displayed several times on the IDRISI screen in order to observe the signature response or gray levels on each band and for comparison among bands. The histogram values are presented in Table 5.2.

Classification of SPOT multispectral data was undertaken by employing both unsupervised and supervised classifiers. In the unsupervised approach the image to be classified was first treated in the same manner as the SAPs; i.e., applying an automated technique and then using a binary mask. For the supervised classification, the procedure was initiated by training representative samples, and generating signatures for each category, followed by image classification. Binary masking was then implemented, the same as that applied to the SAPs. Each classifier is discussed in the following two sections.

5.4.2.1 Unsupervised classification

With the multispectral SPOT data, the clustering algorithm was instructed to include all three bands and to determine sixteen clusters. Thus the image pixels were evaluated in three spectral dimensions to produce cluster map. Again, once the clustering was completed, clusters on the output were then aggregated into the number of categories specified. Due to the unavailability of a 1990 aerial

Tabel 5.2 Histogram statistics of the 1990-SPOT XS data used for land use/land cover classification.

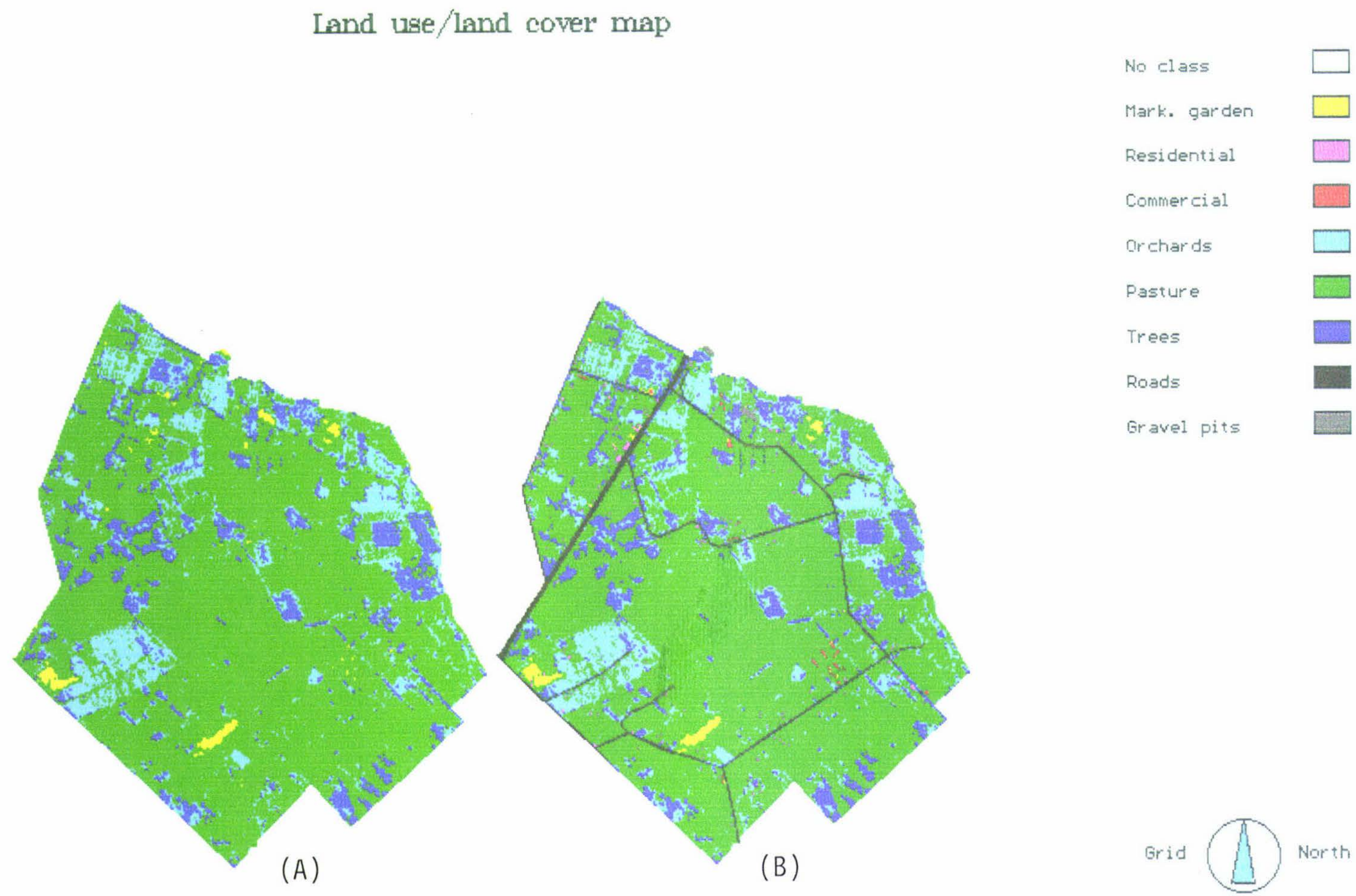
Statistic parameters	Image band		
	1 (0.50-0.59 μm)	2 (0.61-0.68 μm)	3 (0.79-0.89 μm)
1. Gray levels:			
Lowest	28	15	15
Highest	156	108	142
Mean*)	45.90	28.34	84.68
2. Standard deviation	5.83	5.98	13.97
3. No. pixels			
Range	129	94	128
Peak (frequency)	23089	17804	5872

*) Based on 201248 values within range.

photograph, the process of aggregation was made on the basis of the information derived from a visual interpretation of the colour composite imagery displayed on the DRAGON screen (Figure 5.4). This was created by displaying bands 1, 2, and 3 of the XS imagery with the blue, green, and red guns, respectively.

Masks of roads, residential, and commercial features were taken from the 1993 image masks with the assumption that no change to these features had occurred within that period. By applying such a technique as detailed in Section 5.4.3.2, a binary mask of 1990-gravel pits was created employing band 2 (red) of the XS data. The reason for choosing this band was that in the spectral range of band 2 (0.61-0.68 μm wavelength), bare ground is more recognizable and easier to discriminate from other cover types. The results of classification process are presented in Figure 5.5.

Figure 5.4 Colour composite for bands 1, 2, and 3 of SPOT XS imagery showing the study area and environs.



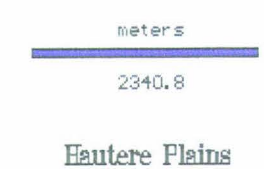


Figure 5.5 Land use/land cover maps of the study area derived from 1990 SPOT XS imagery by applying an unsupervised classifier (A), and with the addition of binary masking (B).

5.4.2.2 Supervised classification

Successful utilization of supervised classification procedures depends largely upon having determined beforehand the spectral structure of the data (Moreton, 1984). This is usually obtained by selecting representative sample sites of known cover type, called *training areas*, prior to the classification stage. The supervised technique involves three major steps: (a) selection of training areas, (b) generation of the spectral signatures for the training areas, and (c) classification of the sub-scene on the basis of signatures generated.

5.4.2.2.1 Selection of training areas

For this application, training areas were based on knowledge derived from the analyses of the displayed XS-colour composite image (Figure 5.4). Such a procedure has often been used for visual analysis of remotely sensed imagery (Astaras and Silleos, 1984; Nicolaos, 1988). On the DRAGON monitor, sample sites were selected to define training areas for each class by means of a moving cursor. Each training area selected was composed of several small circular areas, the number depending on the size and distribution of each land use/land cover type.

Of the eight classes specified, only six were considered: pasture, orchards, market gardens, trees, gravel pits, and roads. Due to the high spectral variability, however, both pasture and market garden classes were subdivided into two subclasses- *dry pasture* and *green pasture* for the pasture category, and *market garden-1* and *market garden-2* for market gardens. Accordingly, with this modification a total of eight classes were examined.

5.4.2.2.2 Signature generation and data evaluation

Each representative sample of features, as defined by the training area, contained the pixels' brightness values for each band of the SPOT XS sub-scene employed. From these values, training statistics involving minimum, maximum, means, variances, and covariances were then calculated to develop spectral signatures representing targets from which the training area was drawn. These parameters enabled the spectral homogeneity of each target to be examined.

Table 5.3 depicts univariate and multivariate statistics of the representative sample chosen. The figures are usually used to indicate the degree of spectral separability within class and between classes (Jensen, 1986). Within-class variability is characterized by such parameters as mean, standard deviation, variance, minimum, and maximum. The spectral range of the XS data used varies from 16 (minimum for trees in band 2) to 117 (maximum for orchards in band 3). The values relate, respectively, to high chlorophyll absorption in red wavelength, and high reflectance of chlorophyll substances in the near infrared band (Hoffer, 1978;

Table 5.3 Univariate and multivariate training statistics for the eight targets selected in the study area using SPOT XS data.

Target :	**Dry pasture**			**Green pasture**		
Image band:	1	2	3	1	2	3
			Univariate statistics			
Mean	56	42	73	45	26	97
Std.dev.	2.67	4.25	5.04	1.01	1.68	9.05
Variance	7.15	18.03	25.40	1.03	2.83	81.87
Minimum	50	34	64	43	23	79
Maximum	61	51	81	48	30	114
			Variance-covariance matrix			
1	7.15			1.03		
2	9.93	18.03		1.40	2.83	
3	-3.20	-11.45	25.40	-7.10	-12.39	81.87

Target:	**Orchards**			**Trees**		
Image band:	1	2	3	1	2	3
			Univariate statistics			
Mean	39	22	91	34	19	76
Std.dev.	1.24	1.52	9.26	1.73	1.39	9.18
Variance	1.54	2.30	85.66	3.01	1.93	84.30
Minimum	36	20	70	29	16	52
Maximum	42	27	117	38	23	94
			Variance-covariance matrix			
1	1.54			3.01		
2	1.05	2.30		1.61	1.93	
3	2.76	-4.57	85.66	10.92	5.03	84.30

Target:	**Market garden-1**			**Market garden-2**		
Image band:	1	2	3	1	2	3
			Univariate statistics			
Mean	59	46	59	85	72	69
Std.dev.	7.26	9.99	13.78	7.12	7.07	4.25
Variance	52.69	99.88	189.78	50.72	50.00	18.10
Minimum	48	29	41	64	44	62
Maximum	75	60	102	97	83	79
			Variance-covariance matrix			
1	52.69			50.72		
2	69.97	99.88		45.60	50.00	
3	-62.30	-94.68	189.78	25.54	20.87	18.10

Table 5.3 Continued.

Target :	Gravel pits			Roads		
Image band:	1	2	3	1	2	3
			Univariate statistics			
Mean	84	64	55	57	38	56
Std.dev.	11.95	10.95	9.53	4.07	3.76	8.74
Variance	142.72	119.85	90.86	16.60	14.17	76.44
Minimum	56	40	35	48	32	45
Maximum	97	78	87	66	48	79
			Variance-covariance matrix			
1	142.73			16.60		
2	128.67	119.85		12.09	14.17	
3	5.76	4.99	90.86	-12.01	-6.54	76.44

Lusch, 1989). The high values of standard deviation and variance indices are an indication of spectral heterogeneity (Adeniyi, 1985).

Similarly, covariance indices- i.e., the tendency of spectral values to vary similarly in two bands (Lillesand and Kiefer, 1987)- are relatively high. The high values of covariance (whether negative or positive) coincident with high variance indices in such a spectral range, describe a case of spectral overlap between the targets. Negative covariance indicates that, in the bispectral plot, the distribution of category values slants down to the right, while positive values mean the data slant upward to the right (Lillesand and Kiefer, 1987). Three-dimensional plots of the representative samples chosen, which illustrate the degree of spectral overlap between categories, are presented in Figure 5.6.

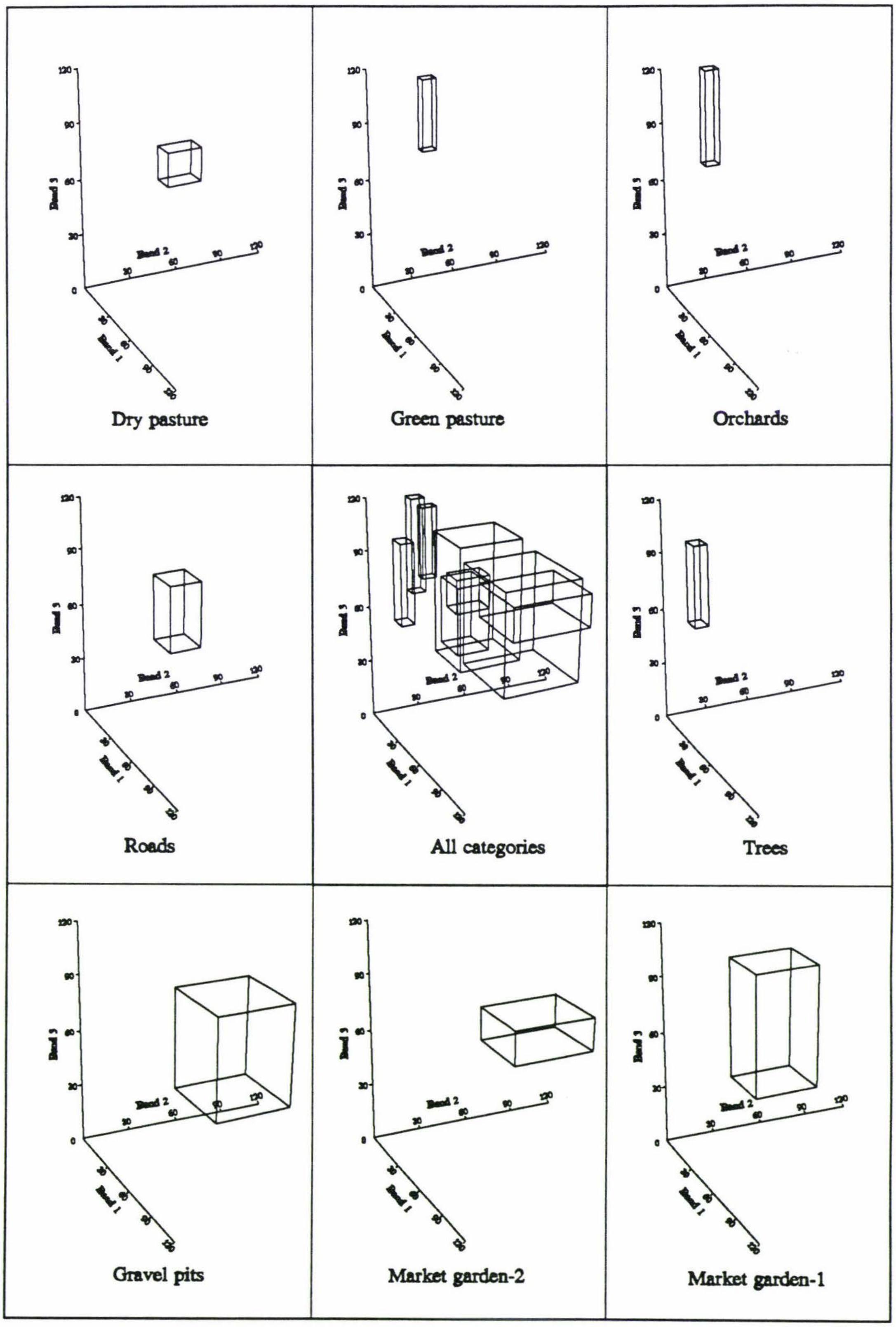


Figure 5.6 Three-dimensional spectral plots of the representative samples used as training signatures for image classification.

Degree of class separability, and thus spectral overlap between categories, could be also estimated by utilizing the percentages of coefficient of variation (Adeniyi, 1985; Satterwhite *et al.*, 1984), or confusion and divergence matrices (Adeniyi, 1985). For this study, confusion matrices were also used as an indicator of the degree of class separability. These measures are presented in the form of an accuracy assessment of the classified images (Section 5.5).

5.4.2.2.3 Sub-scene classification

At this stage, a "minimum distance to mean" classifier was chosen. The reason is that this decision rule is computationally simple (Jensen, 1986), and it could, if operated properly, produce maps with classification accuracies comparable to other more computationally intensive algorithms (Hixson *et al.*, 1980). In addition, this technique of classification does not make use of the covariance information (Richards, 1986), as the trained data of several cover types exhibit a considerably high covariance index. With this classifier, the program only considers the class means; classification is then performed by placing a pixel in the class of the nearest mean (Jensen, 1986; Lillesand and Kiefer, 1987; Richards, 1986). The distance is, in principle, calculated using *Euclidian distance* (Richards, 1993) based on the *pythagorean theorem* (Swain and Davis, 1978).

With the MDM module, image classification was carried out by utilizing the three spectral bands and employing the signature file generated from the training areas.

The RECODE module was then used to aggregate *dry pasture* and *green pasture* to pasture, and *market garden-1* and *market garden-2* to market garden category. Like the techniques applied previously, this stage also involved the use of an automated technique first, and was followed by applying image masking. The outputs are given in Figure 5.7.

It should be noted that the process of training and classification was an iterative procedure. It involved training, preliminary classification, and comparison with the information available, then training, classification, and comparison again several times before completing the land use/land cover map.

5.4.3 Binary masking

A binary mask, sometimes called a "boolean image" is referred to as an image which is used to mask out particular areas on a complex image. This image has only two values, usually one and zero; the 1's represent the areas that meet the stated criteria, while the 0's represent areas that do not meet the criteria (McKendy *et al.,* 1992). Thus, by applying a masking operation, particular features can be isolated from a sub-scene. This procedure together with several GIS-based operations will be demonstrated with the creation of masks for residential and commercial sites, roads and gravel pits.

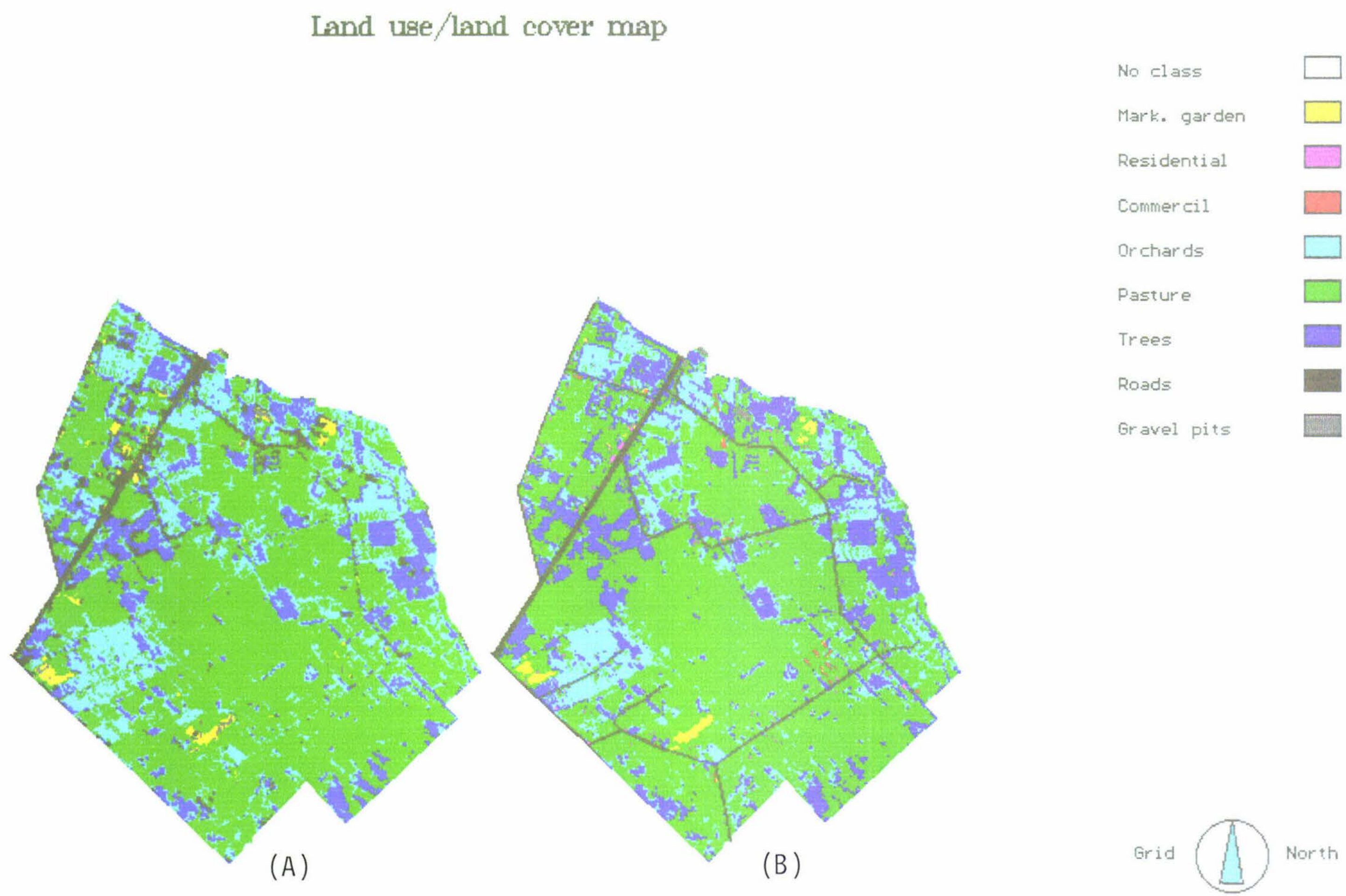


Figure 5.7 Land use/land cover maps of the study area produced from 1990 SPOT XS imagery, by applying a supervised classifier (A), and with the addition of binary masking (B).

5.4.3.1 Mask of residential and commercial sites

As noted earlier, on the plains most residential sites exist as small, widely scattered units, though some blocks of houses were also found especially along the main roads. In the classification process, some difficulties were encountered when an attempt is made to identify houses or blocks of houses in such a spatial setting. Moreover, the area around the residential units tends to have a similar spectral signature to other cover types, particularly pasture grasses.

To cope with these problems, it was possible to make use of the cadastral information to provide a mask of residential sections. Using this technique, only houses residing within polygons less than 0.5 ha in size were considered. However, with those occurring within the larger sized subdivisions, such an approach was not possible. The following procedure describes the techniques used to create the mask of residential and commercial sites.

a. The 1968 imagery was firstly displayed on the IDRISI screen using COLOUR A option. It enabled the coordinates (rows and columns) or the range of coordinates of the features considered to be identified with the aid of a moving cursor. The availability of a 1:20,000 enlargement of the airphoto was very helpful in enabling a visually recognition of these targets. These coordinates were then recorded.

b. The UPDATE module in IDRISI was then employed to assign all the pixels, whose coordinates had been recorded, a new value of zero.

c. Using the RECLASS function, the remaining old values ranging from 1 to 256 were assigned to a new value of one. This gave rise to an image mask with only values of zero (targets) and one (others).

To obtain 1993 residential and commercial masks, the image mask produced from (c) was multiplied by the 1993 imagery using the OVERLAY function. This assigned all the corresponding pixels with 1968 targets to zero, while the remaining pixels of the 1993 sub-scene retained the original values. A similar procedure was then undertaken using the steps (a) to (c) above to identify new targets, and assign them to zero. As would be expected, the procedure adopted here assumed that there were no changes from the existing 1968-residential and commercial sites to other land uses, from 1968 to 1993; instead, it was assumed that these areas increased over that period of time.

5.4.3.2 Mask of roads and gravel pits.

Masks of roads and gravel pits were made with reference to the available cadastral information. The area of interest was firstly extracted from the cadastral polygons employing the technique explained in Chapter III Section 3.2.2. Using such a technique, there was no difficulty separating roads because the property boundaries match precisely both sides of the roads; however, it was not possible to separate gravel pits with exact boundaries. This was dealt with by considering all the subdivisions containing gravel pits and extracting them as a separate layer. The following procedure was then followed.

Binary masks, for each date, were first created from the rasterized polygons (the rasterizing algorithm is discussed in Chapter III Section 3.2.2). This image mask was then multiplied by the original sub-scene to produce the portion of the sub-scene which contains gravel pits. This image was used as a mask to classify the sub-scene of the study area using the CLUSTERING module in DRAGON. Finally, the clusters produced were aggregated for gravel pits only, while other features were assigned to zero. Any scattered pixels which were considered not to be the target were assigned to zero using UPDATE function in IDRISI system. A similar procedure was also applied to the "road layer" to avoid suppressing a number of trees occurring along the road edges.

Using the technique described above, the road layers extracted appeared to be too wide, however. This was because it included the space of unutilized areas between the surface sealed road and the property boundaries.

5.5 Accuracy assessment

There are two major types of accuracy assessment procedures: *non-site-specific* and *site-specific* (Congalton, 1991; Mead and Szajgin, 1982). In the *non-site-*

specific approach, the accuracy is expressed as the similarity between the total areas of land cover types, while ignoring the locational accuracy. The *site-specific* method, however, takes into account the locational accuracy of the data. That is, two spatially defined data sets (one being "ground truth") are compared for the amount of agreement (Mead and Szajgin, 1982). This approach is the most common way of representing the classification accuracy of remotely sensed data and takes the form of an error matrix. Thus, in this study the latter approach was adopted.

5.5.1 Spatial sampling and sample size

The initial concern in performing map accuracy assessment is the selection of a sampling method that would give reliable results. Consideration must be given in the method of constructing a spatial sample and in the determination of sample size. Spatial sampling employed for this application was the *stratified random sample*, i.e., a compromise between random sampling and systematic sampling. This sampling method has good geographical coverage (because of the regular structure of the spatial strata) and is reasonably unbiased because of the random selection of point coordinates with the cell (Eastman *et al*, 1993).

The equation for the approximate sample size, N, is given as:

$$N = Z^2pq/E^2,$$

where Z is the standard score for a specific confidence level, p is the expected percent accuracy, q = 1-p, and E is the confidence interval (Eastman *et al*, 1993; Fitzpatrick-Lins, 1981). In principle, an expected accuracy of 85% (or 15% error) or more is recommended for land use/land cover mapping using remotely sensed data (Anderson *et al*, 1976; Eastman *et al*, 1993). Therefore, the approximate sample size required for a 90% two-sided confidence probability and a confidence interval of 0.04 would be:

$$N = (1.645)^2 x 0.85 x 0.15/(0.04)^2 = 216.$$

Considering the physical setting of the study area, in which residential sites, commercial sites, and trees exist mostly as scattered groups, the accuracy assessment for this study was initiated by generating up to 240 points. Such a sample size has been also used by Rutchey and Vilcheck (1994) for accuracy assessment of a vegetation map using SPOT imagery.

Allocating these points in the study area which has an irregular boundary presented a difficulty, however. This could be dealt with by determining the proportional area of the study region using HISTO function in IDRISI. The area of interest is represented by the proportion of non-zero values in the numeric histogram. For this application, the proportion of the image occupied by the study area is 0.58; thus a total of 240/0.58 = 413 points needed to be allocated in the whole scene to obtain the required number of sample points.

Having applied the stratified random sampling technique using the SAMPLE module in the IDRISI system, it was necessary to identify the number of points actually falling within the study area. This was done by rasterising the vector-point file using the POINTRAS function, and then masking this raster-point image file for only the study region. Again, from the HISTO function, it was determined that 238 sample points fell within the study area. The distribution of sample points, that was applied to all of the maps assessed, is depicted in Figure 5.8.

5.5.2 "Ground truth" generation

Once the number and location of sample points have been determined, the area needs to be visited at each of the locations chosen in the sample. In cases where a large scale image or photograph is available, the condition of these points could be usually determined by visual interpretation (Eastman *et al*, 1993). It is this, the process of establishing "ground truth" data that is considered to be expensive and time consuming.

The procedure of creating a "ground truth" image file for this application involved overlaying the vector-sample points on the original scene of the study area, identifying the category at each point, and assigning the corresponding pixels to the specified values. By overlaying sample points onto the original imagery displayed on the IDRISI screen, it was possible to identify X and Y coordinates

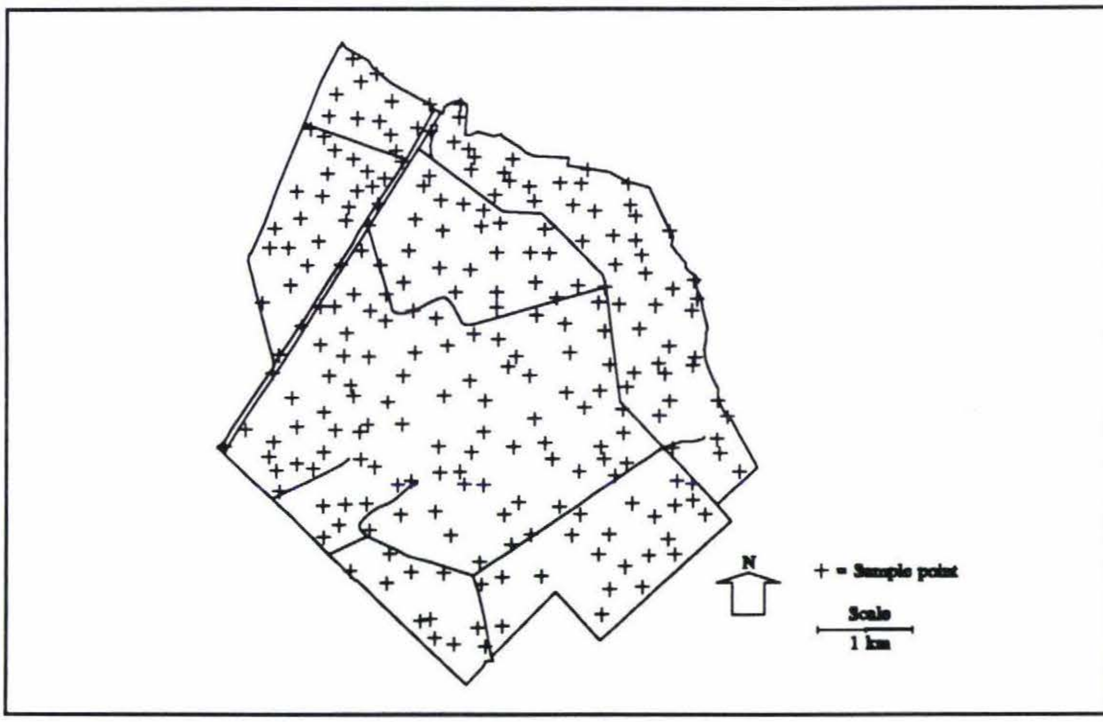


Figure 5.8 Distribution of sample points, in the study area, used for map accuracy assessment (procedure for creating this figure is explained in Appendix 2).

(column and row) of each sample point with the aid of the cursor and zoom operations- these operations were carried out interactively in the COLOUR A module. The land use around each of the sample positions was then carefully examined using the visual interpretation approach, explained in Section 5.6.1, to verify the actual cover type present at a particular test point. Finally, each point of the samples was assigned to a single value according to the land use/land cover type and by reference to the coordinates recorded. This task was done by utilizing the rasterised-sample images, and employing the UPDATE function.

5.5.3 Error matrix

Once the "ground truth" image files have been created, the accuracy assessment algorithm can be implemented to compare the categories generated from the classification of remotely sensed images and the actual ground information. The comparison was made by establishing an "error matrix" which provides information on the accuracy of individual categories as well as the accuracy of the classification as a whole. Thus each error matrix contains *errors of commission, errors of omission,* and *overall accuracy.*

Errors of commission relate to the cases where the algorithm applied has defined categories that do not exist on the ground. This measure, called *user's accuracy* or *reliability*, is indicative of the probability that a pixel classified on the map/image actually represents that category on the ground (Story and Congalton, 1986). On the other hand, errors of omission occur when certain categories that actually exist on the ground are omitted. This accuracy measure is called *producer's accuracy* because the producer of the classification is interested in how well a certain area can be classified (Congalton, 1991). The relationships of such terms could be expressed as follows (Janssen and van der Wel, 1994):

User's accuracy (%) = 100% - error of commission (%)

Producer's accuracy (%) = 100% - error of omission (%).

Overall map accuracy is calculated by dividing the sum of the entries that form the major diagonal (i.e., the number of correct classification) by the total number of samples taken (Rutchey and Vilcheck, 1994; Story and Congalton, 1986).

For the assessment procedure, it was desirable to simplify the descriptions of the image sources and the techniques adopted in this application. Hence, these parameters were presented in an abbreviated form as shown in Table 5.4.

Table 5.4 Description of the techniques employed in the classification.

Image source	*Technique applied*	*Symbol*
1968-SAP	1. Unsupervised	SP68U
	2. Unsupervised+masked	SP68UM
1993-SAP	1. Unsupervised	SP93U
	2. Unsupervised+masked	SP93UM
1990-SPOT XS	1. Unsupervised	XS90U
	2. Unsupervised+masked	XS90UM
	3. Supervised	XS90S
	4. Supervised+masked	XS90SM

The error matrices for the thematic maps generated previously are given in Appendix 3. The figures show that the classification accuracies vary considerably for different image sources and different techniques applied. The use of image masks in the classification algorithms did, in general, improve the overall levels of map accuracies (i.e., producer's, user's and overall accuracies) for all image sources used. In terms of the overall accuracy the increment ranges from 6 to 9 percent.

Classification of the 1968-SAP yielded the highest overall map accuracy, i.e., 82 and 88 percent, respectively, for SP68U and SP68UM (Figure 5.9). It is, however, not necessarily true that it provided good thematic representation for the 1968 classification map. This high value relates to the reduced variability of categories being mapped. As mentioned earlier, at this time, grazed pasture was the dominant land use, with other land uses were found only in a small portion of the area. Many orchards had not yet been planted.

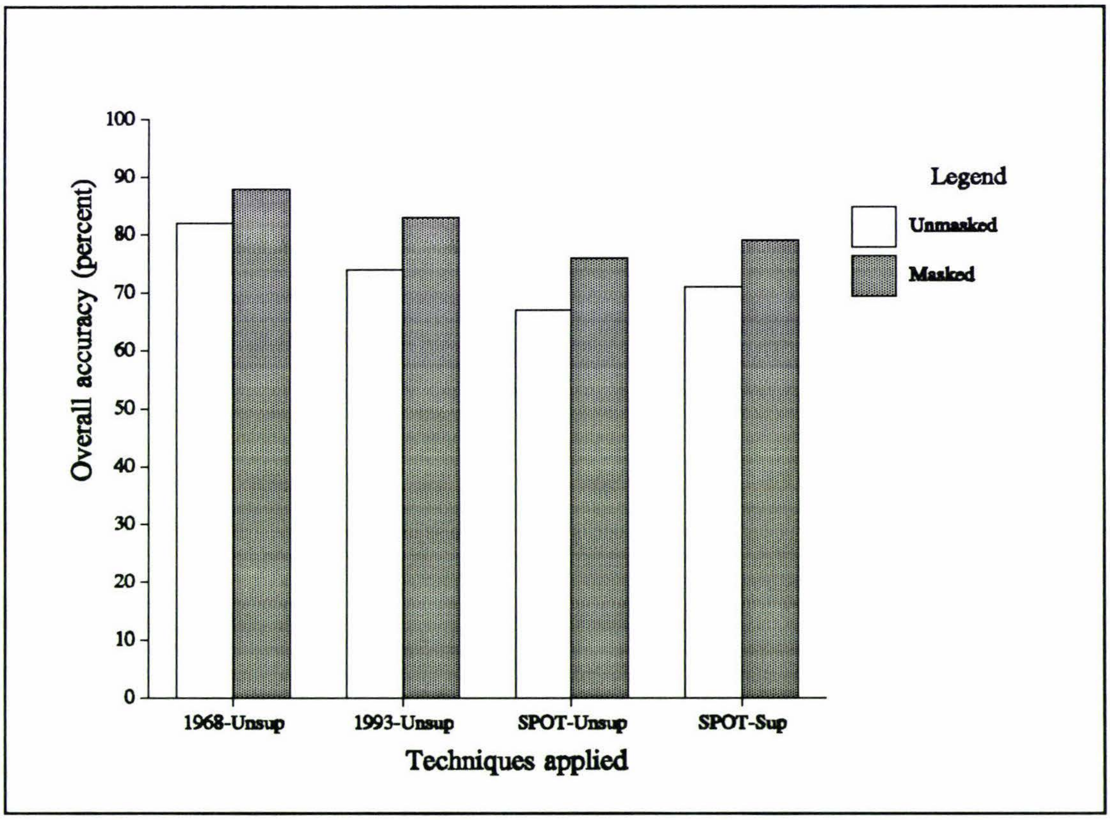


Figure 5.9 Overall accuracy of classification for the eight land use/land cover types according to the technique applied.

In such circumstances, Story and Congalton (1986) suggested using more detailed statements of measures to indicate how the accuracy is distributed across the individual categories. Figures 5.10 and 5.11 provide information on the accuracy of each thematic category being identified. When considering SP68U, 88 percent (producer's accuracy) of market gardening could be identified as such. However, of the areas classified as market gardening, only 29 percent were actually market gardening on the ground; 71 percent of the areas identified as market gardens were actually commercial sites, residential sites, roads, and pasture. Similarly, the

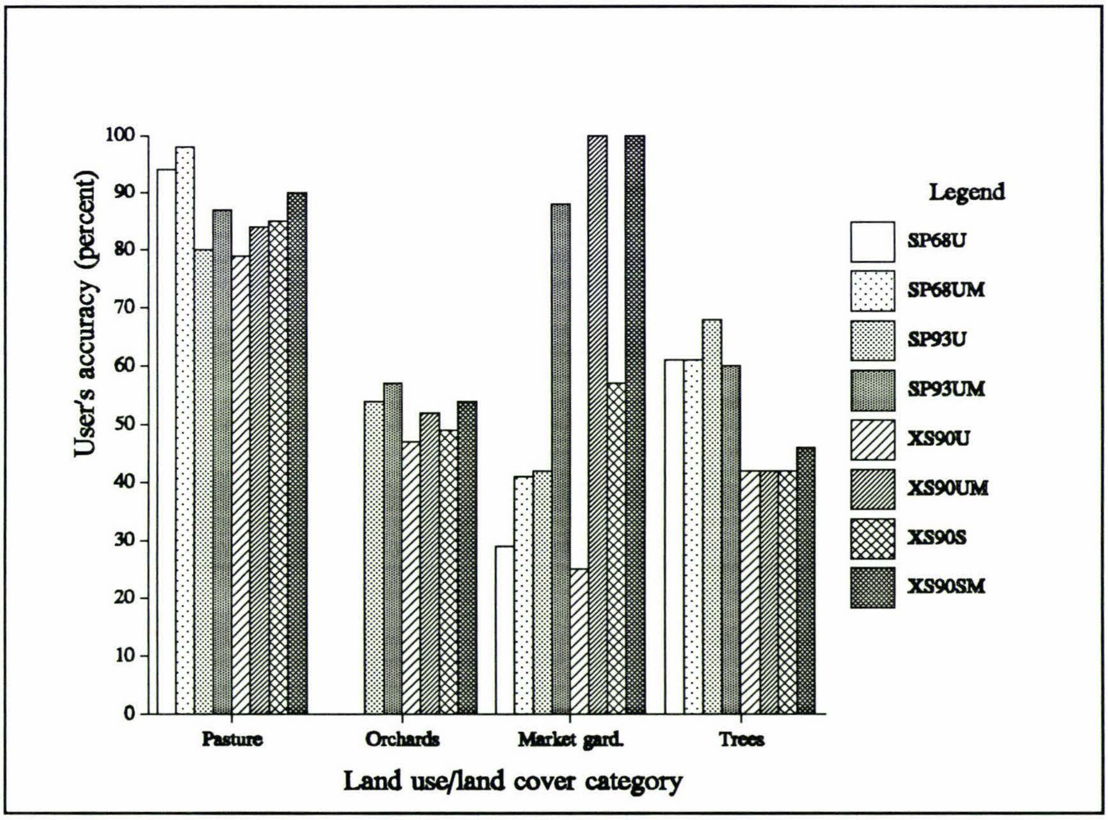


Figure 5.10 User's accuracies of major categories of land use/land cover maps for each technique applied.

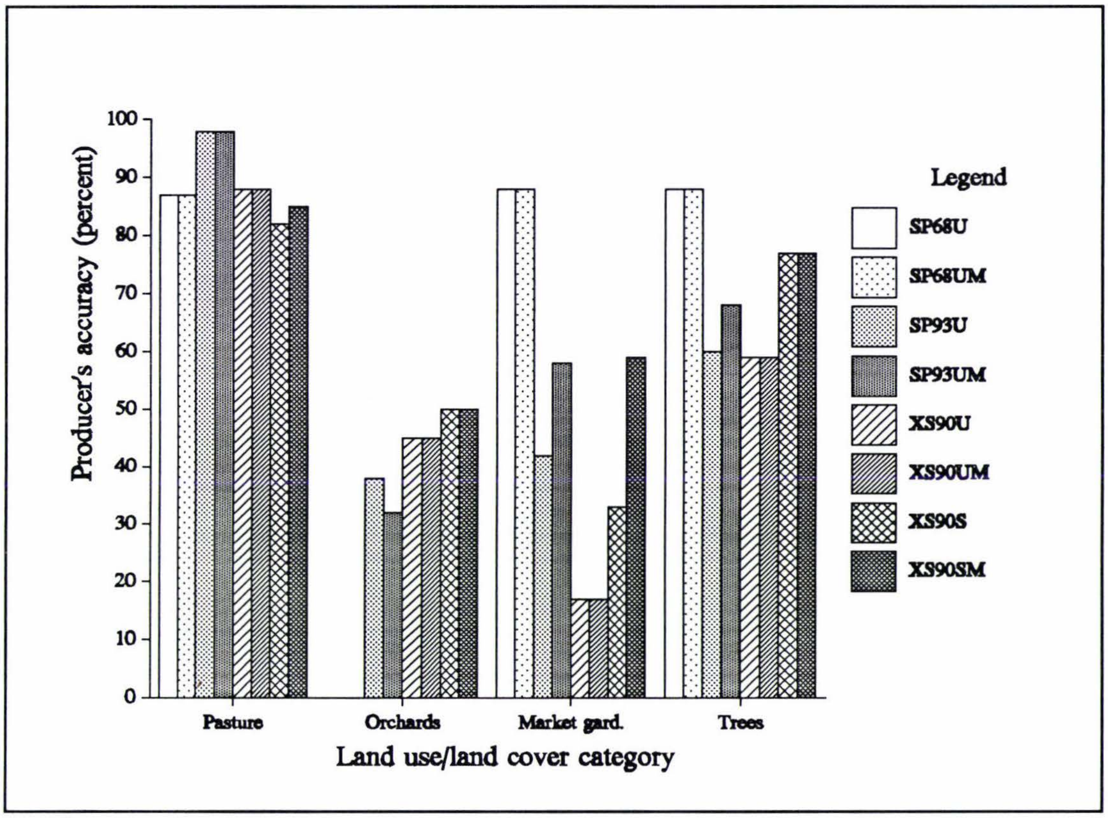


Figure 5.11 Producer's accuracies of major categories of land use/land cover maps for each technique applied.

trees category had a producer's accuracy of 88 percent (only 12 percent omitted), but the probability of trees being recognized as trees was only 61 percent, with 39 percent actually in the pasture class.

Compared to the 1968-SAP, the 1993-SAP gives a better result for category discrimination. Ninety eight percent of the 1993 grazed pasture areas were mapped as pasture, with an 87 percent chance of it actually being pasture land. However, due to the problems of spectral overlapping as well as the limitation in the number of spectral bands employed, trees and orchards tended to have low individual accuracies. For instance, trees could only be discriminated with a producer's accuracy and user's accuracy of 68 and 60 percent, respectively. While 40 percent of these areas were actually orchards on the ground. On the other hand, the error of omission of orchards (68 percent) indicates that of the areas that were found to be orchards on the ground, only 32 percent were mapped as orchards on the map.

Further, because of the similarity in spectral response, 58 percent of 1993-market gardens were classified as other classes (i.e., roads, residential, commercial, gravel pits, and pasture). This results from the fact that market gardening areas tend to have a wider range of spectral values, because they exist in various conditions ranging from barren land (very high brightness values) to a high proportion of ground cover. Using the SP93UM technique, such an error of commission was lowered to 22 percent.

In the case of accuracies of thematic maps derived from SPOT XS imagery, the supervised technique, in general, performs better in terms of category identification compared to those achieved by the unsupervised techniques. Although the producer's accuracy of pasture, using the unsupervised approach, is slightly higher than that when applying supervised technique (i.e., 88 and 85 percent, respectively, for XS90UM and XS90SM), this can be balanced by the higher user's accuracy given by supervised technique (i.e., 90 and 84 percent for XS90SM and XS90UM, respectively). Furthermore, by using the supervised approach 77 percent of trees were mapped, while the probability of them being correctly identified as trees on the ground was 46 percent. The unsupervised algorithm offered somewhat lower values, i.e., only 59 and 42 percent for producer's accuracy and user's accuracy, respectively.

These low values relate to the spectral similarity between trees, full-canopy orchard trees, and areas with high vigour grass. In this instance, 58 percent of trees (using XS90UM technique) were mapped as the latter two classes. Employing the XS90SM technique such an error could only be reduced by four percent. On the other hand, the problem of spectral similarity also led to lower individual accuracies for orchards class. Figure 5.10 and 5.11 indicate that when using the XS90SM technique, only 54 percent of orchards could be separated from other classes, while 46 percent of the area was identified as either trees or grass. Nevertheless, this result is slightly better than that obtained from the 1993 SAP.

One of the important differences between unsupervised and supervised techniques for this particular application is the ability of the algorithms used to discriminate the market gardening areas. As previously noted, this class was difficult to separate spectrally from other groups due to the wider range of spectral values it possesses. As depicted in Figure 4.11, the application of masking in the unsupervised approach (XS90UM) does not improve the producer's accuracy of market gardening (constantly low, i.e., 17 percent). However, using supervised technique, such an accuracy measure could be augmented by 26 percent (i.e., from 33 to 59 percent).

The relatively low producer's accuracy of market gardens may also relate to the time (season) when the imagery was acquired. The high error of commission is mainly due to a high proportion of market gardening area fully covered by crops, as the data were recorded on January, 30th- the time when the crops are likely to give a full ground coverage. As a result many pixels of market gardening area were identified as pasture. This circumstance is different to that found in the thematic maps derived from 1968-SAP and 1993-SAP images. In the latter two cases, market gardening areas were dominantly present as barren land, because the date of data acquisition coincided with the period of land tillage for crop cultivation. Thus, features with high spectral reflectance contribute greatly to the error of commission of market gardens. Such phenomena infer that it is always desirable to consider the seasonal cycle of crops being analyzed.

Given the analysis of thematic map accuracy just undertaken, it is evident that improvement is still needed for the overall levels of map accuracy. Eastman *et al.* (1993) noted that analysis of the error matrix as such, is particularly useful in deciding how to improve the classification accuracy of a map. Attempts to classify land use/land cover in the study area by digital techniques combined with masking the most prominent features have not been totally successful, indicating that a visual analysis method needs to be implemented to produce final land use/land cover maps of the study area. Incorporation of the latter approach ensures an improvement in the accuracy of the classification as a whole.

5.6 Post-classification refinement

A problem of land use/land cover classification, is that one spectral class may often represent either a subset or more than one information class. In post-classification refinement, these problem spectral classes are treated as separate special cases. Based on a sorting rule, individual pixels of the problem spectral class are assigned to the appropriate information class using ancillary data (Hutchinson, 1982; Richards *et al.*, 1982). For instance, the available cadastral polygons could be employed to assist post-classification sorting.

This stage was aimed at generating final land use/land cover maps using a technique of combining both the automated method and visual interpretation. Interpretation criteria were introduced and several related GIS-based operations were implemented to merge thematic category layers, obtained from manual interpretation, into the final land use/land cover maps.

5.6.1 Visual interpretation

As previously stated, digital processing is based solely on the different levels of intensity values (digital numbers). As a result, categories with very similar spectral behaviour are commonly mixed and it is not possible to define statistically the decision boundary around each category. In such a circumstance, visual interpretation is of a valuable aid the analysis procedure.

One of the main advantages of visual over digital analysis involves the possibility of using a greater variety of interpretation criteria. The incorporation of criteria like texture, pattern, association, location, size, shape, and site (Caroll *et al.*, 1977; Estes *et al.*, 1983; Lillesand and Kiefer, 1987; Star and Estes, 1990) makes it possible to differentiate cover types with spectral similarity. In this study, for instance, several tree species offer a similar spectral signature to orchards with full canopies. Some recently planted-orchard trees are spectrally similar to areas where mature crops occur in the market gardens. Likewise, gravel pits and some particular tracts of market gardens are essentially barren land, and have similar spectral reflectance to farm roads, and commercial and residential sites.

In order to obtain a final classification, it was necessary to apply several pertinent and suitable criteria- relating to the physical setting of the study area to assist the analysis procedures. For this purpose, the visual criteria introduced include tone, texture, pattern, and association. It was also possible to consider the seasonal cycle of planting particularly for the market gardening category. The interpretation procedure used to examine the 1968 and 1993 land use/land cover information was carried out on 1:20,000 scale-enlargement photographs. Due to unavailability of 1990 ground cover information photographically for the area of interest, however, visual analysis of 1990 data sets was done by displaying an enhanced three band colour composite of SPOT XS imagery (Figure 5.4). Such a technique has been also demonstrated by several authors (Astaras and Silleos, 1984; Martin and Howarth, 1989; Nicolaos, 1988; Sadar *et al*, 1982) for land use mapping.

Colour has made it possible to discriminate general categories such as vegetation, soil, urban areas, and water (Chuvieco and Vega, 1990), but it is still difficult to distinguish some types of vegetation. Consequently, to analyze vegetation (in this instance, trees, orchards, grass, and market gardening), the colour (tone) criterion needs to be evaluated in conjunction with other spatial characteristics. For example, the tonal characteristic of pastoral farming in the study region ranges from nearly black, in the case of green pasture, to almost white, as exhibited by dry pasture and those on the hilly terrain in the eastern edge of the study area. However, the main indicators such as smoother appearance and more even tone made it possible to distinguish pastoral farming from other land uses. Additionally, there was no evidence of recent cultivation in the area of interest and this contributed to its recognition as pasture.

As a general rule, horticultural areas which consist mainly of orchards and vine-fruits, are characterized by uniformly spaced rows of trees that give the appearance of a grid pattern. The presence of narrow-spaced shelter belts, which surround or bound the area, is additional clue to the identification of orchard areas. Thus the association of orchards and shelter belts was very useful in identifying the boundaries between orchards and either pastoral land or market gardens. Despite the fact that shelter belts also exist within the grazed pasture area, they were not likely to be confused with those occurring between orchards and the other two land use categories, because each class exhibits a discrete characteristic of tones, texture, and pattern.

It was not difficult to separate market gardens from both pasture grasses and orchards, despite the similarity in boundaries which tend to be straight and have definite borders and corners. No matter what the crop, or stage of growth, market gardens are characterized by the dominance of bare soil; even for mature crops, bare soil is present to a limited degree. In addition, mature crops in the market gardening area could be distinguished from recently sown-orchard trees by the finer row pattern, smoother texture, and the absence of cash crops within the area. Trees could be differentiated from other cover types by darker tone, rougher texture, and the fact that they tend to exist as irregularly distributed groups.

5.6.2 Distinguishing between orchards and market gardens

To identify boundaries between each cover type (i.e., orchards and market gardens), it was necessary to overlay the vector-cadastral polygons on the imagery for each date. This was done by using the COLOUR A function in the IDRISI system (described in Chapter IV, Section 4.1.3). As the cadastral boundary marks the property borders, the boundaries of orchards and market gardens were expected to match precisely the vector-lines of cadastral data sets. Any discrepancies were caused predominantly by the in accuracy of the registration process (discussed in Chapter IV).

A supplementary analysis was needed if one subdivision contained both orchards and market gardens. The actual edges of both classes had to be digitized using the ARCEDIT module in ARC/INFO. To do this, the following procedure was adopted:

a. Using the COLOUR A module in the IDRISI system, the original image was displayed, and the vector-cadastral boundary overlaid onto it. It was important to make sure that the *documentation files* had been updated, especially such parameters as resolution, maximum and minimum of X and Y coordinates (in the NZMS grid reference), and position error (i.e., total RMS errors obtained from the registration process). The areas were then examined for the categories of interest by using the WINDOW operation in the COLOUR A module.

b. The boundaries and edges of either the orchard or market gardening areas were then carefully identified, with the help of cursor, by implementing the visual criteria discussed above. Several continuous X and Y coordinates along the boundary line were then recorded (given in the seven digit-NZMS notation). These coordinates were corresponded to the positions at which *nodes* and *vertices* were to be placed to generate line segments in the ARC/INFO system.

c. While displaying the imagery with the vector overlay on the IDRISI screen, the relevant changes were made to the ARC/INFO coverages on a separated, but adjacent system. New coverages were created from the available cadastral arc-coverages using the COPYCOV function. Then, in the ARCEDIT module, changes were made according to the information shown on the IDRISI screen. This was done by firstly identifying the corresponding coordinates recorded using the WHERE function. Line segments were then added to the new coverage by placing nodes and vertices at the appropriate locations. Because the polygon topology had now changed this new coverage had to be cleaned and built to generate new topology.

Having built the new arc-coverages, it was important to identify all those polygons which were either orchards or market gardens. These polygons were then extracted as individual coverages employing the DISSOLVE module (discussed in Chapter IV Section 4.2.2). Subsequently, the rasterizing module (POLYGRID) was employed to create a grid-based data format compatible with IDRISI system.

It should be noted that the procedure described above was not necessary if an area had very distinct spectral characteristics, such as several tracts of market gardens which were dominated by bare ground features. These areas could be discriminated simply by using an automated approach.

5.6.3 Classification refinement of trees and pasture

In some regions of the study area, the discrimination between trees and pasture was very difficult, because the intensity values of high vigour grazed pasture are very low, thus similar to the spectral values of trees (Figure 5.12). Consequently, when applying the automated technique to the sub-scene of the study region as a whole, this particular area could only be identified as trees (see Figure 5.13). To discriminate both features, the technique of image segmentation needed to be employed. Image segmentation refers to the technique of segmenting the scene into homogeneous regions, and assigning each region as a single entity (Landgrebe, 1980; Mason *et al.*, 1988).

The procedure adopted to separate these two classes was similar to that applied to create masks of roads. The area of interest was first extracted from the

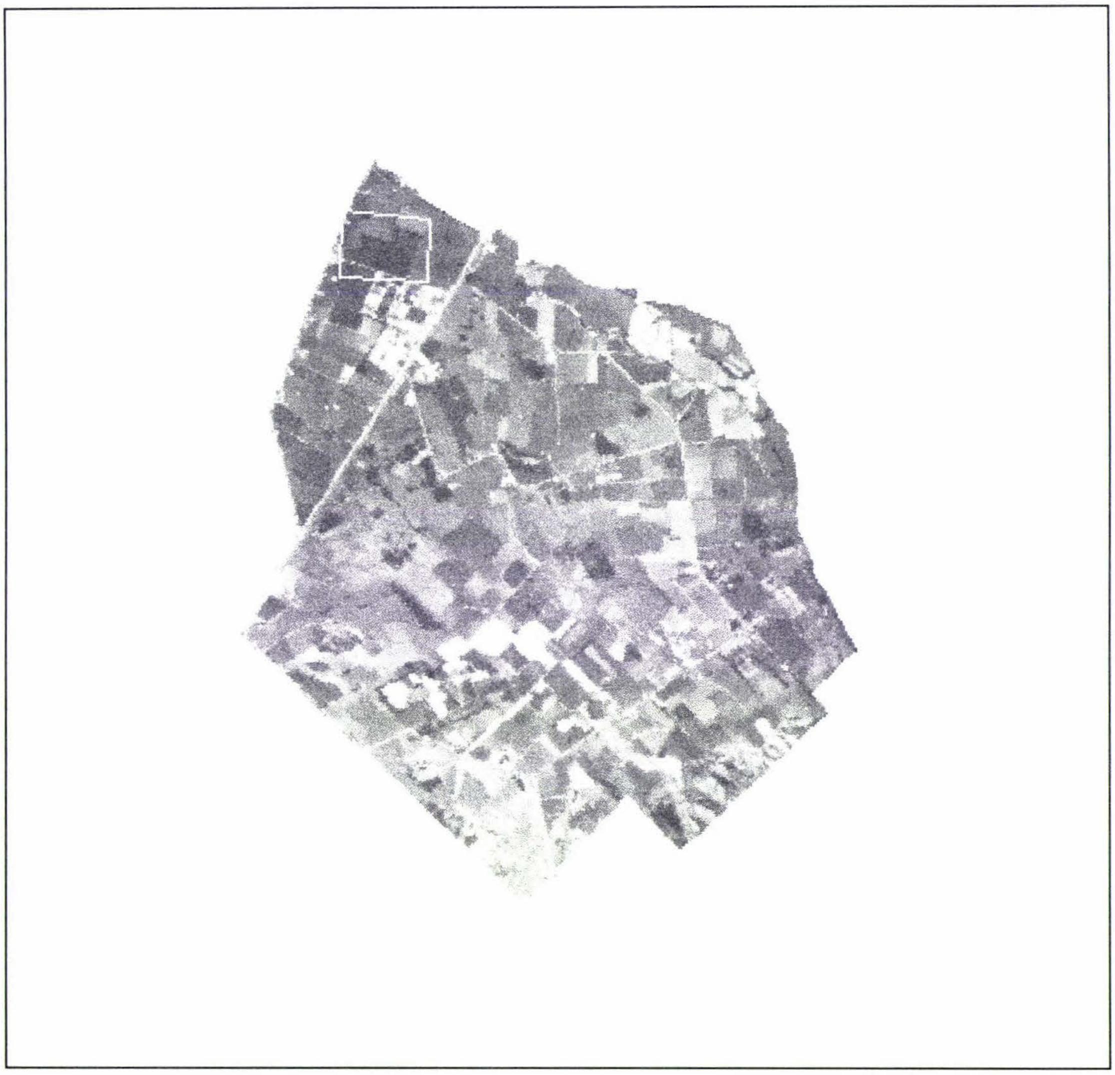


Figure 5.12 The 1968-SAP sub-scene with a small box showing the mixed spectral reflectance phenomenon between trees and pasture.

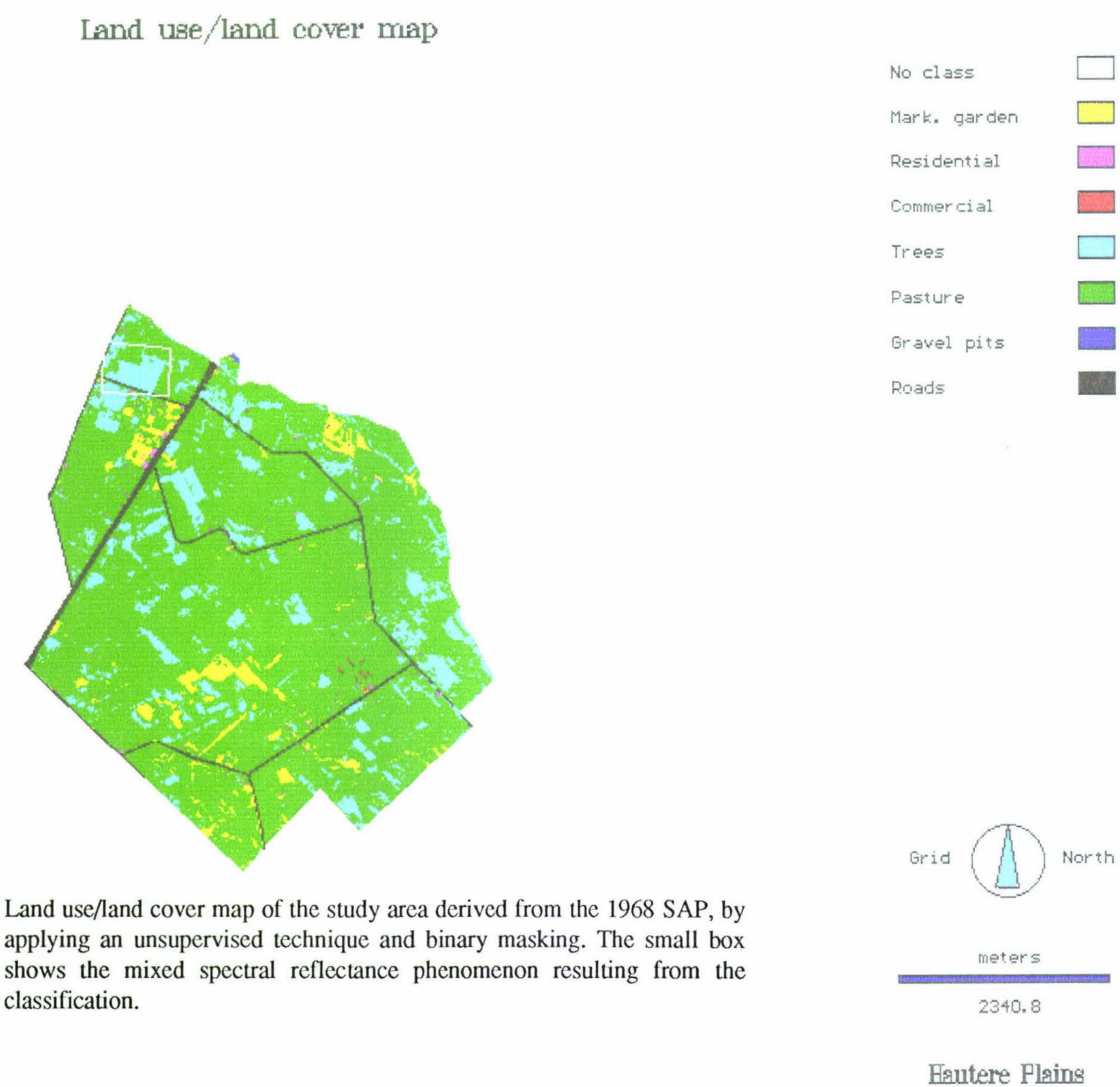


Figure 5.13 Land use/land cover map of the study area derived from the 1968 SAP, by applying an unsupervised technique and binary masking. The small box shows the mixed spectral reflectance phenomenon resulting from the classification.

imagery using an image mask created from the rasterised cadastral polygons. Threshold of the two classes was then identified either by displaying the histogram of this image window in the HISTO function, or by carefully examining the range of digital numbers of each cover type in the COLOR A function. When using the former, the threshold was usually selected in the gap between two histogram peaks. Both classes were then segmented using the RECLASS module. In this instance, values for trees ranged from the minimum to the threshold, and from the threshold to the maximum for pasture. Having completed such an operation, this image window could be merged into the original thematic map by employing the OVERLAY function.

The results of thematic information extraction after applying post-classification sorting are given in Figure 5.14. These images will be used for land use/land cover change detection analyses (Chapter VI). The 1990-thematic map is not presented here because, based on the visual analysis, there was relatively little change of land use/land cover between 1990 and 1993. The only significant change between the two dates occurred in the gravel pits along the river sides (discussed in Chapter VI).

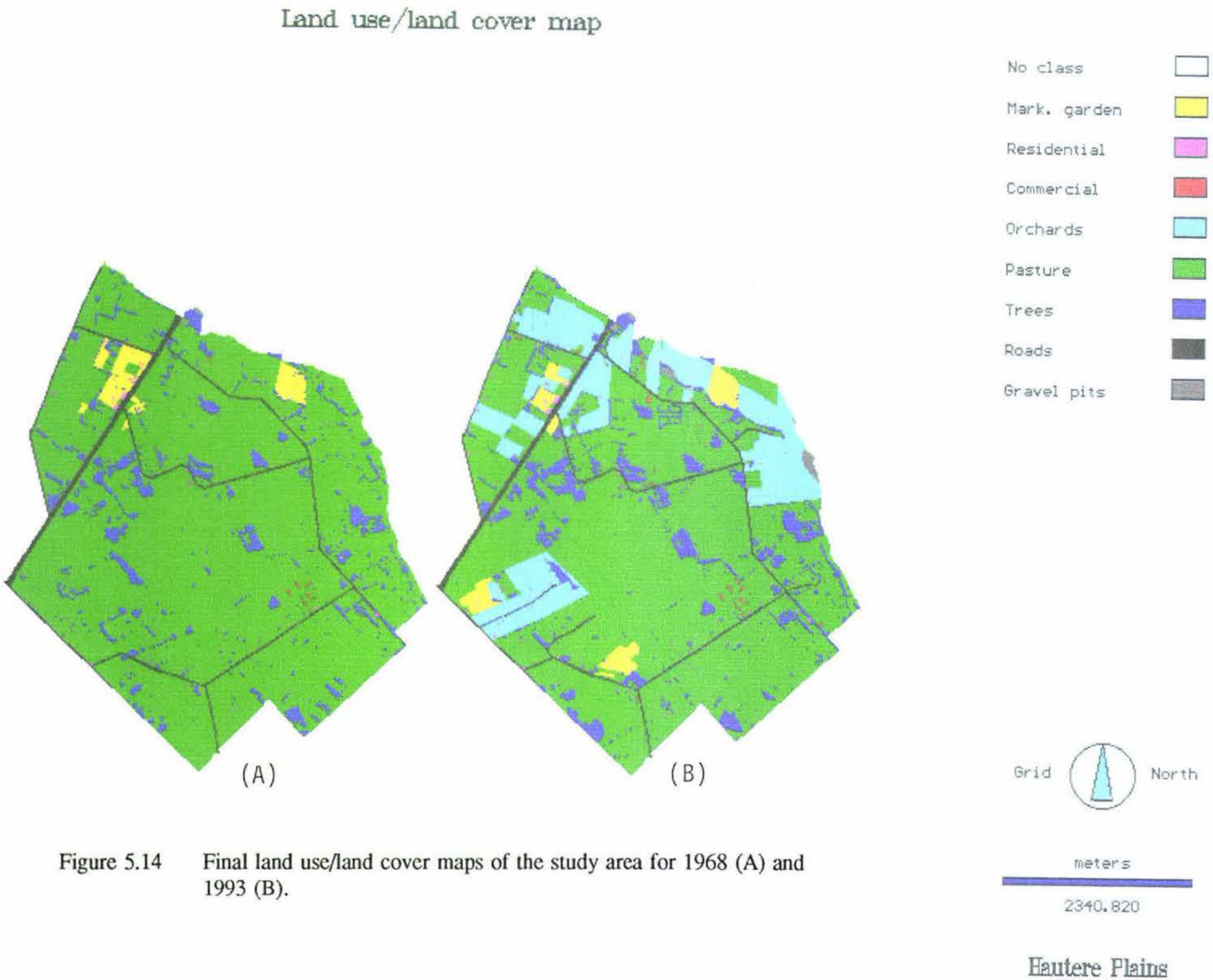


Figure 5.14 Final land use/land cover maps of the study area for 1968 (A) and 1993 (B).

CHAPTER VI

CHANGE DETECTION ANALYSIS

6.1 Introduction

Change detection is a process for determining and evaluating differences in a variety of surface phenomena over time (Mouat *et al.*, 1993). In principle, it also involves the ability to quantify temporal effects using multitemporal data sets (Singh, 1989). Obtaining appropriate multi-temporal data sets is thus indispensable for this purpose, to enable comparative assessment of changes on surface phenomena to be made both quantitatively and spatially.

The methods used for detecting changes may involve either manually interpreting land use/land cover information from aerial photographs for each period of time or digitally comparing and analysing the categories of interest using a computer-based system. The information on change generated from visual interpretation is static and cannot reveal the processes of changes that have taken place in space for each category of land use (Lo and Shipman, 1990). Digital techniques provide powerful tools for spatial-temporal analysis of the change by integrating spatial data collected from different sources and in different formats. Using this approach, the overlaying of two or more data sets can be carried out with ease, and the data are amenable to cross-comparison.

This chapter is intended to evaluate quantitatively and spatially land use/land cover changes in the study area by employing the classification maps produced in Chapter V. Changes in cadastral information between the two dates will be also examined.

6.2 Basic principles and algorithms for digital change detection

The basis of change detection is the comparison of remotely sensed data and map data, or the comparison of remotely sensed data sets taken at two or more different times (Mouat *et al.*, 1993). The digital nature of most remotely sensed data means it is best suited for computer-aided analysis (Singh, 1989), which could automatically correlate and compare a set of images taken of the same area on different dates. Such systems are capable of displaying the temporal differences of the study area. Lillestrand (1972) suggested that a significant increase in speed could be achieved for image processing by representing only the changes rather than expose the interpreter to all of the information in the images used.

A number of methods of digital change detection have been developed to deal with the monitoring of various earth surface resources. The technique chosen for a particular change detection application may depend upon the purpose, availability of data and tools, sources of data, and types of application. Milne (1988) divided the techniques into five broad groups: (a) manual interpretation,

(b) image differencing and ratioing, (c) image classification, (d) image transformation (albedo difference images), and (e) regression analysis. Some other methods which are in common practice include vegetation index difference (Angelici *et al.*, 1977; Howarth and Boasson, 1983; Nelson, 1983), principal component analysis (Byrne *et al.*, 1980; Eastmen and Fulk, 1993; Fung and LeDrew, 1987), an enhanced classification approach (Pilon *et al.*, 1988), post-classification comparison (Estes *et al.*, 1982; Gordon, 1980; Howarth and Wickware, 1981; Weismiller *et al.*,1977), and a GIS approach (Lo and Shipman, 1990; Westmoreland and Stow, 1992). Most of these methods were undertaken on the basis of two principal approaches (Singh, 1989): (a) comparative analysis of independently produced classifications for different dates, and (b) simultaneous analysis of multitemporal data.

In this study, the first approach was employed, i.e., independent production of classification maps (Chapter V), followed by an analysis of the changes between the two classification maps. The procedure used thus adopted the technique of post-classification change detection involving a GIS approach. Such a technique also permitted the analysis of property boundary changes between the two dates by utilizing data sets with a raster-based structure.

6.3 Analysis of land use/land cover changes

The basic data analysis methods used for land use/land cover change dynamic detection are image overlaying and binary masking (Pilon *et al.,* 1988). In this application, these methods were employed, utilizing GIS-based overlay functions facilitated in IDRISI, to evaluate category changes that had taken place in the study region. Only changes between 1968 and 1993 were dealt with, because based on the image classification results and a detailed visual interpretation (Chapter V Section 5.6.1), there was relatively little change in land use/land cover in the area between 1990 and 1993. The only change occurred in the gravel pits around the river meanders.

6.3.1 Detecting general changes

Preliminary evaluations of change involved identification of general changes in land use/land cover areas irrespective of spatial location and the type of change which had occurred. To do this, quantitative areal data of individual categories for the 1968 and 1993 classification maps (Figure 5.14) were first produced using the AREA function. The results are given in Table 6.1.

It is clear that for both years pastoral farming was the dominant land use in the region. However, a comparison of the statistics indicates that there has been

Table 6.1 Land use/land cover types in the Hautere Plains, 1968 and 1993.

Category	Area (ha)		Area (%)		Δ (ha)
	1968	1993	1968	1993	
Pasture	1757.82	1402.97	87.35	69.71	-354.85
Orchards	0.00	309.33	0.00	15.37	+309.33
Market gardens	43.65	50.06	2.17	2.49	+ 6.41
Trees	141.99	169.94	7.06	8.44	+ 27.95
Residential	7.72	12.41	0.38	0.62	+ 4.69
Commercial	1.43	3.25	0.07	0.16	+ 1.82
Gravel pits	0.65	6.87	0.03	0.34	+ 6.22
Roads	59.22	57.65	2.94	2.87	- 1.57
Total	2012.48	2012.48	100.00	100.00	0.00

a substantial decrease in the portion of grazed pasture area during this period; i.e., from 87.35 to 67.71 percent. The figures indicate that a total of 354.85 ha of pasture have been turned to other uses, while the area of orchards has jumped dramatically from none to 309.33 ha accounting for 15.37 percent of the study region. Land for market gardens has increased to some degree, from 43.65 to 50.06 ha, while many trees have also been planted; from 141.99 to 169.94 ha. Other land use/land cover classes, except roads, also exhibit an increase in area during this period. Each of these phenomena will be discussed in the following section, by considering also the spatial locations of the changes that have taken place.

The next stage of detecting change was to evaluate land use/land cover changes according to their spatial locations in the study area. The procedure followed is depicted in Figure 6.1. Pixels belonging to the same land use/land cover category for the two maps to be evaluated were first reassigned to the same value. SUBTRACT-OVERLAY routine was then utilized to subtract each pixel value for the 1968 land use/land cover classification map from its corresponding pixel location on the 1993 classification map. This operation gave rise to an output made up of positive, negative, and zero values. Positive and negative values represent pixels where change has taken place, while zero value denotes pixels

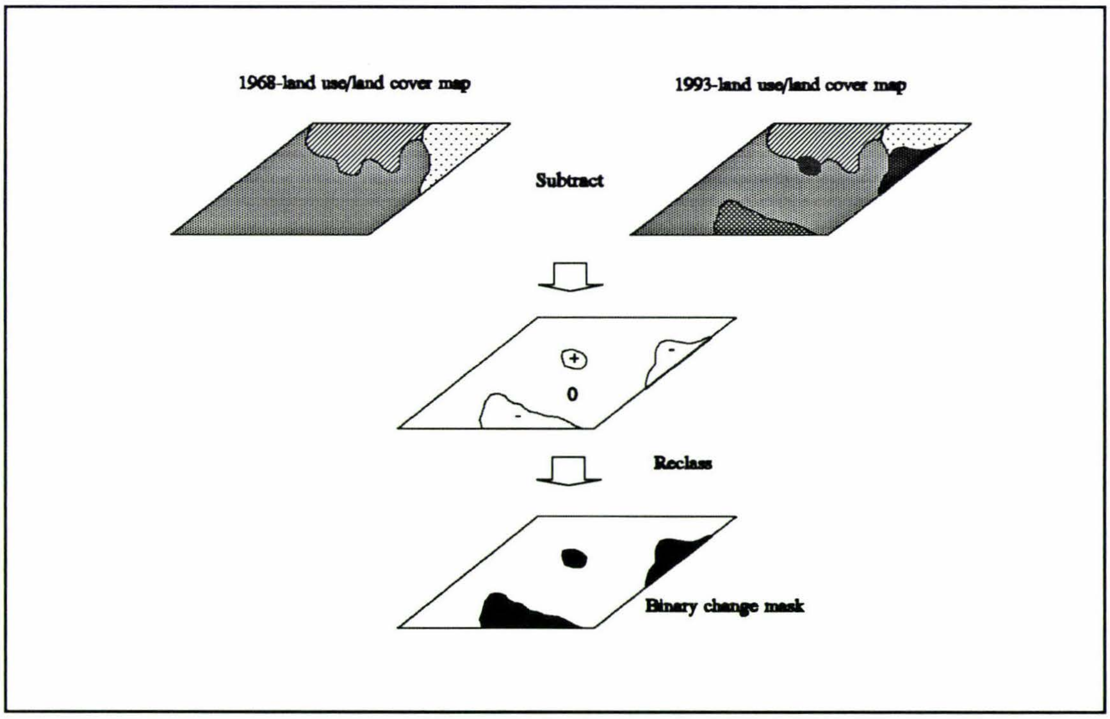


Figure 6.1 Diagram showing the procedure of creating a binary change mask.

where no change has occurred. A binary change mask was then created by reclassifying the positive and negative values to "1", leaving unchanged "0" value. Accordingly, the image produced consisted only value of "ones" representing area which undergone change in land use/land cover since 1968, and "zeros" indicating areas where no change in land use/land cover has taken place (Figure 6.2).

Following these, areas of "change" and "no change" were calculated by utilizing the AREA function. It was found that 493.39 ha, or 24.5 percent of the total study area (i.e., non-zero values in the original classification map) had undergone change. By difference the area of no change was 1519.09 ha or 75.5 percent.

6.3.2 Identification of "losses" and "gains"

This stage identified the nature and dynamics of change within the study area. The procedure was carried out following the flow diagram illustrated in Figure 6.3. A binary change mask prepared (Section 6.3.1) and then multiplied by each of the land use/land cover maps. As a result, all land use/land cover pixels whose corresponding value on the binary change mask image was zero (denoting areas of no change) were eliminated due to their multiplication by zero values. The remaining pixels, which represent areas of changed land use/land cover, retained their original value as a result of multiplication by one. The procedure thus gave rise to a masked classification image for 1968 and 1993 individually.

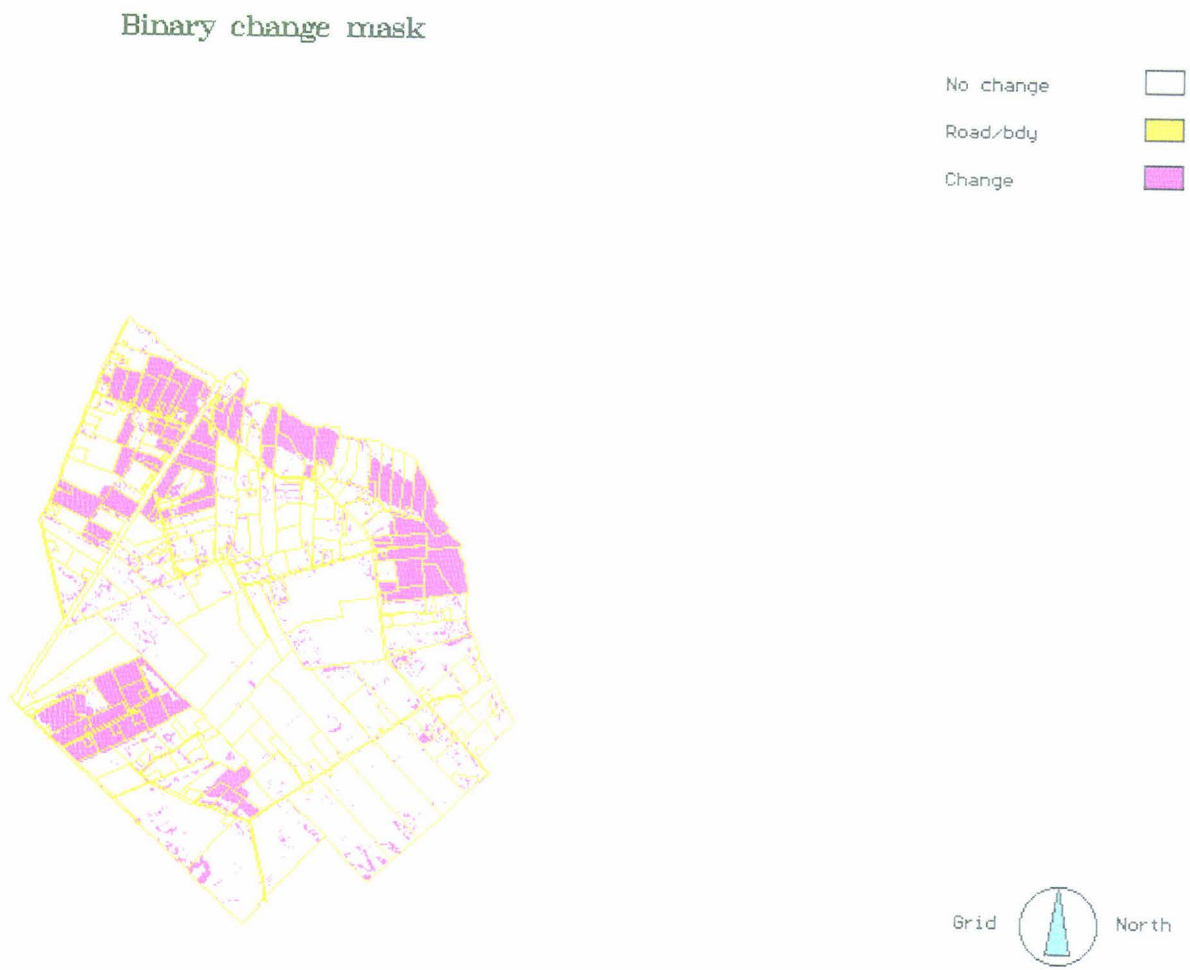


Figure 6.2 Map depicting the spatial location of areas of "change" and "no change" in the study region.

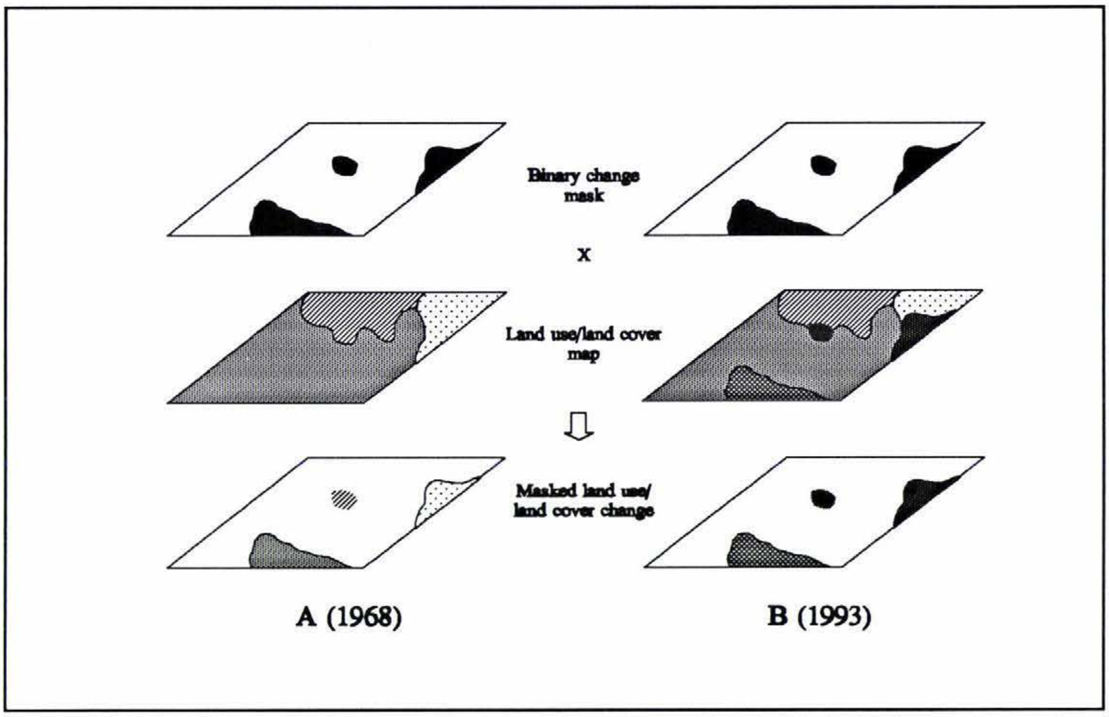
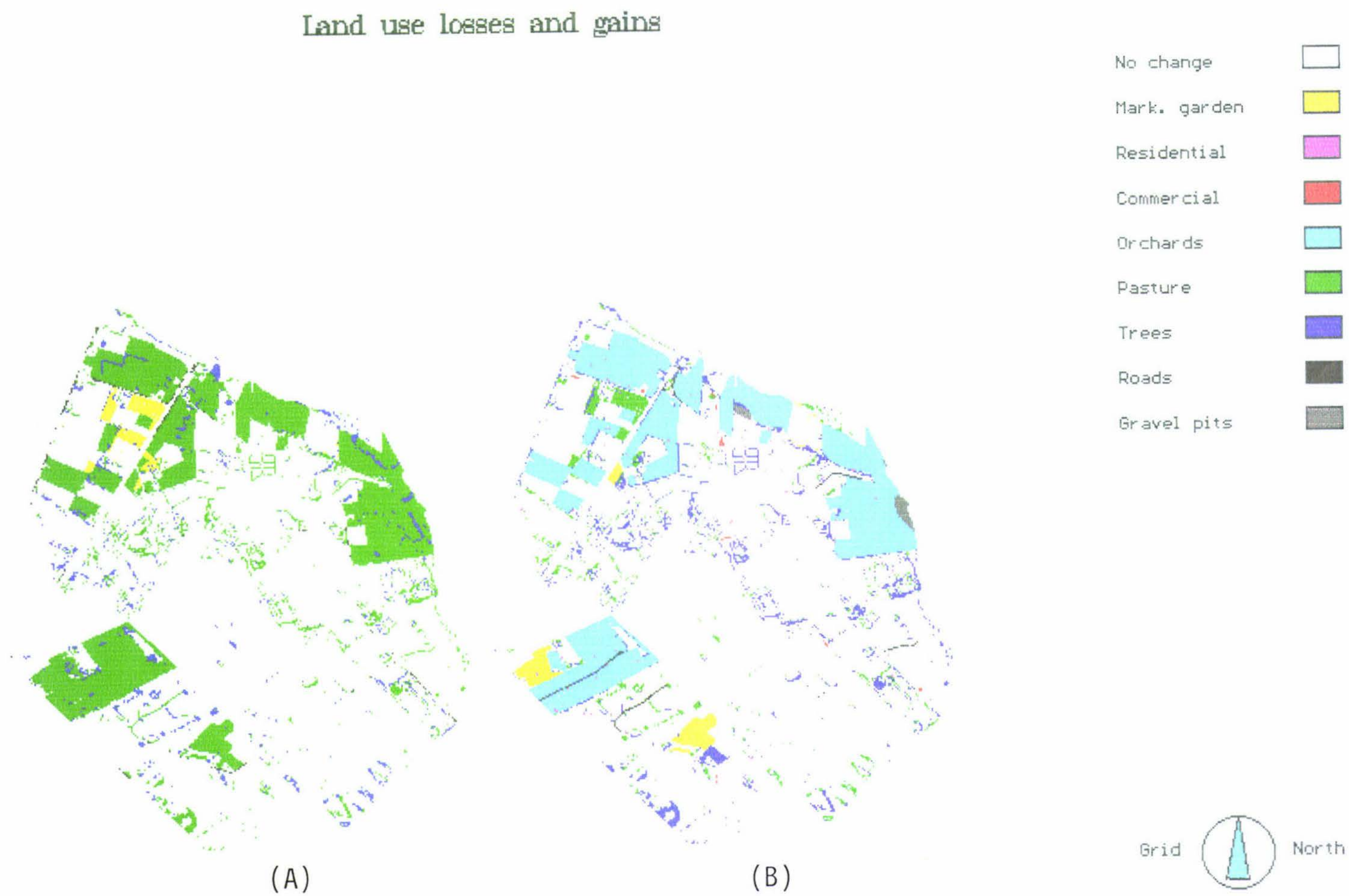


Figure 6.3 Diagram showing the procedure of generating masked classification images for 1968 (A) and 1993 (B).

Figure 6.4 depicts the results of this binary change mask operation employing a MULTIPLY-OVERLAY function. The masked 1968 classification image (Figure 6.4 A) indicates the spatial locations of each of the land use/land cover types, within the study area, that have been lost between the two dates; while masked 1993 image (Figure 6.4 B) shows the corresponding areas of each category that have been gained. Using the AREA function, such areas of losses and gains as well as net change for each category could be quantified. The percentage of net change of individual classes was calculated using the formula:

$$\text{Net change (\%)} = [(\text{gain-loss})/(\text{total area})] \times 100$$



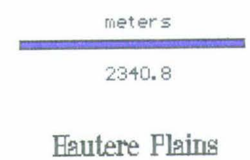


Figure 6.4 Maps showing land use/land cover types that have been lost (A), and those that have been gained (B) between 1968 and 1993.

A tabular summary of gains and losses for each class is given in Table 6.2. The figures indicate that during this period, grazed pasture area underwent a greatest land conversion to other uses (i.e, 407.54 ha) accounting for 20.2 percent (Figure 6.5) of the study region. However, a total of 52.69 ha (2.6 percent) of new pastoral lands were established replacing other land uses. This leads to a net change of -17.632 percent. The negative sign denotes the amount of net losses, while net gains are symbolized with positive values of net change.

The greatest net gain was found to be in the orchards class. As previously stated although orchards were not planted in the study area in 1968, they have been recently established on a total of 309.33 ha or 15.4 percent of the study area,

Table 6.2 Land use/land cover changes (losses and gains) in the Hautere Plains, 1968-1993.

Category	1968 (losses) (ha)	1993 (gains) (ha)	Net change (%)
Pasture	407.54	52.69	-17.63
Orchards	0.00	309.33	+15.37
Market gardens	20.76	27.17	+ 0.32
Trees	57.74	85.69	+ 1.39
Residential	0.00	4.69	+ 0.23
Commercial	0.00	1.82	+ 0.09
Gravel pits	0.02	6.24	+ 0.31
Roads	7.33	5.76	- 0.08
Unchanged	1519.09	1519.09	0.00
Total	2012.48	2012.48	0.00

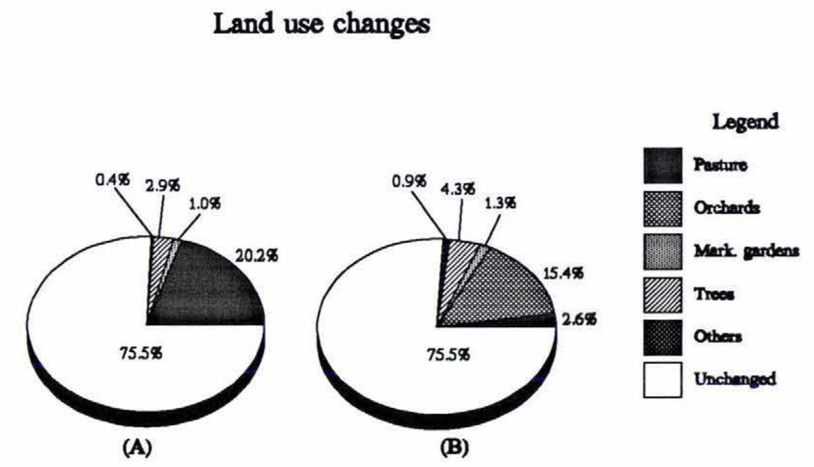


Figure 6.5 Diagram showing the proportion of the study area occupied by each of the land use/land cover categories which have undergone "losses" (A) and "gains" (B) between 1968 and 1993.

resulting in a net change of 15.37 percent. Further, it appears that 20.76 ha (losses) of market gardens were turned to other types of land use, and they tended to shift to other sites within the study region, indicated by 27.17 ha (gains) of new market gardening area opened during this period. Likewise, 57.74 ha (2.9 percent of the study area) of trees had been cut down and the area was converted to other types of land use. Nevertheless, it was found that 85.69 ha of trees had been planted by 1993. A net change of 1.39 percent, or increase by 27.95 ha of the area in trees.

Other classes, i.e., residential sites, commercial sites, gravel pits, and roads underwent little change. It should be noted that zero values for losses from residential and commercial categories are different from that for orchards. The former relate to the case of "no change" from the existing land use (residential and commercial sites) to other uses. In other words, they were present in 1968. While the latter indicates the absence of orchards in 1968. Further, it appears that there was no conversion of gravel pits to other cover types, while 6.24 ha of new exposed gravels have been gained during this period. Total land use/land cover losses for these four categories were 0.4 percent, while the increase represents 0.9 percent (Figure 6.5).

6.3.3. Cross-comparative assessment of individual category changes

The analysis procedures discussed previously have mostly dealt with land use/land cover changes quantitatively by considering the spatial locations of changes that have taken place over the study area. However, It was thought particularly valuable to undertake a cross-comparative evaluation of individual category changes to obtain a more comprehensive understanding of the dynamics of change that had occurred. This task was carried out using the CROSSTAB routine which permitted the identification of the before-change and after-change maps (Figure 6.4) on a category by category basis. The results are presented in Table 6.3.

Table 6.3 Cross-tabulation of the total area for each category showing a 'trend' of changes between land use/land cover types.

Category		1968							
		Past.	M.gar	Tree	Res.	Com	Grav	Rd.	Total
1993	Pasture	1350.28	13.95	38.57	0	0	0	0.17	1402.97
	Orchards	287.77	4.03	17.49	0	0	0	0.04	309.33
	M.gardens	26.43	22.89	0.74	0	0	0	0	50.06
	Trees	75.90	2.66	84.25	0	0	0.02	7.11	169.94
	Residential	4.40	0.11	0.17	7.72	0	0	0.01	12.41
	Commercial	1.77	0	0.05	0	1.43	0	0	3.25
	Gravel pits	5.97	0	0.27	0	0	0.63	0	6.87
	Roads	5.30	0.01	0.45	0	0	0	51.84	57.65
Total		1757.82	43.65	141.99	7.72	1.43	0.65	59.22	**2012.48**

Note: Values underlined indicate the numbers of pixel overlap, of the same class, between the two dates.

From these numeric tabulations it is shown that 1350.28 ha of grazed pasture land remained unchanged (values underlined), 407.54 ha were lost during this period. From the pasture land, 70.61 percent (Table 6.4), were lost to orchards; 6.49 percent to market gardening; or 18.62 percent to trees; 1.08 percent to residential land; and so on. On the other hand, a total of 52.69 ha of pastoral farming in 1993 were gained from 13.95 ha of market gardening (26.48 percent), and 38.57 ha (73.20 percent) of trees (Table 6.5). The remaining land (0.32 percent) was gained from roads.

Table 6.4 Percentage of land use/land cover changes, from one to another, relative to "losses."

Category	Pasture	M.gard.	Trees	Gravel	Roads
Pasture	-	67.20	66.80	0	2.32
Orchards	70.61	19.41	30.29	0	0.55
M.gardens	6.49	-	1.28	0	0
Trees	18.62	12.81	-	100	96.99
Residential	1.08	0.53	0.29	0	0.14
Commercial	0.43	0	0.09	0	0
Gravel pits	1.47	0	0.47	-	0
Roads	1.30	0.05	0.78	0	-
Total	100	100	100	100	100

Table 6.5 Percentage of land use/land cover changes, from one to another, relative to "gains."

Category	Past.	Orch.	M.gr	Tree	Res.	Com	Grav.	Rd.
Pasture	-	93.03	97.28	88.58	93.82	97.25	95.67	92.01
M.gardens	26.48	1.30	-	3.10	2.35	0	0	0.17
Trees	73.20	5.65	2.72	-	3.62	2.75	4.33	7.81
Gravel pits	0	0	0	0.02	0	0	-	0
Roads	0.32	0.01	0	8.30	0.21	0	0	0
Total	100	99.99*	100	100	100	100	100	99.99*

Note : * Total values do not sum up to 100 due to decimal rounding.

From this analysis, it was noted that an anomaly occurs where change is from roads to trees; 7.11 ha of trees have been gained from roads (Table 6.3). In reality, this is not the case of a "real" land use change from road to pasture, but it is related to the many shelter belts which have grown along the road edges- for this particular case, the road edge is the border between property boundaries and roads, thus including the space between surface sealed road and the property boundaries. Consequently, a great portion of roads were covered by the canopies of high-stand shelter belts. This is confirmed by the information on the percentage of category losses in Table 6.4, in which 96.99 percent of road losses were replaced by trees.

Another important data analysis was to evaluate the percentage of overlapping areas of the same category between the two dates, relative to the total area of corresponding class in the 1993-land use/land cover data. This may provide information about the area of each land use/land cover class residing in a particular spatial location in 1968 that is still unchanged in 1993, while non-overlapping areas reveal the areas of "gains" from other types of land use. From this evaluation, it was found that 96.24 percent of grazed pasture land were spatially still unchanged, only 3.76 percent of the area had been gained from the other uses (Figure 6.6). The orchard land exists only in 1993, thus there was no pixel overlap between the two dates, resulting in 100 percent of "gains." In other words, all the current areas in orchards were developed from other land uses.

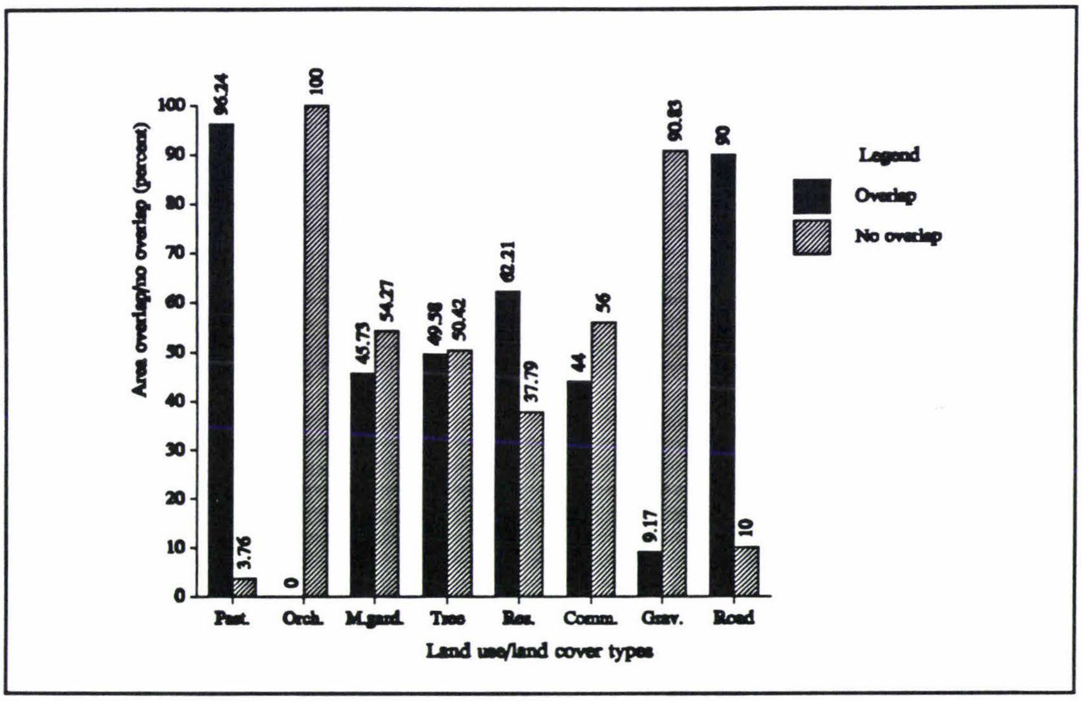


Figure 6.6 Percentage of overlapping and non-overlapping area of each land use/land cover class between 1968 and 1993 relative to the total area of corresponding class in 1993.

Further, it appears that 45.73 percent of the land in market gardening has been retained on the same area, while 54.27 percent of the land has been established from the other land use areas. Similarly, for trees, 49.58 percent were found to be in the same location over this period, while 50.42 percent of the area they occupied in 1993 had been newly grown. The largest portion of land cover "shifting" was identified in the gravel pits category. A total of 90.83 percent of the area had been extended onto the other land use types, while only 9.17 percent of the area of this class remained in the same spatial location during the period between the two dates.

As noted previously there was relatively no change in land use/land cover types between 1990 and 1993. The only change identified was the extensive area of gravel pits in some portions of land near the Otaki River. Using the same procedure as that used for assessing the change between 1968 and 1993, it was found that only 2.93 ha of exposed gravels existed in the study area in 1990; thus increasing 2.28 ha from 1968 (approximately 22 years). This indicated that in only within three years (1990-1993) the area of this cover type had increased by 3.94 ha. From cross-tabulation analysis, it was estimated that 39.24 percent of the area of 1990-and 1993-gravel pits were overlapped, revealing that 60.76 percent of current exposed gravels had occurred in a new area of the study region at the time between 1990 and 1993.

6.4 Analysis of property boundary changes

The initial stage of cadastral data analyses involved the identification of the distribution of land parcels in the study area according to size. This task was first carried out by utilizing the data base (polygon attribute table) of property boundaries stored in PC ARC/INFO, then displaying the results in the IDRISI system. The latter was done after performing a data integration procedure described in Chapter IV Section 4.2.2. Following this, an attempt was also made to identify the spatial locations of changing property boundaries. Like the method applied to analyze land use/land cover changes, only 1968-and 1993-cadastral data sets were used to evaluate property boundary changes, because there was no change in property boundaries between 1990 and 1993.

6.4.1 Distribution of subdivisions according to size

By utilizing the TABLES module of the ARC/INFO package, it was possible to identify the size of all polygons, which is always automatically stored as the *area item* in the cadastral arc-coverages. Because polygon size was calculated automatically in square metres, it had to be converted to hectares to maintain uniformity of the area units used for this application. To do this, a new *item* was first defined using the ADDITEM function, and the CALCULATE command was then employed to convert m^2 to *ha* by dividing each value in the *area item* by 10,000. The results were stored in the new *item* specified.

Using the technique just described, it was also possible to define the ranges of polygon size and store them in a separate *item*. The selected size ranges for evaluating land parcel distribution in the study area were: < 0.5 ha; 0.5-1 ha; 1-2 ha; 2-4 ha; 4-6 ha; 6-10 ha; 10-20 ha; 20-40 ha; and > 40 ha. The numbers of polygons falling within each range were then estimated using the FREQUENCY module. The result, which illustrates the number of subdivisions in 1968 and 1993 within the specified size ranges, is presented in Figure 6.7.

Once the size of all the subdivisions had been identified, it was necessary to produce maps showing the distribution of the different-sized land parcels in the study area; each size range was represented by a discrete colour. The cadastral arc-coverages were first converted to the raster-based data structures compatible with the IDRISI system (discussed in Chapter IV Section 4.2.2). The AREA module was then employed to create cadastral images ("area maps"). Using this

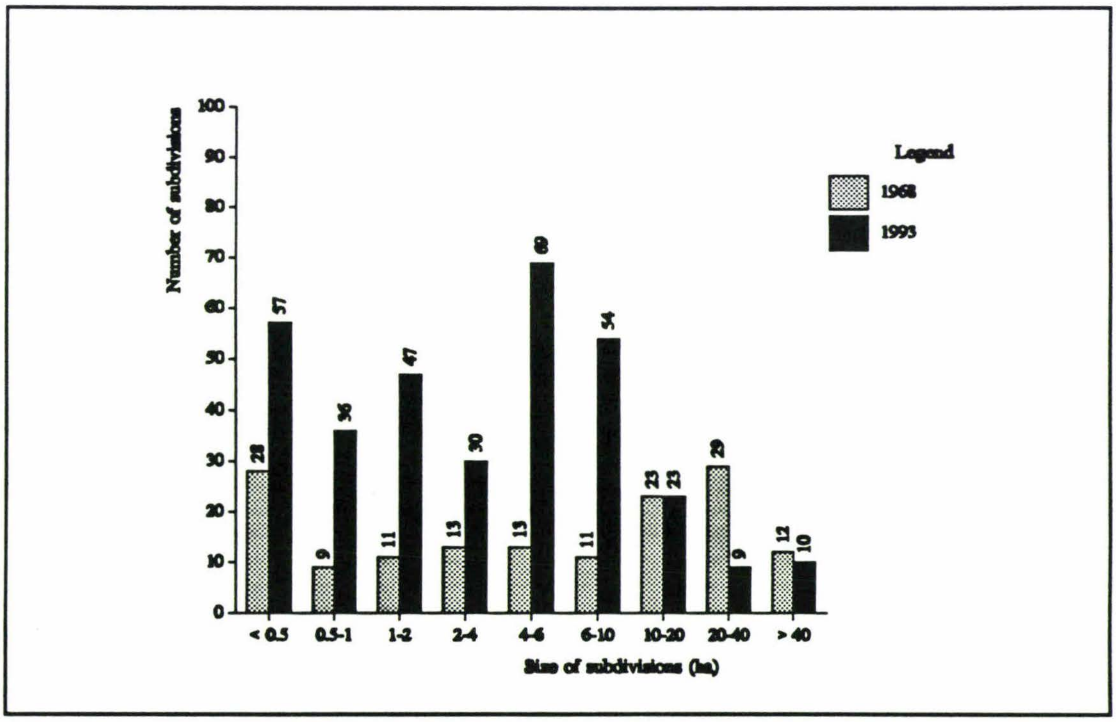


Figure 6.7 The number of subdivisions in each size range in 1968 and 1993.

module, each land parcel was attributed to one colour and one real value. The real value indicated the area which could be given in cells, hectares, acres, square metres, square feet, square kilometres, or square miles. Because the main interest was to display the range of parcel sizes, land parcels in these "area maps" were then grouped according to the value ranges specified above, using the RECLASS function. The resultant images are shown in Figure 6.8.

It is noticeable that the number of smaller land parcels had increased considerably, due to the subdivision of larger sized parcels to smaller parcels, as the numbers of land parcels sized more than 10 ha (25 acres) had decreased. The greatest increment is found for subdivisions sized 4-6 ha and 6-10 ha. Such land parcel sizes (approximately 10 to 25 acres) are considered to be the most favoured size for properties in the study region where orcharding and market gardening are practised (illustrated in Figure 6.9). In addition, these subdivision sizes also appeared to be the popular size for 'lifestyle blocks' in such a rural area as the study region.

6.4.2 Identification of subdivision change

The next step was to identify the spatial location of subdivision change in the study area between the two dates. The procedure was undertaken by creating a binary change mask, like the technique applied to evaluate land use/land cover

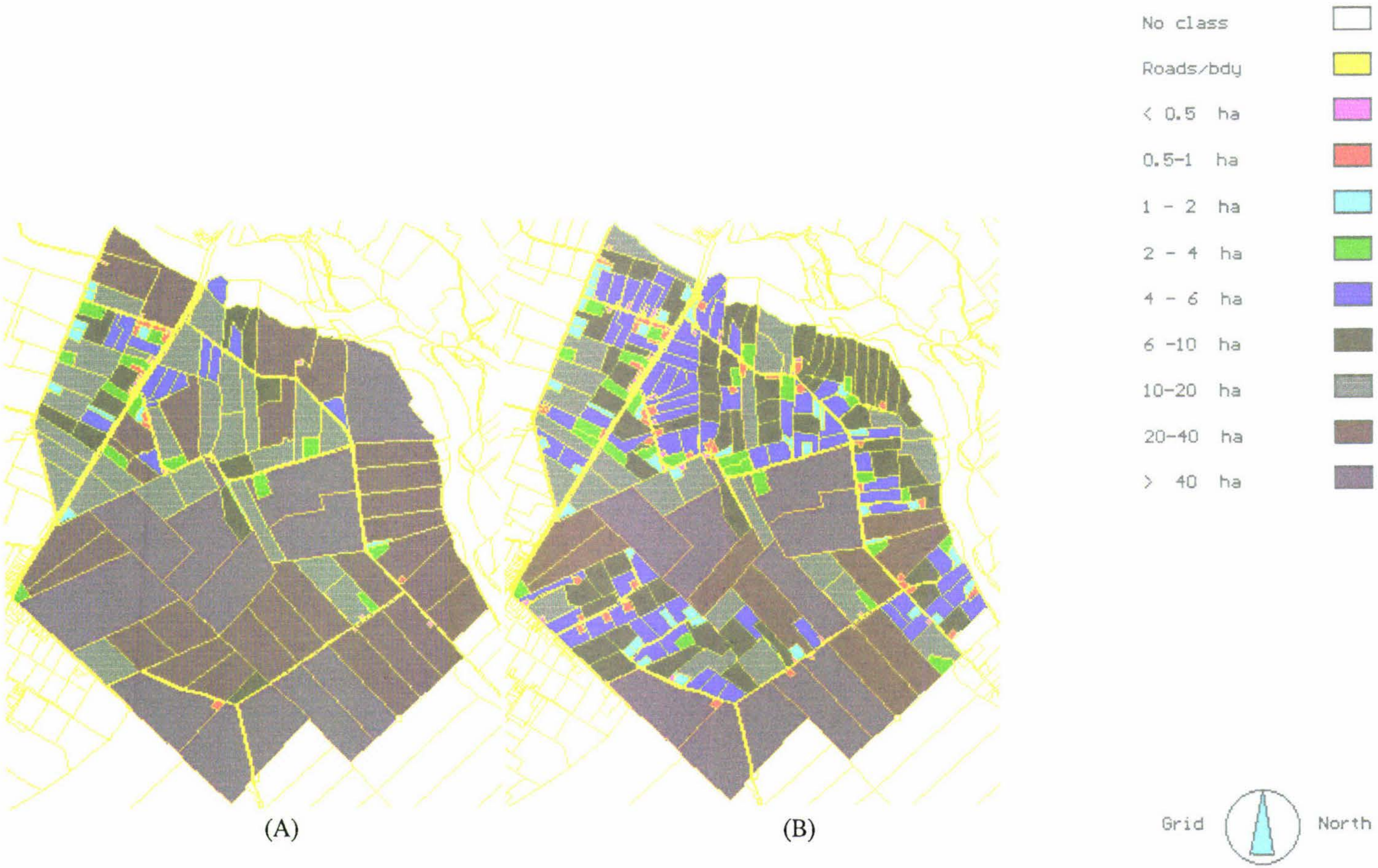


Figure 6.8 Distribution of 1968 (A) and 1993 (B) land parcels in the study area according to size.

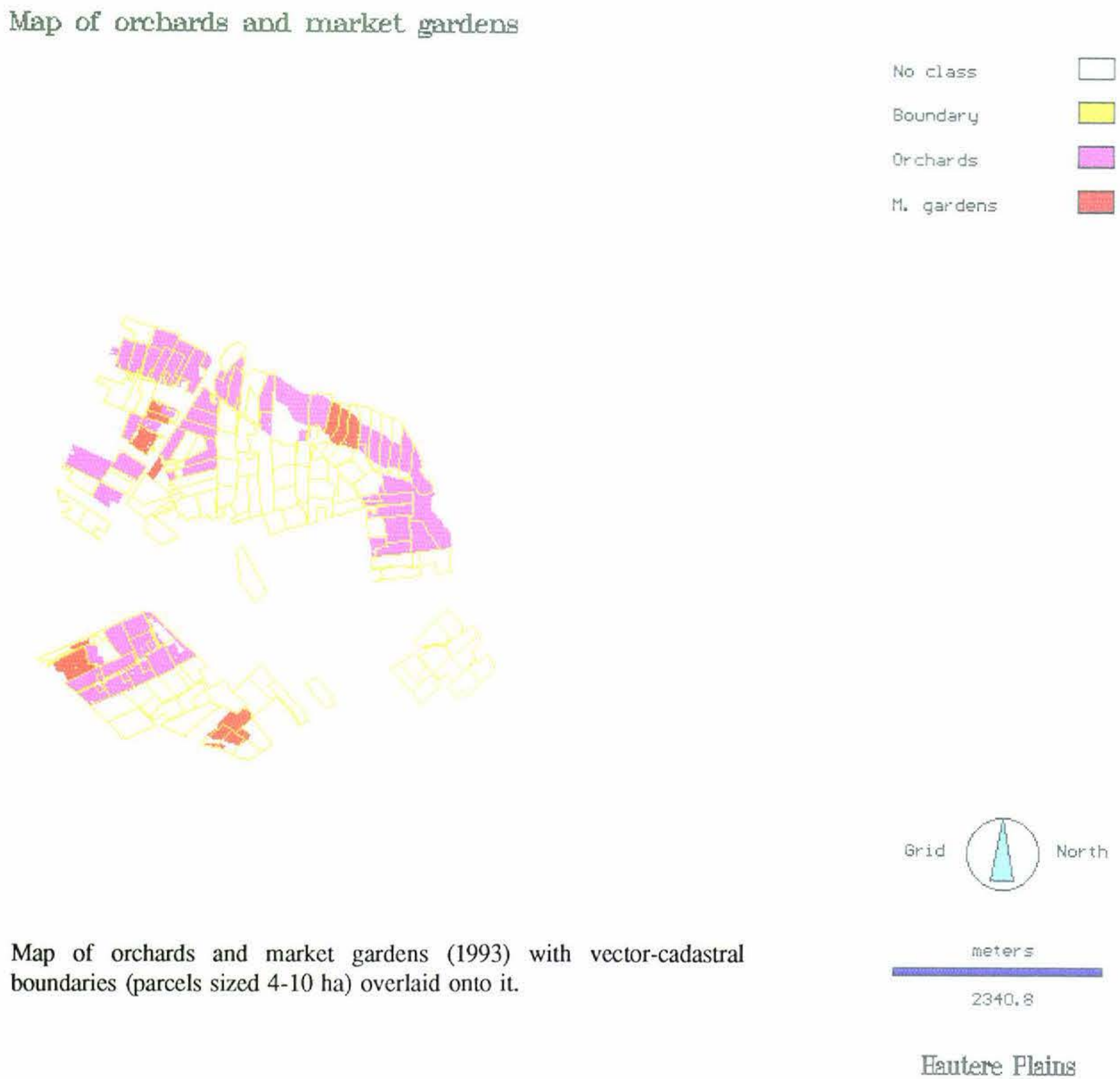


Figure 6.9 Map of orchards and market gardens (1993) with vector-cadastral boundaries (parcels sized 4-10 ha) overlaid onto it.

changes. Thus the image mask displayed all the land parcels that had undergone a change in size. In other words, the change mask depicted the spatial location of all larger sized parcels before change, which had been subdivided to the smaller size in the period between the two dates. The binary change mask of cadastral data sets is presented in Figure 6.10. It revealed that the area of change ("ones" value) is 968.71 ha accounting for 48.14 percent of the study area, while unchanged polygons ("zeros" value) were found to be 1043.77 ha (or 51.86 percent).

6.5 Discussion

As noted earlier pastoral farming has been the predominant land use on the Hautere Plains from the time of bush clearing; in 1968 it accounted for 87.35 percent of the study area. However, since about 1980 land use has been intensified and diversified with a growth in perennial horticulture (mainly orchards) and some market gardening (Simcock, 1986). It has been determined that change from pasture to orchards accounted for 70.61 percent of the total pasture "losses," while market gardens accounted for 6.49 percent. Moreover, the cross-tabulation data in Table 6.5 revealed that the greatest "gains" of all types of land use/land covers were achieved from pasture land ranging from 88.58 percent for trees to 97.28 percent for market gardens. These have brought about a

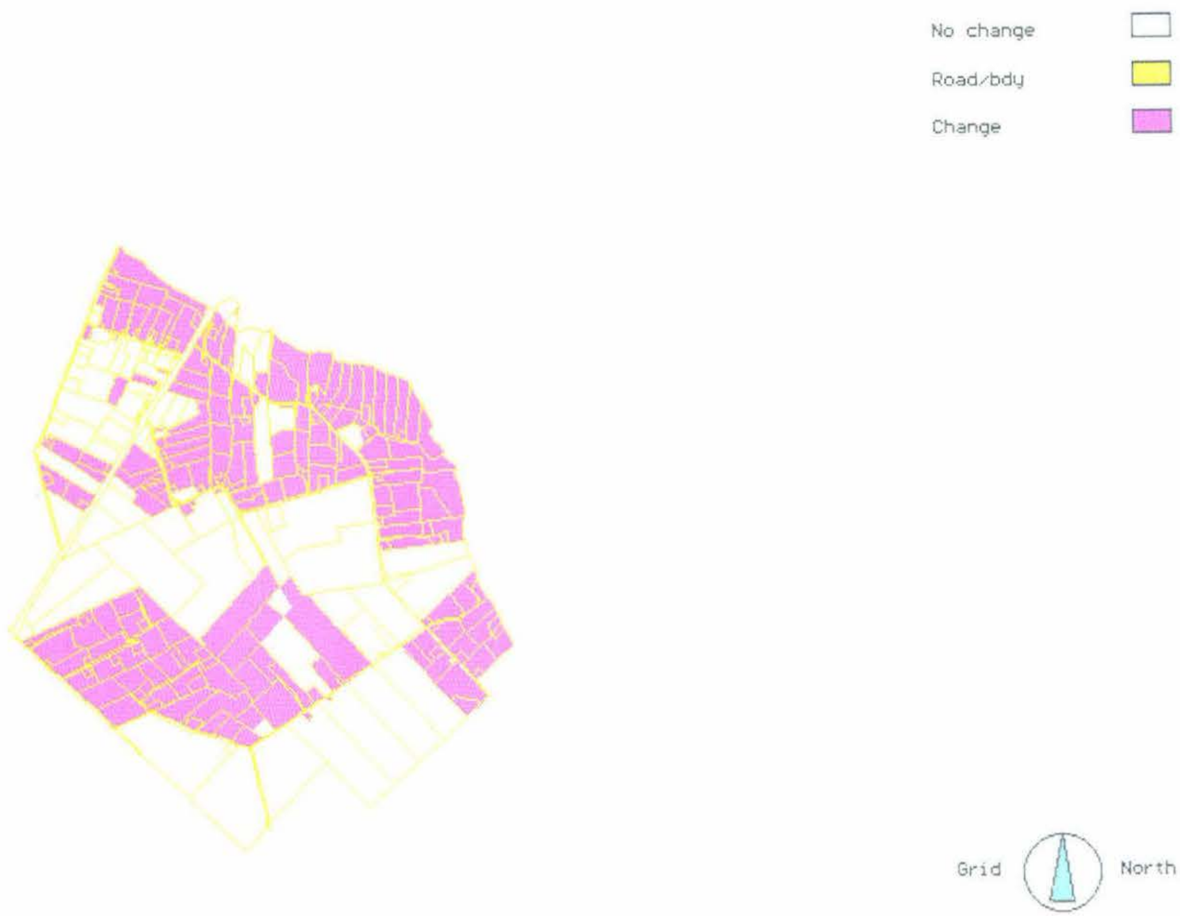


Figure 6.10 Binary change mask of cadastral data showing the spatial location of land parcels that have undergone subdivision (change) and those that remained unchanged.

dramatic reduction of grazed pasture land on the plains for the period between 1968 and 1993.

During this period, there was a significant increase in the area of trees, resulting from an increase in both the number and canopy coverage of shelter belts and bush reserves in the plains. Similarly, the area of exposed gravels had also increased significantly, particularly during the period between 1990 and 1993. According to a local farmer, a major flood of the Otaki River had occurred in 1992, and affected surrounding land, especially the area of horticulture located along the river flat. Apparently, greater flood damage took place on the land around the river's meandering in the northeastern margin of the study region. A comparative analysis of category changes between 1990 and 1993 showed that 3.63 ha of land in orchards had been lost because of the deposition of gravels from floods. This portion of land was still unproductive in 1993.

Based on a careful interpretation of the dynamics of land use/land cover change several anomalies of change between categories, which are due to the technique of data extraction employed, were identified as follows:

1. The change from trees to orchards was not usually entirely attributable to the removal of trees. Numerous scattered groups of trees as well as long-thin shelter belts occurring within the area for orchards were categorized into orchards class. As discussed in Chapter V this was due to the problem

of spectral similarity between the two categories which resulted in a difficulty of separating them spectrally.

2. As expected, there were no "losses" of residential and commercial categories; only "gains" were measured. This was because the after-change (1993) maps of both categories were obtained from the multiplication of the original 1993 sub-scene by each of the binary masks of residential and commercial categories (1968), followed by the addition of category "gains" to the resultant image (1993 sub-scene) manually.

3. Some areas of change from trees to roads were not actually the replacement of roads by trees, but the decline in the area of roads was due to the effect of an increase in the areal coverage of high-stand shelter belts occurring along the road margins.

CHAPTER VII

GENERAL DISCUSSION AND CONCLUSION

7.1 General discussion

The major advantages of a digital approach for land use/land cover classification include the ability of overcoming the limitation in the human eye to discriminate all grey levels recorded in the remotely sensed data, and of analysing ground features in several different bands. Along with these benefits, however, spectral-based classification tends to be more successful for scenes characterised by large areas of homogeneous cover (Westmoreland and Stow, 1992). Often, problems are encountered when dealing with areas composed of heterogeneous land use/land cover types; as in the present study, several defined classes (e.g., high vigour grass, orchards and trees; or gravel pits and recent-cultivated market gardens) tended to exhibit spectral confusion and result in mixed pixels. The solution to these shortcomings is to extend the classification procedure through the incorporation of ancillary data and/or knowledge in pre- and post-classification techniques and the classifier option (Hutchinson, 1982; Janssen *et al.*,1990).

It has been demonstrated that the availability of digital-rectified cadastral data was very helpful in providing binary masks to assist in the discrimination of land use/land cover categories which were difficult to differentiate when using only spectral-based classification. Merging additional information derived from visual interpretation as well as site inspection in the pre- and/or post-classification stages led to more accurate GIS-based operations facilitated in IDRISI, as did employing binary masks created from cadastral data. The use of such a procedure necessitated all the data sets used to be in geometric conformance. Therefore, the technique applied relied heavily on accurate registration between the images and the cadastral layer.

The basic premise of digital change detection is that changes in land use/land cover must result in changes in the spectral response of pixels on the multi-date images (Jensen, 1986; Singh, 1989). However, these changes are often influenced by environmental factors such as atmospheric condition, differences in sun angle, and soil moisture variations (Jensen *et al.*, 1983). These adverse effects could be minimized by selecting the data acquired at the *anniversary dates* (Jensen, 1986) or, as in the present study, by applying a post-classification comparison technique which independently produces classification maps. Moreover, the use of the latter method could also minimize the problem of normalizing for sensor differences between the two dates (Singh, 1989), as in the present application, the data sets used were acquired at a different time of the year and from sensors with different characteristics (Chapter III Section 3.2.1).

In a post-classification change detection the accuracy of the results is dependent upon the accuracies of the geometric registration and the initial classification of the multi-temporal images used (Christensen *et al.*, 1988; Howarth and Wickware, 1981; Lo and Shipman, 1990). If the main purpose is simply evaluating land use/land cover change in the study area, undertaking image-to-image registration is often considered sufficient (Fung, 1992; Fung and LeDrew, 1987; Martin and Howarth, 1989; Pilon *et al.*, 1988) in which one of the images used is chosen as a reference. However, the method adopted in this application is that all the images used were registered to the NZMS grid system, using the same control points. Using this approach several advantages could be achieved. *Firstly,* as demonstrated in Chapter V, the digital cadastral data, available in a form already-rectified to the NZMS, were used to provide a binary mask which was used to improve the final classification map. Thus one can ensure that the thematic features extracted using the binary mask had accurate geographic positions; enabling these data sets to be interchanged.

Secondly, using this approach, the total RMS errors obtained were 0.46, 0.41, and 0.42 pixels, respectively, for the 1968-SAP, 1993-SAP, and 1990-SPOT XS imagery. A similar procedure was followed by Lo and Shipman (1990) who demonstrated that, after two images had been registered to the UTM (Universal Transverse Mercator) projection with total RMS errors of 1.2 and 1.3 pixels, the geometric conformance between the two registered images was then found to be within 0.5 pixels. Therefore, it ensured that this application met the standard prerequisite specified by Jensen (1986) who pointed out that for accurate change detection, rectification should result in the two images being within one-fourth to one-half a pixel of registration one to another. If misregistration is greater than one pixel, this may bring about the identification of spurious areas of change between the two data sets. *Finally,* it was desirable to present all the maps produced in this study so that they conformed to the local grid system.

The importance of accurate classification for each image in a post-classification change detection analysis was emphasized earlier. The classification methods developed in this study were aimed at obtaining land use/land cover maps with an acceptable accuracy for change detection analysis. With the application of post-classification refinements (Chapter V Section 5.6), one can ensure that classification accuracies using a binary mask- i.e., 87.8 and 82.3 percent for the 1968 and 1993 SAP images, respectively- could be improved considerably. Thus the accuracy of the final classification maps produced was better than the 85 percent minimum standard specified by Anderson *et al.* (1976).

7.2 Conclusion

Based on the experience of this exploratory study, digital image analysis proved feasible for classifying land use/land cover types in the study area, and provided rapid and reliable results of category discrimination. Even though the study area was composed of intricate land use/land cover patterns in terms of "spectral reflectance," the problems of spectral confusion among categories could be resolved by incorporating information derived from visual interpretation, with the assistance of a binary masking strategy which made use of the digital cadastral data sets. An accurate geometric registration between images as well as data integration provided an excellent way for the various data sets to be interchanged and/or merged from one to another using the GIS-based overlay functions. The results obtained from an accuracy assessment revealed that the use of such procedures resulted in a significant improvement of the overall levels of classification accuracy.

Some difficulties which were encountered in mapping individual land use/land cover types in the study area included the limitations of the scanned aerial photograph comprising only a single band- black and white, the small size of cultural features (i.e., residential and commercial sites), and the high level of within-class variability of orchards and market gardening. In answer to these shortcomings, some knowledge of traditional photo-interpretation which makes use of a greater variety of discriminatory criteria is considered to be an essential prerequisite for the optimal application of computer-assisted classification of land use/land cover in the study region. Likewise, an appropriate site inspection still needs to be carried out to provide a more reliable way of selecting either training sites or sample points for accuracy assessment, and modifying the results obtained from digital classification. The knowledge of a crop calendar (particularly for market gardens) is also considered a valuable aid in the analysis procedure.

The microcomputer-based GIS-IDRISI proved to be a reliable tool with which to develop a quantitative land use/land cover change assessment and to identify spatial locations of changes on a category-by-category basis. A major advantage of the method adopted is its capability of performing image overlaying in various ways to highlight land use/land cover changes rapidly, accurately, and at low cost. The use of a binary change mask in the post-classification change detection yielded valuable information about the spatial location of land use/land cover changes that have taken place in the study area. The results obtained were largely dependent upon the accuracy of geometric registration between images and the classification accuracies of individual multi-date images.

From the results of the change detection analysis it was shown that a considerable change has taken place in the study region quantitatively and spatially for both land use/land cover types and property boundaries. The greatest land use gain occurred in orchards which, together with other land uses, were developed mainly on the pastoral farming land; leading to a significant reduction of the area for grazed pasture. A great number of larger land tracts have been subdivided to smaller sizes ranging from 4 to 10 ha- i.e., the most popular and favoured size for 'lifestyle blocks' as well as properties where orchards and market gardens are located.

Specifically for the study area, the use of scanned aerial photographs at an appropriate scale and complemented by a wealth of additional site information is considered sufficient for computer-assisted classification of land use/land cover types and further for evaluating their changes over a particular time gap. This was supported by the fact that: (a) the classification results produced were comparable to those derived from the three bands-SPOT XS imagery; (b) it requires far less investment compared to using satellite imagery; and (c) given the flat nature of the study region the problem of relief displacement, which is a common obstacle of using aerial photograph imagery in the area with complex topographic units, could be eliminated. However, it should be noted that for a particular region, data acquisition of aerial photographs could be limited; makes the information needed for desired dates or time intervals not be available. For this particular case, it is possible to use satellite imagery as one of the image pair for monitoring land use/land cover changes. Approaches demonstrated in this study will be of value for similar applications using the technique of merging data sets derived from different sensors with different spatial resolution.

REFERENCES

Abiodun, A.A., Camacho, S., and Chernicov, S. 1990. Thematic mapper data applications: experience in a number of developing countries. *Geocarto International. Vol.5, No.1*. pp.3-7.

Adeniyi, P.O. 1985. Digital analysis of multitemporal Landsat data for land use/land cover classification in a semi-arid area of Nigeria. *Photogrammetric Engineering and Remote Sensing. Vol.51, No.11*. pp.1761-1774.

Anderson, J.R. 1977. Land use and land cover changes- A framework for monitoring. *Journal of Research U.S. Geologic Survey. Vol.5, No.2*. pp.143-153.

Anderson, J.R., Hardy, E.E., Roach, J.T. 1972. Land use classification system for use with remote sensor data. *US. Geological Survey Circular 671*. Washington D.C.

Anderson, J.R., Hardy, E.E., Roach, J.T., and Witmer, R.E. 1976. A land use and land cover classification system for use with remote sensor data. *US. Geological Survey Professional Paper 964*. Washington D.C.

Angelici, G.L., Brynt, N.A., and Friedman, S.Z. 1977. Techniques for land use change detection using Landsat imagery. *Proceedings of the American Society of Photogrammetry*. pp.217-228.

Aronoff, S. 1991. *Geographic Information Systems: A Management Perspective*. WDL Publ. Ottawa.

Astaras, T., and Silleos, N. 1984. Land classification of part of Central Macedonia (Greece) by the use of remote sensing techniques. *International Journal of Remote Sensing. Vol.5*. pp.289-302.

Baker, R.D., de Steiguer, J.E., Grant, D.E., Newton, M. 1979. Land use/land cover mapping from aerial photographs. *Photogrammetric Engineering and Remote Sensing. Vol.45, No.5*. pp.661-668.

Barret, E.C. and Curtis, L.F. 1992. *Introduction to Environmental Remote Sensing. 3rd Edition.* Chapman and Hall. New York.

Benny, A.H. 1983. Automatic relocation of ground control points in Landsat imagery. *International Journal of Remote Sensing. Vol.4, No.2.* pp.335-342.

Bernstein, R. 1983. Image geometry and rectification. (Chapter 21). Colwell, R.N., Simonet, D.S., Ulaby, F.T. (eds). *Manual of Remote Sensing. Vol.1. 2nd Edition.* American Society of Photogrammetry. pp.873-922.

Berry, J.K. 1988. Maps as data: fundamental considerations in computer-assisted map analysis. *GIS/LIS'88 Proceedings: Accessing the World. Third Annual International Conference, Exhibits and Workshops. Vol.1.* pp.273-284.

Brougham, G.G. 1976. *Shingle Management Report on the Otaki River.* Manawatu Catchment Board and Regional Water Board. New Zealand. 5 pp.

Burrough, P.A. 1989. *Principles of Geographic Information Systems for Land Resources Assessment.* Monographs on Soil and Resources Survey. No.12. Clarendon Press. Oxford.

Byrne, G.F., Crapper, P.F., and Mayo, K.K. 1980. Monitoring land cover change by principal component analysis of multitemporal Landsat data. *Remote Sensing of Environment. Vol.10.* pp.175-184.

Campbell, J.B. 1983. *Mapping the Land: Aerial Imagery for Land Use Information.* Association of American Geographers. Washington D.C.

Caroll, D.M., Evans, L. and Bendelow, V.C. 1977. *Airphotos-interpretation for soil mapping*. Rothomsted Experimental Station. Harpenden. 85 pp.

Chen, L. and Lee, L. 1992. Progressive generation of control frameworks for image registration. *Photogrammetric Engineering and Remote Sensing. Vol.58, No.9.* pp.1321-1328.

Christensen, E.J., Jensen, J.R., Ramsey, E.W., Mackey, H.E. 1988. Aircraft MSS data registration and vegetation classification for wetland change detection. *International Journal of Remote Sensing. Vol.9, No.1.* pp.23-38.

Chuvieco, E., and Vega, J.M. 1990. Visual versus digital analysis for vegetation mapping: Some examples on Central Spain. *Geocarto International. Vol.5, No.3.* pp.21-30.

Claasen, D. van R. 1992. The utilisation of remote sensing in the South Pacific. *ACIAR Technical Reports 19.* Canberra. 59 pp.

Clarke, K.C. 1986. Advances in geographic information systems. *Computers, Environment and Urban Systems. Vol.10.* pp.175-184

Coleman, T.L. 1992. Three-dimensional modelling of an image-based GIS to aid land use planning. *Geocarto International. Vol.7, No.4.* pp.47-53.

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.

Cowie, D.J., and Wide, R.H. 1980. Suitability of the soils of the Hautere Plains, Otaki for irrigation. *Soil Bureau District Office Report PN 1. DSIR New Zealand.* 15 pp.

Cowie, D.J., and Wilde, R.H. 1981. Soil of the Te Horo area, Otaki, and their suitability for irrigation. *Soil Bureau District Office Report PN 2. DSIR. New Zealand.* 17 pp.

Cowen, D.J. 1987. GIS vs CAD vs DBMS: What are the differences? *GIS'87-San Francisco Second Annual International Conference, Exhibits and Workshops in Geographic Information Systems. Vol.1.* Berkeley. pp.46-56.

Cracknell, A.P. 1981. Background: The physical basis of remote sensing. (Chapter 1). Cracknell, A.P. (ed.). *Remote Sensing in Meteorology, Oceanography and Hydrology.* John Wiley and Sons. New York. pp.27-44.

Curran, P.J. 1985. *Principles of Remote Sensing.* Longman Scientific and Technical. New York.

Dangermond, J. 1993. A classification of software components commonly used in geographic information systems. (Chapter 3). Peuquet, D.J. and Marble, D.J. (eds). *Introductory Readings in Geographic Information Systems.* Taylor & Francis. New York. pp.31-51.

Davidson, G.J. 1986. Ground control pointing and geometric transformation of satellite imagery. *International Journal of Remote Sensing. Vol.7, No.1.* pp.65-74.

Druce, A.P. 1962. Secondary totara-matai-titoki forest on the Otaki Plain. *Bulletin of the Wellington Botanical Society. Vol.33.* pp.22-39.

Dymond, J. 1991. *Digital Rectification of Aerial and Satellite Imagery.* (Unpublished). DSIR. Land Resources. Palmerston North. New Zealand. 9 pp.

Eastman, J.R., and Fulk, M. 1993. Long sequence time series evaluation using standardized principal components. *Photogrammetric Engineering and Remote Sensing. Vol.59, No.8.* pp.1307-1312.

Eastman, J.R., Kyem, P.A.K., Toledano, J., and Jin, W. 1993. *Exploration in Geographic Information Systems Technology: GIS and Decision Making. Vol.4.* United Nations Institute for Training and Research. 112 pp.

Ehlers, M. 1990. Application of SPOT data for regional growth analysis and local planning. *Photogrammetric Engineering and Remote Sensing. Vol.56, No.7.* pp.175-180.

Ehlers, M., Edwards, G., and Bedard, Y. 1989. Integration of remote sensing with geographic information systems: A necessary evaluation. *Photogrammetric Engineering and Remote Sensing. Vol.55, No.11.* pp.1619-1627.

Estes, J.E., Hajic, E.J. and Tinney, L.R. 1983. Fundamentals of image analysis: Analysis of visible and thermal infrared data. (Chapter 24). Colwell, R.N., simonet, D.S., Ulaby, F.T. (eds). *Manual of Remote Sensing. Vol.1. 2nd Edition.* Am. Soc. of Photogrammetry and Remote Sensing. pp.987-989.

Estes, J.E., Stow, D., and Jensen, J.R. 1982. Monitoring land use and land cover changes. (Chapter 10). Johannsen, J.C., and Sanders, J.L. (Eds). *Remote Sensing for Resource Management.* Soil Conservation Society of America. pp.100-110.

Fischer, M.M. and Nijkamp, P. 1993. Design and use of geographic information systems and spatial models. Fischer, M.M. and Nijkamp, P. (Eds).

Geographic Information Systems, Spatial Modelling and Policy Evaluation. Springer-Verlag. Berlin. pp.1-13.

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.

Ford, G.E. and Zanelli, C.I. 1985. Analysis and quantification of errors in the geometric correction of satellite images. *Photogrammetric Engineering and Remote Sensing. Vol.51, No.11.* pp.1725-1734.

Fung, T. 1992. Land use and land cover change detection with Landsat MSS and SPOT HRV data in Hongkong. *Geocarto International. Vol.7, No.3.* pp.33-40.

Fung, T. and LeDrew, E. 1987. Application of principal component analysis to change detection. *Photogrammetric Engineering and Remote Sensing. Vol.53, No.12.* pp.1649-1658

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.

Gates, D.M. 1970. Physical and physiological properties of plants. (Chapter 5). Shay, J.R. (ed). *Remote Sensing with Special Reference to Agriculture and Forestry.* National Academy of Sciences. Washington D.C.

Gordon, S. 1980. Utilizing Landsat imagery to monitor land use change: a case study of Ohio. *Remote Sensing of Environment. Vol.9.* pp.189-196.

Green, W.B. 1989. *Digital Image Processing: A System Approach. 2nd Edition.* Van Nostrand Reinhold. New York.

Haralick, R.M. 1973. Glossary and index to remotely sensed image pattern recognition concepts. *Pattern Recognition. Vol.5.* pp.391-403.

Hardy, E.E. 1982. Remote sensing for land and water resource management. (Chapter 1). Johannsen, C.J., and Sanders, J.L. (eds). *Remote Sensing for Resource Management.* Soil Conservation Society of America. pp.3-15.

Harrison, B.A. and Jupp, D.L.B. 1990. *Introduction to Remotely Sensed Data.* Division of Water Resources CSIRO. Australia.

Hixson, M., Scholz, D., Fuhs, N., and Akiyama, T. 1980. Evaluation of several schemes for classification of remotely sensed data. *Photogrammetric Engineering and Remote Sensing. Vol.46, No.12.* pp.1547-1553.

Hoffer, R.M. 1978. Biological and physical consideration in applying computer aided analysis techniques to remote sensor data. (Chapter 5). Swain, P.H. and Davis, S.M. (eds). *Remote sensing: The Quantitative Approach.* McGraw-Hill. New York.

Hord, R.M. 1986. *Remote Sensing: Methods and Application.* John Wiley & Sons. New York.

Howarth, J.P., and Boasson, E. 1983. Landsat digital enhancement for change detection in urban environment. *Remote Sensing of Environment. Vol.13.* pp.149-160.

Howarth, J.P., and Wickware, G.M. 1981. Procedure for change detection using Landsat digital data. *International Journal of Remote Sensing. Vol.2.* pp.277-291.

Hutchinson, C.F. 1982. Techniques for combining Landsat and ancillary data for digital classification improvement. *Photogrammetric Engineering and Remote Sensing. Vol.48, No.1.* pp.123-130.

Janssen, L.L.F., Jaarsma, M.N., and van der Linden. 1990. Integrating topographic data with remote sensing for land cover classification. *Photogrammetric Engineering and Remote Sensing. Vol.56, No.11.* pp.1503-1506.

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.

Jensen, J.R., Bryan, M.L., Friedman, S.Z., Henderson, F.M., Holz, R.K., Lindgren, D., Toll, D.L., Welch, R.A., Wray, J.R. 1983. Urban/suburban land use analysis. (Chapter 30). Colwell, R.N., Estes, J.E., and Thorley, G.A. (eds). *Manual of Remote Sensing. Vol.2. 2nd edition.* American Society of Photogrammetry. pp. 1571-1666.

Jensen, J.R. 1986. *Introductory Digital Image Processing: A Remote Sensing Perspective.* Prentice-Hall, Englewood Cliffs. New Jersey.

Johannsen, C.J. and Sanders, J.L. (eds). 1982. *Remote Sensing for Resource Management.* Soil Conservation Society of America.

Kenk, E., Sondheim, M., and Yee, B. 1988. Methods for improving accuracy of Thematic Mapper ground cover classification. *Canadian Journal of Remote Sensing. Vol.14, No.1.* pp.17-31.

Landgrebe, D.A. 1980. The development of a spectral-spatial classifier for earth observational data. *Pattern Recognition. Vol.12.* pp.165-172.

Langran, G. 1992. *Time in Geographic Information Systems.* Taylor & Francis. New York.

Lillesand, T.M. and Kiefer, R.W. 1987. *Remote Sensing and Image Interpretation. 2nd Edition.* John Wiley and Sons. New York.

Lillestrand, R.L. 1972. Techniques for change detection. *I.E.E.E. Transactions on Computers. Vol.21.* pp.654-659.

Lo, C.P. 1986. *Applied Remote Sensing.* Longman Inc. New York.

Lo, C.P., and Shipman, R.L. 1990. A GIS approach to land-use change dynamics detection. *Photogrammetric Engineering and Remote Sensing. Vol.56, No.11.* pp.1483-1491.

Lusch, D.P. 1989. Fundamental considerations for teaching the spectral reflectance characteristics of vegetation, soil and water (Chapter 1). Nellis, M.D., Loungeay, R. and Lulla, K. (eds). *Current Trends in Remote Sensing Education.* Geocarto International Centre. Hongkong. pp.5-23.

Maguire, D.J., Kimber, B., and Chick, J. 1991. Integrated geographic information systems: the importance of raster. *ACSM/ASPRS Annual Convention Technical Paper. Vol.4.* pp.107-116.

Marble, D.F. 1993. Geographic information systems: An overview. (Chapter 1). Peuquet, D.J., and Marble, D.J. (eds). *Introductory Readings in Geographic Information Systems.* Taylor & Francis. New York. pp.9-17.

Marble, D.F., and Peuquet, D.J. 1983. Geographic information systems and remote sensing. (Chapter 22). Colwell, R.N., Simonet, D.S., Ulaby, F.T. (eds). *Manual of Remote Sensing. Vol.1. 2nd edition.* American Society of Photogrammetry. pp.923-958.

Martin, L.R.G. and Howarth, P.J. 1989. Change detection accuracy assessment using SPOT multispectral imagery of the rural-urban fringe. *Remote Sensing of Environment. Vol.30.* pp.55-66.

Mason, D.C., Corr, D.G., Cross, A., Hogg, D.C., Lawrence, D.H., Petrow, M., and Tailor, A.M. 1988. The use of digital map data in the segmentation and classification of remotely sensed images. *International Journal of Geographical Information Systems. Vol.2, No.3.* pp.195-215.

McKendry, J.E., Eastman, J.R., Martin, K.S., and Fulk, M.A. 1992. *Explorations in Geographic Information Systems Technology: Applications in Forestry. Vol.2.* United Nations Institute for Training and Research. Switzerland. 157 pp.

Mead, R.A., and Szajgin, J. 1982. Landsat classification accuracy assessment procedures: an account of a national working conference. *Photogrammetric Engineering and Remote Sensing. Vol.48, No.1.* pp.139-141.

Milne, A.K. 1988. Change detection analysis using Landsat imagery: A review of methodology. *Proceedings of IGARSS'88 Symposium.* Edinburgh. Scotland. pp.541-544.

Milne, A.K., and O'Neill, A.L. 1990. Mapping and monitoring land cover in the Willandra Lakes World Heritage region. *International Journal of Remote Sensing. Vol.11, No.11.* pp.2035-2049.

Moreton, G.E. 1984. Irrigated crop inventory by classification of satellite image data. *Photogrammetric Engineering and Remote Sensing. Vol.50, No.6.* pp.729-737.

Mouat, D.A., Mahin, G.G., and Lancestar, J. 1993. Remote Sensing technique in the analysis of change detection. *Geocarto International. Vol.8, No.2.* pp.39-50.

Nelson, R.F. 1983. Detecting forest canopy change due to insect activity using Landsat MSS. *Photogrammetric Engineering and Remote Sensing. Vol.49, No.2.* pp.1303-1314.

Nguyen, P.T. and Ho, D. 1988. Multiple source data processing in remote sensing. (Chapter 9). Muller, J.P. (ed). *Digital Image Processing in Remote Sensing.* Tailor & Francis. London. pp.135-176.

Nichols, D. 1983. Digital hardware. (Chapter 20). Colwell, R.N., simonet, D.S., Ulaby, F.T. (eds). *Manual of Remote Sensing. Vol.1. 2nd Edition.* American Society of Photogrammetry and Remote Sensing. pp.841-873.

Nicolaos, S. 1988. Visual and digital classification of Landsat TM data for soil, physiography and land use mapping in Axios alluvial plain, Thessaloniki, Greece. *Geocarto International. Vol.3, No.4.* pp.55-66.

Palmer, A.S., Barnett, R., Pillans, B.J., and Wilde, R.H. 1988. Loess, river aggradation terraces and marine benches at Otaki, southern North Island. Eden, D.N., and Furkert, R.J. (eds). *Loess: Its Distribution, Geology and Soils.* A.A. Balkema. pp.163-174.

Palmer, A.S., and Wilde, R.H. 1986. *Soil Definitions and Interpretations for Otaki District Soil Resources Study, North Island New Zealand.* (Unpublished). 169 pp.

Parker, H.D. 1988. The unique quantities of a geographic information system: a commentary. *Photogrammetric Engineering and Remote Sensing. Vol.54, No.13.* pp.1547-1549.

Peuquet, D.J. and Marble, D.F. 1993. ARC/INFO: An example of a contemporary geographic information system. Peuquet, D.J. and Marble, D.F. (eds). *Introductory Readings in Geographic Information Systems.* Taylor & Francis. New York. pp.90-99.

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.

Rhind, D.W. and Green, N.P.A. 1988. Design of a geographic information system for a heterogeneous scientific community. *International Journal of Geographic Information Systems. Vol.2, No.2.* pp.171-189.

Richards, J.A. 1986. *Remote Sensing Digital Image Analysis: An Introduction.* Springer-Verlag. Berlin.

Richards, J.A. 1993. *Remote Sensing Digital Image Analysis: An Introduction. 2nd, Revised and Enlarged Edition.* Springer-Verlag. Berlin.

Richards, J.A., Landgrebe, D.A., and Swain, P.H. 1982. A means for utilizing ancillary information in multispectral classification. *Remote Sensing Environment. Vol.12.* pp.463-477.

Rutchey, K., and Vilcheck, L. 1994. Development of an Everglades vegetation map using a SPOT image and the global positioning system. *Photogrammetric Engineering and Remote Sensing. Vol.60, No.6.* pp.767-775.

Sadar, S.A., Linden, D.S., and McGuire, M. 1982. Fuels mapping from Landsat imagery and digital terrain data and fire suppression decision. *Proceedings of ACSM-ASP Congress.* pp.345-351.

Satterwhite, M., Rice, W., and Shipman, 1984. Using landform and vegetative factors to improve the interpretation of Landsat imagery. *Photogrammetric Engineering and Remote Sensing. Vol.50, No.1.* pp.83-91.

Simcock, R.C. 1986. *A Study of Land Use and Land Development with Reference to Soil Types in the Hautere Plains North Island, New Zealand.* B.Agr.Sc Honours Project (Unpublished). 92 pp.

Singh, A. 1989. Digital change detection techniques using remotely sensed data. *International Journal of Remote Sensing. Vol.10, No.6.* pp. 989-1003.

Smith, T.R., Menon, S., Star, J.L., and Estes, J.E. 1987. Requirements and principles for the implementation and construction of large-scale geographic information systems. *International Journal of Geographic Information Systems. Vol.1, No.1.* pp.13-31.

Star, J. and Estes, J.E. 1990. *Geographic Information System: An Introduction.* Prentice Hall, Englewood Cliff. New Jersey.

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.

Suits, G.H. 1983. The nature of electromagnetic radiation. (Chapter 2). Colwell, R.N., simonet, D.S., Ulaby, F.T. (eds). *Manual of Remote Sensing. Vol.1. 2nd edition.* American Society of Photogrammetry and Remote Sensing. pp.37-60.

Swain, P.H., and Davis, S.M. (eds). 1978. *Remote Sensing: The Quantitative Approach.* McGraw-Hill. New York.

Thomas, I.L., Benning, V.M., and Ching, N.P. 1978. *Classification of Remotely Sensed Images.* Adam Hilger, Bristol. England.

Thomas, I.L. and Timmins, S.M. 1982. Spectral characteristics of vegetation and the generation of band ratios and principal components. *Physical Engineering Laboratory Report No. 773.* DSIR. New Zealand. 21 pp.

van der Laan, F.B. 1988. Improvement of classification results from a satellite image using context information from a topographic map. *Proceedings of the Sixteenth International Congress of ASPRS. Vol.27.* pp.98-108.

Walsh, S.J. 1987. Geographic information systems for natural resource management. Ripple, W.J. (ed.). *Geographic Information Systems for Resource Management: A Compendium.* American Society of Photogrammetry and Remote Sensing. pp.190-193.

Weismiller, R.A., Kristof, S.J., Scholz, D.K., Anuta, P.E., and Momin, S.M. 1977. Evaluation of change detection techniques for monitoring coastal zone environments. *Proceedings of the Eleventh International Symposium on Remote Sensing of Environment.* Ann Arbor, Michigan. pp. 1229-1238.

Westmoreland, S. and Stow, D.A. 1992. Category identification of changed land-use polygons in an integrated image processing/Geographic Information System. *Photogrammetric Engineering and Remote Sensing. Vol.58, No.11.* pp.1593-1599.

APPENDICES

Appendix 1 Corresponding coordinates used for image registration.

1. Corresponding coordinates used for 1968-SAP image registration.

16			
151	270	2689736	6043487
78	342	2689160	6044511
349	276	2691946	6042934
461	276	2693170	6042581
213	115	2689920	6041580
180	456	2690659	6045450
404	208	2692328	6042014
314	488	2692240	6045377
257	213	2690728	6042531
89	460	2689674	6045785
282	337	2691405	6043821
467	383	2693592	6043738
105	170	2688903	6042540
336	128	2691340	6041345
363	536	2692930	6045753
266	437	2691552	6044967

2. Corresponding coordinates used for 1993-SAP image registration.

16			
236	373	2689736	6043487
127	457	2689160	6044511
502	416	2691946	6042934
651	434	2693170	6042581
344	174	2689920	6041580
246	627	2690659	6045450
586	335	2692328	6042014
421	693	2692240	6045377
388	316	2690728	6042531
124	618	2689674	6045785
402	487	2691405	6043821
643	579	2693592	6043738
190	232	2688903	6042540
508	215	2691340	6041345
478	765	2692930	6045753
365	617	2691552	6044967

Appendix 1 Continued.

3. Corresponding coordinates used for 1990-SPOT XS image registration.

9			
299	263	2690728	6042531
211	409	2689674	6045785
317	333	2691405	6043821
425	353	2693592	6043738
210	242	2688903	6042540
342	211	2691340	6041345
370	444	2692930	6045753
312	391	2691552	6044967
457	265	2693830	6041864

Appendix 2 The procedure of format conversion (from IDRISI to WORD PERFECT GRAPHIC) to create Figure 5.8.

1. The vector sample point file generated using the SAMPLE routine in IDRISI was first exported to an UNGEN-ASCII file using the ARCIDRIS function.

2. Using the GENERATE module in the PC ARC/INFO, the UNGEN file was converted to a point coverage. To maintain the position of these sample points at the appropriate geographic coordinates (i.e., NZMS reference system), the COPYCOV command was employed to copy the TICS and BOUNDARY files of cadastral coverage. Using the GENERATE POINTS command and the UNGEN file name the empty coverage was then filled with point features.

3. In the ARCEDIT module several easily-recognizable features such as roads and the boundary of study area were added with reference to the cadastral data. This ensured this line coverage had the same TICS and BOUNDARY as the point coverage created in step 2.

4. Both point and line coverages were then overlaid in ARCPLOT to create the PLOT file.

5. Using the PLOT function, the plot file (created in step 4) was converted to an HPGL file compatible with DRAW PERFECT. The line command used was:

   C:\>[ARC] Plot infile.plt 1 h a 0 > outfile.hpg

   where,
   Infile.plt = input file created in step 4
   Outfile.hpg = output file in HPGL format
   1 = scale, the same as that of the infile.plt
   h = the plot file will be displayed on Hewlett-packard plotters that support HPGL

Appendix 2 Continued.

a = advance; i.e., the plotter will advance the paper after plotting to begin a new plot
0 = port 0; i.e., direct output to the standard output which is usually the PC monitor
> = DOS redirection meaning 'export to'

6. The outfile.hpg was then displayed on the DRAW PERFECT screen and saved as .WPG file which is compatible with WORD PERFECT.

7. It should be noted that this procedure provided a basic method of converting any point and line coverages in ARC/INFO to WORD PERFECT GRAPHIC.

Appendix 3 Error matrices for the land use/land cover maps generated from two image sources (SAP and SPOT XS), classified using unsupervised and supervised decision rules.

1. 1968 SAP-Unsupervised (SP68U)

		"Ground truth" (based on the visual interpretation)								
	Category	Ps	Mr	Tr	Rs	Cm	Rd	Gr	Total	ErrorC
M	Ps	165	1	3	1	0	5	0	175	0.06
a	Mr	1	7	0	2	1	4	0	24	0.71
p p	Tr	1	0	23	0	0	0	0	38	0.39
e d	Gr	0	0	0	0	0	0	1	1	0.00
	Total	190	8	26	3	1	9	1	238	
	ErrorO	0.13	0.13	0.12	1.00	1.00	1.00	1.00		**0.18**

2. 1968 SAP-Unsupervised+mask (SP68UM)

		"Ground truth" (based on the visual interpretation)								
	Category	Ps	Mr	Tr	Rs	Cm	Rd	Gr	Total	ErrorC
M	Ps	165	1	3	0	0	0	0	169	0.02
a	Mr	10	7	0	0	0	0	0	17	0.59
p	Tr	15	0	23	0	0	0	0	38	0.39
p	Rs	0	0	0	3	0	0	0	3	0.00
e	Cm	0	0	0	0	1	0	0	1	0.00
d	Rd	0	0	0	0	0	9	0	9	0.00
	Gr	0	0	0	0	0	0	1	1	0.00
	Total	190	8	26	3	1	9	1	238	
	ErrorO	0.13	0.13	0.12	0.00	0.00	0.00	0.00		**0.12**

Appendix 3 Continued.

3. 1993 SAP-Unsupervised (SP93U)

	"*Ground truth*" (based on the visual interpretation)										
	Category	Ps	Or	Mr	Tr	Rs	Cm	Rd	Gr	Total	ErrorC
M	Ps	144	17	7	0	3	2	6	2	181	0.20
a p	Or	1	14	0	9	0	0	2	0	26	0.46
p e	Mr	1	0	5	0	0	2	1	3	12	0.58
d	Tr	0	6	0	13	0	0	0	0	19	0.32
	Total	146	37	12	22	3	4	9	5	238	
	ErrorO	0.01	0.62	0.58	0.41	1.00	1.00	1.00	1.00		0.26

4. 1993 SAP-Unsupervised+masked (SP93UM)

	"*Ground truth*" (based on the visual interpretation)										
	Category	Ps	Or	Mr	Tr	Rs	Cm	Rd	Gr	Total	ErrorC
	Ps	142	15	5	0	0	0	0	0	162	0.12
M	Or	2	12	0	7	0	0	0	0	21	0.43
a	Mr	2	0	7	0	0	0	0	0	9	0.22
p	Tr	0	10	0	15	0	0	0	0	25	0.40
p	Rs	0	0	0	0	3	0	0	0	3	0.00
e	Cm	0	0	0	0	0	4	0	0	4	0.00
d	Rd	0	0	0	0	0	0	9	0	9	0.00
	Gr	0	0	0	0	0	0	0	5	5	0.00
	Total	146	37	12	22	3	4	9	5	238	
	ErrorO	0.03	0.68	0.42	0.32	0.00	0.00	0.00	0.00		0.17

Appendix 3 Continued.

5. 1990 SPOT XS-Unsupervised (XS90U)

		"Ground truth" (based on the visual interpretation)									
	Category	Ps	Or	Mr	Tr	Rs	Cm	Rd	Gr	Total	ErrorC
M a	Ps	128	12	10	2	1	2	7	1	163	0.21
p	Or	9	17	0	7	1	0	2	0	36	0.53
p	Mr	0	0	2	0	1	2	0	3	8	0.75
e d	Tr	9	9	0	13	0	0	0	0	31	0.58
	Total	146	38	12	22	3	4	9	4	238	
	ErrorO	0.12	0.55	0.83	0.41	1.00	1.00	1.00	1.00		**0.33**

6. 1990 SPOT XS-Unsupervised+masked (XS90UM)

		"Ground truth" (based on the visual interpretation)									
	Category	Ps	Or	Mr	Tr	Rs	Cm	Rd	Gr	Total	ErrorC
	Ps	128	12	10	2	0	0	0	0	152	0.16
M	Or	9	17	0	7	0	0	0	0	33	0.48
a	Mr	0	0	2	0	0	0	0	0	2	0.00
p	Tr	9	9	0	13	0	0	0	0	31	0.58
p	Rs	0	0	0	0	3	0	0	0	3	0.00
e	Cm	0	0	0	0	0	4	0	0	4	0.00
d	Rd	0	0	0	0	0	0	9	0	9	0.00
	Gr	0	0	0	0	0	0	0	4	4	0.00
	Total	146	38	12	22	3	4	9	4	238	
	ErrorO	0.12	0.55	0.83	0.41	0.00	0.00	0.00	0.00		**0.24**

Appendix 3 Continued.

7. 1990 SPOT XS-Supervised (XS90S)

	"*Ground truth*" (based on the visual interpretation)										
	Category	Ps	Or	Mr	Tr	Rs	Cm	Rd	Gr	Total	ErrorC
M	Ps	120	10	6	1	1	2	1	0	141	0.15
a	Or	12	19	2	4	0	0	2	0	39	0.51
p	Mr	0	0	4	0	1	0	0	2	7	0.43
p	Tr	14	9	0	17	1	0	0	0	41	0.59
e	Rd	0	0	0	0	0	0	6	0	6	0.00
d	Gr	0	0	0	0	0	2	0	2	4	0.50
	Total	146	38	12	22	3	4	9	4	238	
	ErrorO	0.18	0.50	0.67	0.23	1.00	1.00	0.33	0.50		**0.29**

8. 1990 SPOT XS-Supervised+masked (XS90SM)

	"*Ground truth*" (based on the visual interpretation)										
	Category	Ps	Or	Mr	Tr	Rs	Cm	Rd	Gr	Total	ErrorC
	Ps	124	10	4	1	0	0	0	0	139	0.11
M	Or	11	19	1	4	0	0	0	0	35	0.46
a	Mr	0	0	7	0	0	0	0	0	7	0.00
p	Tr	11	9	0	17	0	0	0	0	37	0.54
p	Rs	0	0	0	0	3	0	0	0	3	0.00
e	Cm	0	0	0	0	0	4	0	0	4	0.00
d	Rd	0	0	0	0	0	0	9	0	9	0.00
	Gr	0	0	0	0	0	0	0	4	4	0.00
	Total	146	38	12	22	3	4	9	4	238	
	ErrorO	0.15	0.50	0.42	0.23	0.00	0.00	0.00	0.00		**0.21**

Note: 1. Ps = Pasture
Or = Orchards
Mr = Market gardens
Tr = Trees
Rs = Residential sites
Cm = Commercial sites
Rd = Roads
Gr = Gravel pits
ErrorC = Errors of commission (expressed as proportions
ErrorO = Errors of omission (expressed as proportions).

2. Values underlined indicate the number of pixels that were correctly classified.