

**A surface-based approach to the handling of uncertainties  
in an urban-orientated spatial database**

**Thesis presented in partial fulfilment of the requirements of  
the University of Edinburgh for the degree of  
Doctor of Philosophy (Ph.D.)**

**Jingxiong Zhang  
Department of Geography  
The University of Edinburgh**

**September 1996**



## **Declaration**

I declare that this thesis represents my own work, and that where the work of others has been used it has been accredited.

## Acknowledgements

The author is extremely grateful to his departmental supervisors, Dr. R. Kirby and Dr. N. Stuart, for their inspiring encouragement and advice. In addition, many other people at the University of Edinburgh have been helpful with respect to the research ideas, the computing aspects and the field work, in particular, Dr. P. Furley, Mr. S. Dowers, Mr. C. Place, Miss C. Jarvis and Mr. Wan Mohd.

The author's Ph.D. work is funded by an SBFSS (Sino-British Friendship Scholarship Scheme) award, which was established by Sir Yue-Kong Pao Foundation (Hong Kong), Chinese and British governments in 1986, and administered by The State Education Commission of China and The British Council.

The author's home university, Wuhan Technical University of Surveying and Mapping, has given great assistance with respect to his pre-departure for study in Edinburgh, and is readily helpful to enable him to be settled down to his work in Wuhan. The Ordnance Survey of Great Britain is acknowledged for the provision of National Grid reference coordinates.

Professor P. Burrough (University of Utrecht, The Netherlands), Dr. J. Drummond (University of Glasgow), Professor M. Goodchild (NCGIA, USA), Dr. K. Lowell (University of Laval, Canada) and Dr. H. Veregin (Kent State University, USA) have all kindly offered valuable advice and comments towards the author's research.

Professor Andersen (Norwegian University of Agriculture), Dr. Carson (Carto Instrument Service, USA) and Mrs. Payne (Cartographic Engineering Ltd) offered prompt help in the use of the BLOKK package. Dr. Yates (USDA) supplied free the geostatistical software program GEOPACK for the author's research. Dr. Dungan

(NASA Ames Research Centre) provided advice about geostatistical methods and packages.

The author feels deeply indebted to Shuxia and Zhe for their increasing and lasting understanding. His extended families, especially his parent, have shown their profound affection and endurance during the author's absence from the hometown. Lastly but not the least, this thesis is dedicated to the loving memory of the author's late aunt Joan, uncles De-Yan, De-An and De-Li.

## Abstract

The work presented in this thesis is built on three assertions: (1) uncertainties should be perceived as integral components of GIS spatial databases; (2) as such and given the importance of uncertainties at all stages of processing spatial data by digital methods, an integrating strategy is needed to provide more direct access to the uncertainties of spatial data during data collection, update, spatial analysis and during the creation of output products; (3) surface-based models and methods are capable of such an integral strategy, by which many kinds of uncertainties of spatial data can be well represented and handled.

A cumulative description is given of various uncertainties occurring in geographical abstraction and spatial data acquisitions with special reference to one common area of geographical studies, that is, land cover mapping. Two alternative forms of geographical abstraction or spatial data modelling are introduced: discrete objects and continuous fields. The uncertainties are then discussed with respect to their description, estimation and representation under object and field perspectives. For categorical data, in particular, uncertainties are represented as fuzzy surfaces, whose derivation and analysis are described in detail.

To provide an evaluation of the integrated approach and to show how such an integrated strategy can be used to advantage, a case study is developed in the context of suburban land cover mapping, based on a local Edinburgh area. The case study begins with the construction of a co-registered hierarchy of test data with a corresponding hierarchy of accuracies, and continues to the generation and analysis of fuzzy surfaces using the suite of methods introduced previously. The various graphical maps and quantitative data produced show that surface based approaches are well suited to the representation and handling of uncertainties of spatial data, because they are effective and flexible.

# Contents

	Page
Chapter 1 Introduction	1
1.1 Background to the research topic	1
1.2 A review of the research on error issues	4
1.3 Problems from the past and the aims of this research	10
1.4 An overview of the contents of this thesis	16
Chapter 2 Uncertainties in spatial data modelling and acquisition	20
2.1 Introduction	20
2.2 Spatial data modelling and uncertainties	21
2.3 Spatial data acquisition and uncertainties	29
2.3.1 Land surveying	29
2.3.2 Photogrammetric techniques	31
2.3.3 Remote sensing	35
2.3.4 Secondary data acquisition	41
2.4 Discussion	43
Chapter 3 Object and field perspectives of uncertainties	46
3.1 Introduction	46
3.2 Uncertainties in objects	46
3.3 Uncertainties in fields	57
3.4 Representing uncertainties in objects and fields	61
3.5 Discussion	65
Chapter 4 Deriving and analysing fuzzy surfaces	68
4.1 Introduction	68
4.2 Deriving fuzzy surfaces from remotely sensed images	73
4.2.1 Supervised methods	73

4.2.2 Unsupervised methods	76
4.3 Generating fuzzy surfaces from photogrammetric data	80
4.3.1 Resolving proportions of component land cover types in mapping units	81
4.3.2 Interpolation approaches	82
4.4 Analysing fuzzy surfaces	89
4.5 Discussion	96
Chapter 5 Empirical study I: data and uncertainties in objects	100
5.1 Introduction	100
5.2 Study area	100
5.3 Data sources	103
5.4 Methods used to acquire test data in required form	110
5.4.1 Land surveying	114
5.4.2 Control densification using block adjustment based on 1:5,000 scale aerial photographs	116
5.4.3 Map digitising to supply control points for remote sensing data	120
5.4.4 Photogrammetric digitising of aerial photographs at 1:24,000 scale	122
5.4.5 Remote sensing image classification	126
5.4.6 Discussion of the results	137
5.5 Discussion	140
Chapter 6 Empirical study II: the surface approach to uncertainties	141
6.1 Introduction	141
6.2 Generation of fuzzy surfaces	143
6.2.1 Fuzzy surfaces built on remotely sensed data	143
6.2.2 Fuzzy surfaces generated from photogrammetric data	151

6.3 Analysing fuzzy surfaces	161
6.4 Deriving categorical maps from fuzzy surfaces	166
6.5 Evaluating the accuracies of fuzzy classified remote sensing data	172
6.6 Discussion	181
Chapter 7. Conclusions	185
7.1 Handling uncertainties in spatial databases	185
7.2 Prospects for the surface-based approach to uncertainties in GISs	192
References	196
Appendix 1. The list of ground control points by field surveying	207
Appendix 2. The list of densified control points by photogrammetric block adjustment	209
Appendix 3. The class statistics used in the supervised fuzzy classification described in section 6.2.1	212
Appendix 4. The FORTRAN 77 program for the fuzzy <i>c</i> -means clustering	214
Appendix 5. The FORTRAN 77 program for deriving proportions of sub-pixel component land cover types	223
Appendix 6. The FORTRAN 77 program for transforming point sample data to GSLIB data format	225
Appendix 7. The FORTRAN 77 program for transferring kriging outputs to ASCII format suitable for ERDAS LAN data files	227



## List of Tables

	Page
Table 1.1 Error issues and strategies for their analysis	5
Table 2.1 A classification and comparison of stereoplotters	33
Table 2.2 Optical and radar sensors on spaceborne platforms	36
Table 3.1 Definitions of different types of uncertainties in spatial data with examples in digital land cover mapping	48
Table 3.2 An error matrix for classification accuracy assessment	54
Table 3.3 A numerical example of Table 3.2	54
Table 3.4 A topological data structure for lines as drawn in Figure 3.1(a) (the 0s denote outer areas)	62
Table 3.5 An extended relational data structure for incorporating uncertainties of lines as drawn in Figure 3.1(a)	63
Table 4.1 A classification of methods for deriving fuzzy surfaces	72
Table 4.2 Initial and final U matrix (i.e., FMVs) for the pixels as shown in Figure 4.2	78
Table 4.3 Labelled pixels and their probabilities (based on Table 4.2 in Section 4.2)	93
Table 4.4 Pure pixels selected by slicing at a threshold of 60 % (based on Table 4.3)	93
Table 5.1 The data sources used to build the test data	106
Table 5.2 The hierarchy of test data	110
Table 5.3 The equipment for test data automation	114
Table 5.4 The main features of the analytical plotter AP190	117
Table 5.5 Photogrammetric block adjustment results (unit: metres)	119
Table 5.6 Test results of manual digitising on OS plans (unit: metres)	121
Table 5.7 Photogrammetric point digitising accuracy estimate (unit: metres)	124
Table 5.8 A summary of remote sensing data processing	131

Table 5.9 Error matrix for SPOT HRV data obtained by simple random sampling	132
Table 5.10 Error matrix for Landsat TM data obtained by simple random sampling	132
Table 5.11 Error matrix for SPOT HRV data obtained by superimposed comparison	135
Table 5.12 Error matrix for Landsat TM data obtained by superimposed comparison	135
Table 5.13 The summary result of evaluation of classification accuracies	136
Table 5.14 The hierarchies of test data and their accuracies	138
Table 6.1 The correlation coefficients ( $r$ ) resulting from semivariogram model fitting (omnidirection semivariograms unless otherwise specified, sample size = 24)	156
Table 6.2 The spherical model semivariogram parameters (omnidirection semivariograms unless otherwise specified, 1 unit = 2.5 metres)	158
Table 6.3 Correlation coefficient ( $r$ ) values between fuzzy classified data and their fuzzy reference data	166
Table 6.4 Error matrix of hardened fuzzy classified SPOT HRV data, using the original hard photogrammetric data as reference data	173
Table 6.5 Error matrix of hardened fuzzy classified Landsat TM data, using the original hard photogrammetric data as reference data	173
Table 6.6 Error matrix of hardened fuzzy classified SPOT HRV data, using the hardened fuzzy reference data derived from indicator kriging as the reference data	174
Table 6.7 Error matrix of hardened fuzzy classified Landsat TM data, using the hardened fuzzy reference data derived from indicator kriging as the reference data	174

Table 6.8 Error matrix of sliced fuzzy classified SPOT HRV data, using sliced fuzzy reference data as the reference data	176
Table 6.9 Error matrix of sliced fuzzy classified Landsat TM data, using sliced fuzzy reference data as the reference data	177
Table 6.10 Error matrix based on the soft comparison between fuzzy classified SPOT HRV data and fuzzy reference data derived from indicator kriging	178
Table 6.11 Error matrix based on the soft comparison between fuzzy classified Landsat TM data and fuzzy reference data derived from indicator kriging	178
Table 6.12 Summary results of the classification accuracies for the categorical maps derived in Section 6.4 using different methods of comparisons	180
Table 7.1 Object-based versus field-based approaches to GIS uncertainties	190

## List of Figures

	Page
Figure 2.1 NCDCDS classification of cartographic objects	22
Figure 2.2 Field-based models of geographical abstraction	26
Figure 2.3 Slivers created from polygon overlay	28
Figure 3.1 Data structures in GISs: (a) vector structure, (b) raster structure	62
Figure 3.2 A set of fuzzy surfaces with hypothetical data: (a) class P1, (b) class P2, and (c) class P3, distributed as in Figure 3.1	65
Figure 4.1 The principle of a fuzzy classification	74
Figure 4.2 The process of a fuzzy 3-means clustering	79
Figure 4.3 A mapping unit dominated by class denoted by rectangles, with other classes denoted by circles and triangles	82
Figure 4.4 The process of interpolating fuzzy surfaces	83
Figure 4.5 The features of an idealised semivariogram (unit of distance: miles)	86
Figure 4.6 The process of estimating the occurrence of classes at a point $x_0$ in the neighbourhood within the search radius $r$	88
Figure 4.7 The process of analysing fuzzy surfaces by slicing along the transect	95
Figure 5.1 The study area as marked by the rectangle (this reprint covers the southern parts of the City of Edinburgh)	102
Figure 5.2 The spatial coverages of aerial photographs and satellite images used in the case study (not to scale)	107
Figure 5.3 Sub-scenes of remotely sensed images used in this case study (scale 1:20,000 approximately): (a) SPOT HRV data, and (b) Landsat TM data. The same location is shown but with different boundaries	109
Figure 5.4 The scheme to build test data for the case study	113

Figure 5.5 The distribution of control points produced from land surveying in the local Edinburgh area (points are only approximately located)	115
Figure 5.6 Block configuration, Blackford Hill, Edinburgh	118
Figure 5.7 Land cover map derived from the 1:24,000 scale aerial photographs	123
Figure 5.8 Land cover classification derived from the SPOT HRV image	129
Figure 5.9 Land cover classification derived from the Landsat TM image	130
Figure 6.1 The improved algorithm on the supervised fuzzy <i>c</i> -means clustering, using the same data set as in Figure 4.2	146
Figure 6.2 Perspective views of fuzzy surfaces of land cover derived from the SPOT HRV data: (a) grass, (b) built, (c) wood, (d) shrub, and (e) water. The peaks on diagrams indicate higher certainty, i.e., increasing likelihood of presence	148
Figure 6.3 Perspective views of fuzzy surfaces of land cover derived from the Landsat TM data: (a) grass, (b) built, (c) wood, (d) shrub, and (e) water. The peaks on diagrams indicate higher certainty, i.e., increasing likelihood of presence	149
Figure 6.4 Histograms of entropy measures of fuzzy classification based on: (a) the SPOT HRV data; and (b) the Landsat TM data	150
Figure 6.5 Classified pure point samples from photogrammetric data: (a) grass, (b) built, (c) wood, (d) shrub, and (e) water	155
Figure 6.6 Semivariograms (Y-axis is semi-variance; X-axis is distance, in units of 2.5 metres): (a) grass, (b) built, (c) wood in north-south direction, (d) wood in east-west direction, (e) shrub, (f) water in north-south direction, and (g) water in east-west direction	157

Figure 6.7 Perspective views of fuzzy surfaces of land cover derived from the photogrammetric data: (a) grass, (b) built, (c) wood, (d) shrub, and (e) water. The peaks on diagrams indicate higher certainty, i.e., increasing likelihood of presence	160
Figure 6.8 Histogram of entropy measures for fuzzy surfaces generated via indicator kriging	161
Figure 6.9 Histograms for cross-entropy measures between fuzzy classified remote sensing data and fuzzy reference data derived by using sub-pixel component proportions: (a) the SPOT HRV data, and (b) the Landsat TM data	163
Figure 6.10 Histograms for cross-entropy measures between fuzzy classified remote sensing data and fuzzy reference data derived using indicator kriging: (a) the SPOT HRV data, and (b) the Landsat TM data	164
Figure 6.11 Categorical maps derived from the maximisation operation and draped on to their underlying certainty surfaces: (a) SPOT HRV data, (b) Landsat TM data, and (c) photogrammetric data	168
Figure 6.12 Categorical maps derived from the slicing operation based on the map shown in Figure 6.11 (c) at a succession of thresholds: (a) 40 %, (b) 55 %, (c) 70 %, and (d) 85 %	170

# Chapter 1

## Introduction

### 1.1 Background to the research topic

Geographical Information Systems (GISs) are powerful sets of tools for collecting, storing, retrieving, transforming, and displaying spatial data from the real world for a particular set of purposes, a definition provided by Burrough (1986). It is generally agreed that GISs are decision support systems involving the integration of spatially referenced data in a problem solving environment (Abler 1987; Cowen 1988).

As Rhind (1987) has said, in a general introduction to developments of GISs in the UK, a GIS is probably one of the most innovative and influential technologies in the information era, especially in the handling of geographical information. Due to increased investment and research, GISs have become more robust and efficient with reinforced algorithms and improved hardware capacity (NCGIA 1989). A GIS has been a useful tool for environmental studies and applications in a variety of disciplines and fields. Comprehensive examples have been described by Tomlinson (1987).

The developments and applications of GISs in a modern world have been very impressive. However, as Berry (1987) has argued, a GIS, though powerful in managing and processing massive amounts of spatial data, also provides a vehicle for generating seemingly accurate maps with minimum understanding of the actual spatial relationships. To some extent, automating the map data processing by using a GIS may lead to false perceptions about the quality of the results, as Bailey (1988), van der Wel, Hootsmans and Ormeling (1994) have all described. In other words, the powerful hardware capacity and high software precision emphasised by mainstream GIS systems tend to convey a false impression about the accuracy of the spatial data

held in GISs.

The various inaccuracies (errors) that arise in GISs do so for a number of reasons. Most obviously, geographical reality is inherently so complex that it needs to be abstracted, generalised and approximated in order to facilitate geographical studies, in particular when using digital methods, according to Burrough (1992) and Goodchild (1989a). The process of abstracting and generalising real geographical variations in order to express them in a discrete digital store is commonly known as spatial data modelling (Goodchild 1989a). The process of data modelling produces a conceptual model of the real world. It is unlikely that a highly complex geographical reality can be easily represented by a model with absolute accuracy. In other words, there exist differences between the model and the underlying geographical reality. These differences are known as conceptual errors (Bedard 1987; Goodchild 1989b; Veregin 1989).

Modelling represents a complex interaction of human and instrumental factors and acquiring the raw components of each model, the spatial data themselves, is also subject to error (Goodchild 1991). Depending on the skill of data analysts and the sophistication of instruments concerned, spatial data acquisitions will have varying levels of errors. Though advances in field technologies such as Global Positioning Systems, and improvements in laboratory techniques such as digital photogrammetry and digital image processing have greatly improved the quality of spatial data, errors will not simply be eliminated. The errors occurring in spatial data acquisition are called measurement errors (Burrough 1986). Errors also occur during geo-processing. A good example is the conversion of a vector data structure to a raster data structure or *vice versa*: the converted data rarely duplicate the original data, even if the original data are error-free. Such kind of error is called processing error (Goodchild 1989a).



Errors occurring in geographical abstraction and measurements will be made more complicated during the combined uses of spatial data, as will be developed below. The ability of a GIS to integrate diverse spatial data is frequently cited as its major defining attribute, and as its major source of power and flexibility in meeting user needs, as discussed by Maguire (1991). Spatial data integration can bring together all kinds of data necessary to derive certain desirable products or to assemble raw data required for spatial analysis or modelling in an automated fashion. However, the data integration may also be misleading. This is because spatial data have their inherent (and often limited) accuracies and different natural levels of detail. But, the usual data integration procedures in GISs take no account of the varying levels of accuracy and detail of the spatial data being merged, assuming scaleless and precise digital data during geo-processing (Abler 1987). As a result, when digital maps of different scales, for example, a detailed and accurate city map and a highly generalised stream network map, are merged by adopting a common scale, the results may be meaningless or potentially dangerous. For example, streams can be represented as running above buildings. Error propagation, or more specifically, error concatenation occurs during such geo-processing as map overlay (Veregin 1989).

Therefore, research into accuracy and error issues in GISs is well justified by the need for better knowledge about the nature of the spatial data and proper handling of the various errors occurring in geo-processing (Burrough and Frank 1996; Guptill and Morrison 1995). Moreover, in the 1990s, users will demand reliability and confidence as part of the acceptance of GISs as scientific and professional tools (Stuart 1996, pers. comm.).

With the ever-increasing amount of GIS applications, error handling is an increasingly urgent necessity for the reliable use of any GIS. This is more so in the case of urban applications where many complex spatial entities are often superimposed and compressed into small areas. Such a situation will in turn impose challenges to

research into the errors occurring in data on urban areas. This is because there is a need for large scale map data on the one hand (such as maps recording buildings with high precision and contouring data) and less detailed data at small scales on the other hand (such as highly generalised coverages of soil maps) in urban land use and land cover mapping. Such a need will inevitably lead to great complexity not only in terms of the resulting data products but also in error propagation when data layers of varying accuracies and details are used in combination. For example, Kirby and Zhang (1993) described their initial research on error modelling in an urban-orientated GIS, where data layers of different accuracies were incorporated.

As Tomlinson (1987) predicted, future developments in GISs will not only depend on better algorithms, data structures and continuing improvements in hardware, but will also need research leading towards a better understanding of the nature of spatial data themselves, their accuracy and errors. Therefore, this thesis aims for new strategies for handling errors in GISs, which are expected to improve the handling of spatial data themselves.

## **1.2 A review of the research on error issues**

The goals of research on error issues in GISs, according to Burrough (1994) and NCGIA (1989), are to investigate how errors arise, or are created and propagated in the GIS information process, and what the effects of these errors might be on the results of subsequent decision making. An ideal solution might be to have available the necessary information on errors and uncertainties intrinsic to spatial data and data layers during geo-processing, and to devise new algorithms that can track the propagation of errors (Goodchild 1991; Lanter and Veregin 1990). In this way, both producers and users of composite maps can have knowledge about the accuracy obtainable with a certain map operation by digital methods. Similar observations were made by Drummond and Ramlal (1992).

To facilitate a systematic review of research on error issues in GISs, Table 1.1 provides a framework in which errors are dealt with from their sources, through their propagation to their management and reduction. A similar framework is also provided in Veregin (1989).

Table 1.1 Error issues and strategies for their analysis

Error issues	Strategies for analysis
Identification	<p>distinction among error sources:</p> <ul style="list-style-type: none"> <li>• conceptual error (geographical abstraction)</li> <li>• measurement error (inaccuracy in measurement of positions, heights and attributes of spatial entities)</li> <li>• processing error (geo-processing)</li> </ul>
Detection and measurement	<p>accuracy assessment according to the types of data:</p> <ul style="list-style-type: none"> <li>• points - error ellipses</li> <li>• lines - the epsilon error band model</li> <li>• areas <ul style="list-style-type: none"> <li>quantitative attributes - standard deviations</li> <li>qualitative attributes - error matrices <ul style="list-style-type: none"> <li>- fuzzy methods</li> </ul> </li> </ul> </li> <li>• surfaces - root mean square errors (RMSE) <ul style="list-style-type: none"> <li>- kriging variances</li> </ul> </li> </ul>
Propagation	<ul style="list-style-type: none"> <li>• numerical attribute data with arithmetic operations <ul style="list-style-type: none"> <li>- standard stochastic theory</li> </ul> </li> <li>• categorical attribute data with map overlay <ul style="list-style-type: none"> <li>- Boolean logic model (probability theory)</li> <li>- fuzzy methods</li> </ul> </li> </ul>
Management and reduction	<ul style="list-style-type: none"> <li>• decision-making in the presence of errors</li> <li>• reducing the errors in GIS data products</li> </ul>

The first and basic issue is identification of the sources of errors. The following discussion is based on the definitions given in Bedard (1987), Chrisman (1991), Goodchild (1989a) and Veregin (1989). As has already been mentioned in the previous section, there is a distinction among three general sources of error: conceptual error, which is associated with the process of abstraction or generalisation about real world phenomena; measurement error, or error in position and height and attributes of entities; and processing error, which is involved in geo-processing. As an example of conceptual error, there may be a difficult decision to make about whether a small patch, which differs from the surrounding region, should be picked up as a separate class, or blended into the surrounding region. There are many occasions where measurement errors occur: the operator can't track precisely the lines to be digitised during a map digitising; the light dot is not accurately kept to the ground in sampling contours on a photogrammetric plotter; an agricultural field is wrongly classified as a recreational area by a classification of remote sensing images. For processing error, Goodchild and Lam (1980) showed that errors occur in spatial interpolation of area data. Aronoff (1989) provides a good summary of the common sources of errors encountered in GIS applications, which shows how errors are occurring from the data sources, through data storage and manipulation, to data output and use of GIS data results.

The next issue is to detect and measure the errors. Error detection and measurement are concerned with methods of assessing accuracy levels in spatial data. These methods can be differentiated by the data types under study, whether point, line, area or surface data. For points, the simplest method is based on the Gaussian distribution, which describes the variation in sample measurements. Like most least-squares models, an adjustment of survey data can estimate an error ellipse for each point along with its computed coordinates. In some applications, such as cadastral maps, the location of boundary corners may also follow these rules (Chrisman 1989; Cooper and Cross 1988).

For line data, Blakemore (1984) suggested that an epsilon error band model can be used to indicate the accuracy level about a digitised line, so that the true line will occur at some displacement (not more than  $\epsilon$ ) from the measured position. Epsilon  $\epsilon$  is defined as the displacement distance, and the epsilon error band is defined by twice the epsilon distance, that is,  $2\epsilon$  (see also Chrisman 1989). The epsilon band model can be used to provide overlay operation in GIS software such as ARC/INFO with specific tolerance for boundaries of individual combinations of classes (Aspinall and Pearson 1994). This adjusted overlay is more flexible and reliable than the globally specified tolerance at the outset, because it can be adapted to the underlying uncertainty levels of lines.

Area data refer to the spatial data depicting areal distributions, and are usually represented as irregular polygons or regular grid cells. For numerical (or quantitative) attributes of individual areas, certain statistical measures such as means and standard deviations can be used, as was described by Beard (1994). For categorical (or qualitative) attributes, on the other hand, a measure of classification accuracy is required. The assessment of classification accuracy as in remote sensing is based on a confusion or error matrix, which is derived after the comparison of the result against the "ground truth" based on a certain number and distribution of samples. One obvious measure of agreement is the sum down the diagonal as a proportion of the total sample. To take account of the magnitudes of the "marginal" probabilities and to remove the effect of chance from the measure, the Kappa statistic is appropriate, defined as the proportion of agreement over and above chance agreement (Rosenfield and Fitzpatrick-Lins 1986).

Fuzzy methods are useful for addressing the uncertainty in, often categorical, attributes (Altman 1994). Fisher and Pathirana (1989) reported on their derivation and mapping of fuzzy maps using satellite data in an urban change detection environment.

Lowell (1994) gave some general indications as regards the generation of fuzzy maps in a forest-orientated spatial database, where map data, aerial photography and remotely sensed data are incorporated. Goodchild, Sun and Yang (1992) described a model for uncertainties in categorical maps, which uses fuzzy maps and a spatial autocorrelation parameter together to characterise generic problems of heterogeneity within polygons and transition across polygon boundaries. The fuzzy maps can be derived from an image classifier, from the legend of a soil map or subjective assessment of uncertainties of interpretations.

In addition to point, line and area data as discussed briefly above, there is another type of geographical data: surfaces, of which Digital Elevation Models (DEMs) are examples. Published DEMs are sometimes supplied with a report of the accuracy level as a reference for users, as described in Guptill (1989). For example, the accuracy of the DEMs by USGS is measured by root-mean-square-error (RMSE). When a surface is expected to be essentially continuous, geostatistics can be usefully explored (Burrough 1986; Chrisman 1989; Webster and Oliver 1990). For instance, kriging, as a spatial prediction and interpolation technique based on geostatistical approaches, generates variance surface as well as the surface estimated from certain distribution of point observations or samples (Bregt 1991; Oliver and Webster 1990).

After the identification of errors and their measurements, attention is given to the third issue: error propagation. This means error propagation *per se* or error production. Error propagation refers to the process in which errors present in spatial data pass through a GIS operation and accumulate in output products, while error production refers to a situation in which errors in output products are attributable mainly to the operation itself (in other words, errors present in spatial data have little effect on the error in output products). Error production has been described previously as processing errors.

For modelling the effects of uncertainties as they propagate through GIS operations and to report these effects in connection with the results of GIS analysis, there is an accumulation of useful work. Heuvelink, Burrough and Stein (1989) used standard stochastic theory of error propagation to model the propagation of errors in raster GIS numerical data when applying continuous-differentiable arithmetic operations. The inputs of their model include the coefficients and their standard errors, maps of numerical variables and their prediction errors. While the prediction errors were obtained from kriging or calculated from repetitive sampling in a mapping unit such as a polygon, the errors associated with model coefficients were supplied by experts, which would otherwise be difficult to derive.

On the other hand, for categorical data as opposed to numerical data, Newcomer and Szajgin (1984) discussed propagation of uncertainties during simple Boolean overlay using probability theory. Their model accounts explicitly for the degree of error coincidence based on the computation of a conditional probability. They found that map overlaying using AND operations tends to degrade the accuracy of the derived map product when the number of overlaid layers increase. Walsh, Lightfoot and Butler (1987) applied this model for the AND operations to a set of raster land cover, slope-angle and soil type layers. The uncertainties in the source layers were measured by field-checking a sample of cells. The empirical results conform to the model of Newcomer and Szajgin (1984). Veregin (1989) extended the error propagation model to both logical AND and logical OR operations. In contrast to AND operations, OR operations tend to inflate the accuracy in the derived map products. Veregin (1995) presented his recent work on testing an alternate model that propagates the entire classification error matrix rather than a single index of error derived from this matrix such as overall classification accuracy, which has been used as the sole basis for propagating uncertainties in previous work. Simulation results tend to be consistent with actual levels of propagation uncertainties. Model output provides more information on the error characteristics of derived data than a single index of error.

Recognising the deficiencies of Boolean logic model such as unrealistic sharp class boundaries and spatial independence of uncertainties across the map layers concerned, Heuvelink and Burrough (1993) described a method for propagating uncertainties in cartographic modelling involving the intersection of several maps using fuzzy methods. They defined suitable fuzzy membership functions to transfer Boolean classes into fuzzy classes, and found that fuzzy methods are less sensitive to uncertainties in the data.

The final issue refers to strategies for error management (methods for decision-making in the presence of errors) and strategies for error reduction (reduction or elimination of errors in output products), which are of equal importance. These two interconnected problems go beyond error assessment and are concerned with the inferences that may be drawn from the results of error propagation. A detailed discussion is well beyond the scope of this thesis.

### **1.3 Problems from the past and the aims of this research**

Despite the increasing research on accuracy aspects of spatial data, as broadly discussed in the previous section, there are many issues remaining to be studied at depth, as the following section makes clear.

The error ellipse model for points has already been discussed, and will not be discussed further. As an extension of the error ellipse model for points, the epsilon error band model for lines has largely relied on traditional statistics for its proper interpretations. This model may be appropriate for modelling errors in reasonably homogeneous regions, where uncertainties involved in linear entities can be approximated by “buffer zones” defined by epsilon distances. However, the true nature of spatial error and spatial variability for geographical boundaries tends to be



far more complicated than homogeneously defined epsilon error band models would suggest, as admitted by Pullar and Beard (1989). Moreover, with the epsilon band model, there is an analytical complication arising from the need to model polygonal boundary errors while maintaining topological consistency (Aspinall and Pearson 1994; Goodchild 1989a).

For the study of uncertainties in categorical attribute data, a misclassification matrix implicitly assumes that classification accuracy or probability is constant over the mapped area, which is unlikely to be true, as coverages often have varying accuracies within their extent (Goodchild, Sun and Yang 1992). In other words, a misclassification matrix provides no information on the spatial variation in uncertainty below the class level. Additional information is also needed on the correlation between positional and classification errors (Estes 1992), since these two broad types of errors are difficult to isolate from each other in practice (Chrisman 1989).

In the error model for categorical data as described by Goodchild, Sun and Yang (1992), the use of a parameter of spatial autocorrelation, though explicit, has only an indirect control of the size of inclusions, but no control over their positions within a patch. As a result, more complex and specific models would be required when the spatial dependency is anisotropic, or where detailed information is available on the forms of spatial dependence present. Moreover, using this model, it is not possible to simulate the effects of correlations of uncertainties between maps, though it is possible to compute the uncertainties in outputs by overlaying on suits of simulated maps.

For error propagation in map overlay, even in the recent work as published by Veregin (1995), the propagation of an error matrix is still based on the assumption of independence of combinations of actual and estimated values between sources layers. However, this may not be appropriate in many situations where spatial dependence is

central to the process of map overlay. Besides, synthesis of individual maps representing specific factors of interests may not be able to honour the original geographical reality, from which map layers of individual factors are abstracted (Bailey 1988).

Though more suitable than Boolean logic, the fuzzy methods as described by Heuvelink and Burrough (1993) have several shortcomings. One shortcoming is that users will need to specify the parameters of fuzzy membership functions such as the class boundary values and the widths of the transition zones. Besides, though exceptionally implemented in their work, Heuvelink and Burrough (1993) admitted that spatial correlations of uncertainties across map layers are rarely quantified by the majority of researchers.

For continuous data such as DEMs, the variance surfaces, produced by kriging as by-products of spatial interpolation, have been taken to represent and model the spatial variability of errors occurring in the underlying data (Beard 1994). However, the variance surfaces need to be validated against an independent sample of field surveyed data, as warned by Burrough, van Rijn and Rikken (1993).

In overview, error issues, except for geometric accuracy, are, in general, under-researched (Burrough 1994; Rhind 1988). Moreover, current spatial databases designs, though being particularly flexible and useful for well-defined entities, are extremely difficult in modelling of natural variations evident in geographical processes (Burrough 1992). The limitations in data modelling have, however, greatly affected the extent possible to model and understand uncertainties of spatial data, as Goodchild (1992) said. Therefore, further research is needed.

A good framework or a systematic strategy for research on error issues in GISs may be based on the two alternative forms (or perspectives) of geographical abstraction,

which, as mentioned previously, create a conceptual model of the real world. These two perspectives of geographical abstraction are called field-based and object-based, depending on whether reality is regarded as consisting of a set of single-valued functions defined everywhere, or as occupied by a discrete collection of objects (Goodchild 1989a).

Different perspectives of geographical abstraction create different data models. Field-based models are favoured in physical geography and environmental applications, while object-based models are more suitable for cartography and facility management (Goodchild 1996; pers. comm.). The choice of object-based and field-based models depends on the specific nature of the underlying phenomenon, is limited by the data acquisition techniques and implementations of spatial databases, and will affect the possibilities for modelling the errors in spatial databases (Goodchild 1993). A detailed account of object-based and field-based data models will be provided in the next chapter.

Errors are better modelled in fields rather than in objects, because a field perspective facilitates the integration of roles of spatial data acquisition and spatial data analysis, thus allowing for the raw data with their spatial heterogeneities to be retained before compiling a specific map. These raw data are usually vital for deriving necessary information on errors in the geo-processing (Goodchild 1989b). A similar observation was implied in the work by Bailey (1988) and Drummond (1987). Moreover, errors in objects could then be seen as outcomes or realisations of a stochastic model defined on fields (Goodchild 1991).

The advantages of field-based models are particularly significant during the combined use of different data layers that possess different levels of accuracies. Moreover, the field-based models provide an open strategy for research on error issues in GISs, especially when one recognises the potential held by geostatistics (Goodchild 1996;

pers. comm.). Therefore, this thesis will pursue error issues in GISs from a field perspective, though object-based models and methods will be discussed for comparison.

Up until this point in the text, concepts of error and uncertainty seem to have been used interchangeably. Errors include inexactness and imprecision, and are mainly seen as characteristics of the underlying data. Uncertainties, on the other hand, imply fuzziness, heterogeneity and ambiguity, and tend to be concerned more with subjectiveness and the complexity of the geographical world (Altman 1994). It is thus argued that uncertainties seem to be more relevant, in particular for geographical studies concerning spatial variabilities such as land cover mapping. Therefore, the term uncertainty will be used as a superset of all sorts of errors and uncertainties mentioned above.

Pragmatically, a continuously varying field is perceived as a continuous surface. It appears clear from the account above on the merits of fields as opposed to objects in the handling of uncertainties that a surface-based method provides an integrating strategy. Under this integral strategy, access to the information on uncertainties in spatial databases will be more direct during data collection, data update, spatial analysis and output of final products (Stuart 1996, pers. comm.).

In particular, this thesis will show how a surface approach to data offers a common method for representation and operational handling of information on uncertainties, irrespective of the initial types of the underlying data. The utility of the proposed methods and models is illustrated by examples showing how the data commonly used in the context of suburban land cover mapping are stored, how the resulting final maps contain implicit information about their variable accuracies, how a sense of uncertainties can be visualised, and how a variety of analyses and interpretations can be performed to make fuller use of data on uncertainties.

To illustrate the theory, a case study is presented. The case study was chosen in the context of suburban land cover mapping for the following reasons. Suburban land cover mapping involves uses of diverse and sometimes disparate data sets that vary in dates, resolutions, scales, details, formats and accuracies. For example, it is common for spatial data at large scale and small scale and of different dates to be incorporated in land cover mapping and land use planning. As a result, inconsistencies arise from temporal, spatial and thematic discrepancies and mismatches, which will bring enormous pressure and complication to the digital handling of combined spatial data (Chrisman 1991). In the case of suburban land cover mapping, a strategy for handling uncertainty will need to be very effective and efficient to survive the challenges mentioned above. Therefore, the combined use of spatial data at a wide range scales with implied varying details and accuracies makes suburban land cover mapping an ideal environment to try out the proposed surface methods (Kirby 1996, pers. comm.).

There are useful methods recommended for modelling uncertainties via surfaces using probabilistic and fuzzy concepts (Burrough 1994; Goodchild 1993). These methods are not new, but nevertheless, they offer a sound basis for conducting research as presented in this thesis. More exactly, there are specific novel contributions in this thesis, which are now discussed.

Firstly, earlier work on using surface-based methods for representing and handling uncertainties is mostly orientated towards the integration of environmental modelling with GIS (Goodchild 1993). In the context of urban and suburban land cover mapping, where spatial data of various sources, different accuracies and diverse formats are usually involved, surface-based methods are seen to be lacking. This thesis incorporates one of the few novel uses of surface-based methods for urban and suburban land use and land cover mapping.

Secondly, there has been a conspicuous lack of appreciation of the two distinctive perspectives of geographical abstraction, i.e., objects versus fields, with respect to the research on uncertainties in spatial databases, and how such a distinction may help foster a more systematic strategy in handling uncertainties in spatial databases. So, this thesis will synthesise the past work pertinent to both the object-based and field-based models of uncertainties and express it in a rigorous way.

Furthermore, this thesis will contribute by clarifying probabilistic and fuzzy concepts by drawing examples from land cover mapping, by investigating the relationships between fuzzy surfaces and object-based models of errors such as the epsilon error band model and error matrix, and by discussing the research links between uncertainty concepts in GIS and geostatistics.

Finally, this thesis is characterised by its flexible exploitation of the established functionality provided by photogrammetric, remote sensing and GIS packages, though some programming was necessary to perform certain data transformations to allow for data integration. Such a strategy will help expose the uncertainties of spatial data without tremendously increasing the added expense of additional values.

#### **1.4 An overview of the contents of this thesis**

The first chapter provides an introduction to the research topic, followed by a short review of the research on error issues in the past with specific reference to error models pertinent to different types of spatial data and error propagation methods. Following the exposure of the shortcomings inherent in the existing methods for the representation and handling of uncertainties in spatial databases, the objectives pursued in this thesis are then identified and outlined.

Chapter 2 is a generalised introduction to spatial data and uncertainties in a GIS environment with special reference to urban and suburban land cover mapping. Two alternative perspectives of geographical abstraction, that is, discrete objects and continuous fields, are described, with their differences highlighted. It is also shown how various uncertainties occur as a compounded result of geographical abstraction and data acquisition. It is discussed that greater uncertainties occur when using a collection of discrete objects to model a continuously varying underlying phenomenon than using a set of continuous fields. Besides, a grid type of field models is the most suitable for representing and handling uncertainties in categorical maps, which are the common representation of spatial data for urban and suburban land cover mapping.

Chapter 3 moves to the description of uncertainties in spatial databases with respect to their measures and representation in objects and fields. While uncertainties in objects have been relatively well-researched with a suit of methods for their classification, measurements and estimation, uncertainties in fields are hardly studied in a systematic manner, in particular for urban and suburban land cover mapping, leading to a gap in research. Fuzzy surfaces are proposed for unified representations of uncertainties in categorical maps typically used in urban and suburban land cover mapping.

Chapter 4 stands as the main theoretical and methodological part of this thesis. It synthesises how the fuzzy surfaces are derived from remotely sensed images by using a suite of fuzzy classification methods, and from aerial photographs by using component land cover proportions in mapping units, and by using geostatistics, in particular indicator kriging. An account of the construction of fuzzy surfaces is followed by a section discussing the variety of methods for analysing the various fuzzy surfaces. Lastly, probabilistic and fuzzy measures used in estimating fuzzy surfaces are clarified in a discussion section.

Chapter 5 begins the introduction of the case study undertaken. This case study is presented with respect to the study area chosen, the data sources available and accessed, and the test data compiled. The hierarchy of test data is co-registered with a corresponding hierarchy of accuracies: ground truth data, photogrammetric data digitised from 1:24,000 scale aerial photographs, classified SPOT HRV data and Landsat TM data. While it will be shown therein that spatial data acquisition is a very costly process in terms of time and manpower, especially in a situation where a diversity of spatial data is involved, a dual purpose for this chapter is to show how uncertainties in objects are measured and represented. It is also seen that, though object-based models are suitable for uncertainties relevant to well-defined entities, they are not suitable for fuzzy phenomena, which are characterised by spatial heterogeneities. This is where a field-based model works well, as developed in Chapter 6.

Chapter 6 sets out to generate and analyse the fuzzy surfaces by using methods described in Chapter 4. While fuzzy surfaces are derived from remotely sensed images by using an improved version of the fuzzy *c*-means clustering algorithm, fuzzy surfaces are produced from photogrammetric data by firstly using sub-pixel proportions of component land covers, and then by using indicator kriging. The nature of the remote sensing data means that fuzzy surfaces produced from these data will be more varied than fuzzy surfaces produced from photogrammetric data. The analysis of fuzzy surfaces generated from photogrammetric data confirms that they are better derived by using indicator kriging than by using sub-pixel component proportions. Further, categorical maps are derived from fuzzy surfaces by using “maximising” and “slicing” operations, which are seen to contain richer information on the spatially varying nature of uncertainties than traditional categorical maps. Results confirm that uncertainties are better researched from fields than from objects. Finally, an evaluation of the accuracies of classification of remote sensing images is performed comparatively, by using “fuzzy” and “hard” methods, and suggests the



extraordinary usefulness of fuzzy methods in situations where mixed pixels are significant and where fuzziness is intrinsic to both the remotely sensed data and the ground data.

The concluding chapter, Chapter 7, examines in retrospect the particular implementation and application of the surface approach in this study. In prospect, this chapter strongly recommends that further work should be pursued in order to develop the fuller potential of surface-based methods. This becomes more convincing when it is recognised that natural links between fuzzy surfaces and geostatistics have been reinforced in this thesis.

## Chapter 2

### Uncertainties in spatial data modelling and acquisition

#### 2.1 Introduction

The GIS communities have seen a proliferation of general and specific purpose GISs and spatial data handling packages made available to them over the past decade. However, the fundamental research on spatial data themselves tends to be overlooked by those seeking to harvest the benefits from GIS applications. Increasingly refined GIS hardware and software will not guarantee an unbiased and complete understanding of the nature of spatial data and their uncertainties. Thus, this chapter is dedicated to the description of spatial data and uncertainties.

As mentioned in the previous chapter, the case study has been contextualised in suburban land cover mapping, where, because the spatial data are so diverse, it becomes more necessary to elaborate on spatial data and uncertainty issues. Attention is therefore given to various uncertainties that result from geographical abstraction (also known as spatial data modelling) and spatial data acquisition, with particular reference to land cover mapping. It becomes clear therein that various uncertainties occur as a compounded effect of spatial data modelling and acquisition. In the final section, it is seen that it is sensible to begin the study of uncertainties from the fundamental issue of spatial data modelling, because such a strategy provides valuable insights into how a choice of object-based and field-based spatial data models may affect the possibilities with which uncertainties are handled and usefully explored.

## 2.2 Spatial data modelling and uncertainties

As mentioned in Chapter 1, geographical reality is inherently complex, and needs to be abstracted, generalised and approximated in order to facilitate geographical studies. This obvious fact is stressed in the context of modern digital methods by many authors, for example by Burrough (1992) and Goodchild (1992). Usually, the process of abstracting and generalising real geographical variations in order to express them in a discrete digital store is known as spatial data modelling (Goodchild 1989a). The process of spatial data modelling produces conceptual models of the real world. As has already been described in Chapter 1, there are two kinds of spatial data modelling: object-based and field-based. An object-based model views the real world as occupied by a collection of discrete objects, while a field-based model regards the reality as consisting of a set of single-valued functions defined everywhere (Goodchild 1989a).

For an object-based model, it is necessary to introduce a few concepts about entities, objects, features and attributes. An entity is a real world phenomenon that is not subdivided into phenomena of the same kind. An object is a digital representation of all or part of an entity. A feature is a defined entity and its object representation. Further, an attribute is defined as a characteristic of an entity, and an attribute value is a specific quality or quantity assigned to an object. These definitions were supplied by NCDCDS (1988). Based on these definitions, it becomes clear that an object-based data model expresses the real world as consisting of objects associated with their attributes. NCDCDS (1988) published an illustrated classification of cartographic objects, part of which is presented as Figure 2.1.

Firstly, as shown in Figure 2.1, for points, there are two types of generic point objects: points (including entity points, label points and area points) and nodes. An entity point is used principally for identifying the location of a point entity such as a

tower (the location is usually specified by a set of coordinates). A label point is used principally for displaying map and chart text to assist in feature identification. An area point is a point within an area carrying attribute information about that area. A node is a topological junction or end point that may specify geometric location (again, specified by a set of coordinates).


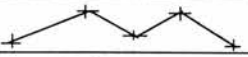



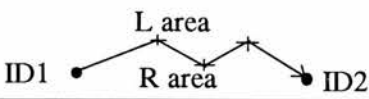
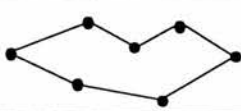
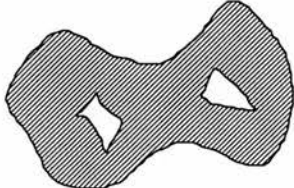
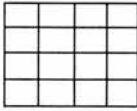
Point (0-dimensional)	Point	Entity point Label point Area point	+
	Node		•
Line (1-dimensional)	Line segment		
	String		
	Arc		
	Link		
	Directed link		
	Chain	Simple chain Area chain Network chain	
	Ring (created from)	Strings Arcs Links Chains	
Area (2-dimensional)	Interior area		
	Polygon (simple or complex, i.e., with inner rings)		
	Pixel Grid cell		

Figure 2.1 NCDCCDS classification of cartographic objects (source: NCDCCDS 1988)

Secondly, for lines, there are line segments, strings, arcs, links, chains and rings. A

line segment is a direct line between two points. A string is a sequence of line segments that does not have nodes, node IDs, Left/Right IDs, any intersect itself or other strings. An arc is a locus of points that forms a curve, defined by a mathematical function. A link is a connection between two nodes, while a directed link is a link with one direction specified. A chain is a directed sequence of non-intersecting line segments and/or arcs with nodes at each end. It may be a complete chain having node IDs and Left/Right area IDs, an area chain having Left/Right area IDs but no nodes IDs, a network chain having nodes IDs but no Left/Right area IDs. A ring is a sequence of non-intersecting chains, strings, links or arcs with closure. It represents a closed boundary, but not the interior. As an example, a ring created from links is shown in Figure 2.1.

Lastly, an area object is a bounded continuous two-dimensional object which may or may not include its boundary. There are irregularly-shaped and regularly-shaped (usually rectangular) area objects. Irregularly-shaped area objects include interior areas and polygons. An interior area does not include its boundary. A polygon consists of an interior area, one outer ring and zero or more non-intersecting, non-nested inner rings: a simple polygon has no inner rings, while a complex polygon has one or more inner rings. There are two types of rectangular area objects: pixels and grid cells. A pixel is a 2-dimensional picture element that is the smallest non-divisible element of an image. A grid cell is a 2-dimensional object that represents an element of a regular or nearly regular tessellation of a surface.

So, under an object-based model, spatial data exist in the forms of discrete points, lines and areas. Points occupy very small areas in relation to the scale of the data base (e.g., towns and cities on small scale maps). Lines are used for real linear phenomena, and are exemplified by entities such as highways, rivers, pipelines, and power lines. Area objects are used to represent distributions such as building blocks, lakes and land parcels and other patterns that occupy areas at the scale of the GIS. Area data are

represented usually as categorical maps, which take the form of a finite number of discrete nominal classes represented either by polygons with homogeneous values or collections of adjacent raster cells with the same values (Campbell 1987).

For individual objects, their various characteristics will be analysed by measuring and analysing corresponding attributes. Attributes can be measured on different types of scale. The scale of measurement is a system of classification of measurements depending on the mathematical operations permitted: the simplest scale is the nominal scale which only permits a test of equivalence; ordinal scale adds the property of order and the operations greater than or less than; interval scale defines addition and subtraction; and ratio scale substitutes the earlier ones with the inclusion of multiplication and division (Stevens 1946). Chorley (1966) provided specific examples in geomorphological research.

Examples of nominal data are the name of a city, the material of a road surface and the ownership of a land parcel. Ordinal data are used where a set of categories has a natural ranking associated with it (e.g., grades of agricultural land), or may be ranked from first to last according to some criterion (e.g., the ranking of residential properties for taxing). Interval data are referred to a purely arbitrary zero value, for which good examples are elevation above mean sea level and temperature measurement on a centigrade or Fahrenheit scale. Examples for data on ratio scales are given by city population, road speed limits and parcel areal extent (Flowerdew 1991). Sometimes, nominal and ordinal data are referred to as categorical or qualitative data, while interval and ratio data are called numerical or quantitative data.

As seen above, with an object-based model, the real world is represented as discrete points, lines and areas with associated attributes. An object-based model is suitable for well-defined entities such as roads, buildings and land parcels, and is widely used in cartography and facility management. However, it is important to note that

representing well-defined entities by using object-based models is considered suitable only on relative conditions and within tolerance, because pure points, lines and areas do not exist in the real world. Besides, entities have varying and irregular sizes and shapes. Thus, it is rare to find an object-based model that duplicates the real world accurately. The difference between a modelled world and the real world represents the uncertainty due to geographical abstraction. Such an uncertainty is complicated by the limitations in the process of spatial data acquisition: inaccuracies in the positions and attributes of objects. This is discussed in the next section.

Attention is now given to the field-based models for spatial data. According to Goodchild (1989b and 1993), there are six different types of field-based models: irregular points, regular points, contours, polygons, grid and TIN (triangulated irregular network), as illustrated in Figure 2.2. Two types of variables are involved in field models: categorical and numerical. Categorical variables are measured on discrete scales (nominal and ordinal), while numerical variables are measured on continuous scales with interval and ratio properties. Categorical variables are exemplified by land cover and soil type, while examples of numerical variables include elevation, noise level, annual rainfall and atmospheric pressure.

There are four field models for categorical variables of which two are used for areas: the grid model, where variation is described by determining its value within each rectangular cell; and the polygon model, where the plane is divided into irregular polygons. For both models, spatial variation within cells and polygons are ignored. Besides, a polygon often removes some of the geometric complexity of boundary lines. This is evident in a classification of remote sensing images, where each pixel is assigned a single dominant class, and where contiguous patches of pixels with identical classes are further smoothed to generate polygon-like data. Therefore, there are uncertainties occurring in using grid and polygon model, where spatial variations within cells and polygons are generalised, allowing only for a single and dominant

class for each cell or polygon.

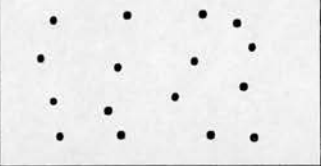
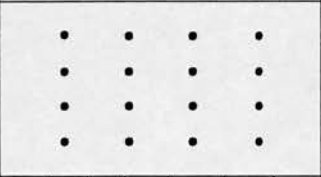
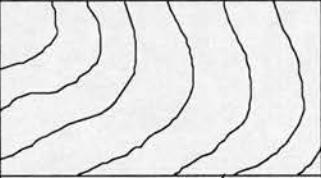


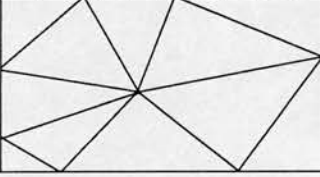
Field models	Graphical representations	Valid types of variables	Examples
Irregular points		numerical categorical	weather station data
Regular points		numerical categorical	elevation data
Contours		numerical	noise level data
Polygons		numerical categorical	land cover data
Grid		numerical categorical	remote sensing data
TIN		numerical	topographic data

Figure 2.2 Field-based models of geographical abstraction

For numerical variables, all the six models are commonly used. Using a TIN model, only critical points defining local discontinuities are required to model the spatial variation under study. Thus, TINs are often “reserved” for elevation data, where linear features and breaks of slope along triangle edges fit well with any naturally occurring topography (see Theobald 1989 for discussion). As shown in Figure 2.2, with polygon and grid models it is possible to query the value of a variable at any



point in the plane, whereas, with others, values can be obtained at arbitrarily located points only through the use of an additional process called interpolation. Moreover, because interpolation is not standard, it is more difficult to address the problem of uncertainties for those field-based models which require an interpolation process, as the uncertainties in an estimate of a variable such as elevation from a set of regularly spaced points depends on the interpolation method used. This is more so for interpolations based on irregularly spaced points and contours, in which the spatial pattern of uncertainties is highly non-uniform, with lowest levels of uncertainties close to the points and contour lines. Thus, only grid and polygon models are fully field-based models in this sense, and a further distinction between these is made in the discussion section.

As with object-based models, uncertainties occur also with field-based models. This is because it is rarely the case that complex spatial variations can be easily and accurately modelled by mathematical functions. Thus, the spatial variations need to be generalised and simplified in order that they can be approximated by a finite number of observations. Besides, observations themselves will suffer further uncertainties in spatial data acquisition, as will be established in the next section.

It has been seen so far in this section that the concepts of objects and fields, as embodied during the course of object-based and field-based geographical abstraction, provide necessary mechanisms, by which geographical reality is abstracted, simplified, and measured to permit computer storage, modelling and analysis. However, as mentioned above, geographical abstraction leads to a limitation in the accuracy of spatial data, acknowledgement of which has been largely omitted in the past. Besides, past research has somehow unduly emphasised precision rather than accuracy. Such a misplaced emphasis reinforced a spurious precision, which contributes to the creation of “slivers”, that is, small polygons occurring in polygon overlay, a topic reviewed by Goodchild (1991). Figure 2.3 provides an example of slivers created from overlaying

polygons, which are digitised slightly differently, as represented by the long, narrow shaded areas, where the boundary separating polygons P1 and P2 is digitised as a1-6, on one map layer, and as a1, b2-5, a6 on the other.

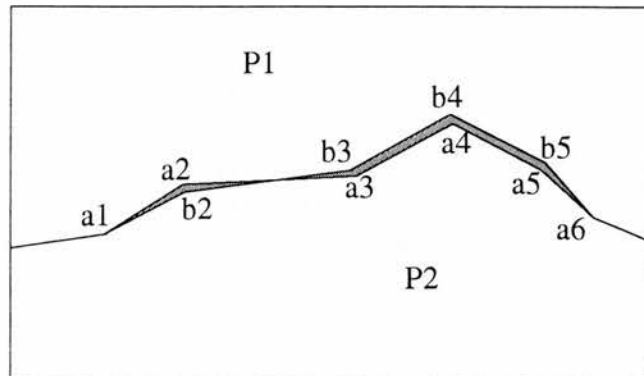


Figure 2.3 Slivers created from polygon overlay

If polygons boundaries were represented in accordance to their underlying levels of accuracies, that is, if there were a tolerance adopted for the differences existing between the two digitised “versions” of the polygon boundary, slivers would not occur. This can be easily understood from Figure 2.3, where the slivers would collapse, resulting in the boundary lines coinciding, if a threshold is set big enough to resolve the discrepancy existing between the two versions of the boundary.

It is therefore important that further research on uncertainties in spatial databases be based on the understanding of the process of geographical abstraction. This is because it is generally agreed that uncertainties resulting from the process of geographical abstraction tend to be the most influential ones. Similar observations have been made by Goodchild (1989a) and Veregin (1989).

## **2.3 Spatial data acquisition and uncertainties**

Spatial data acquisition is performed in accordance with abstraction of geographical reality. As will be shown below, the spatial data acquisition will have further impacts on the accuracies in the measured data sets. Since this research has been set in the context of urban and suburban land use and land cover mapping, an introduction to this topic is necessary. Land cover refers to the material cover on the surface of the Earth, while land use refers to the cultural use of land. A good example is an area which is forested in terms of land cover, but may be used as habitat for wild life or as recreational land in terms of land use (Jensen 1983).

Land cover mapping is a fundamental activity, necessary for a variety of applications: land use planning, agricultural development, forestry management and environmental monitoring (Townshend 1992). For urban applications, land use and land cover mapping provides basic data for purposes ranging from inventory to prediction (Jensen 1983; Jensen et al. 1994). As mentioned in the introductory chapter, land cover mapping usually incorporates various data of different origins, resolutions, accuracies and formats. These diverse spatial data are typically derived from a variety of techniques ranging from land surveying to satellite remote sensing, as is described below. Besides, as this thesis seeks to understand the uncertainties existing in spatial data more fully, this section also provides initial clues to the various uncertainties involved in the data acquisition process, following the discussion presented in the previous section of uncertainties occurring in geographical abstraction.

### **2.3.1 Land surveying**

Instrumental land surveying plays an important role in the spatial data collection process. It is used to derive adequate and well-distributed ground control points (GCPs), which are required to geo-reference the separate data layers or contiguous

data sets involved in a particular GIS application properly. This function is particularly valuable in an urban and suburban land cover mapping, where many layers of data need to be co-registered for spatial analysis and planning. Moreover, land surveying is used to provide ground truth data for interpretation and measurement based on aerial photographs and remotely sensed images, as will be described in the following sub-sections.

Traditional land surveying requires skill and experience in the observation of angles, distances and height differences. Introduction of the electronic theodolite or the total station (i.e. an electronic theodolite integrated with or attached to an EDM) has made conventional surveying easier and faster. Some established procedures are made available as standard programmes and are even built into GIS packages as sub-systems (e.g., COGO commands in ARC/INFO), computerising the entry of angle and distance measurements and subsequent coordinate calculation. Global positioning systems (GPSs) have been developed since the 1980s and increasingly are replacing traditional land surveying systems as efficient ground data collection tools, with improved precision and lowered cost. The link between GPS and GIS is rapidly developing (Kennedy 1996).

It is known that the sophistication of instruments and the refinement of surveying procedures lead to efficient treatment of the errors inherent in surveying process and measurements. Most importantly, land surveyors check the accuracy of their work directly in the field using highly refined surveying procedures or rules. Also, by survey adjustment techniques (Cooper and Cross 1988), field measurements can be modified very precisely into most probable values. Thus, it is appropriate to say that uncertainties are relatively easy to measure and control in land surveying. The understanding and the handling of errors will become more important as the automation of surveying work will generate increasing amounts of data, whose errors may accumulate and propagate very quickly, especially when land surveying is carried

out over large areas with high density of detail.

### **2.3.2 Photogrammetric techniques**

While land surveying is required to provide accurate GCPs for many GIS applications, it would be impractical and prohibitive to employ land surveying to perform point by point detail mapping in urban areas, where efficiency and currency are more crucial for fast and updated information on urban land use development. This is where photogrammetric techniques can help.

According to the International Society for Photogrammetry and Remote Sensing (ISPRS), photogrammetry is defined as the art, science and technique of obtaining reliable information about physical objects and the environment through recording, measuring and interpreting photographic images and patterns of electromagnetic radiant energy and other phenomena.

As a discipline typically using aerial photography for the production of topographical maps, DEMs and other forms of geometric land information (Torlegard 1986), photogrammetry now has a major role in spatial data acquisition (topographical and thematic) for GISs over a wide range of scales: urban land cover mapping and urban land use planning (Forster 1983), resources management and environment monitoring (Kirby 1992; Welch 1987).

There is a variety of aerial photographic techniques, which provide different sources to derive geometric and thematic information for urban applications. Panchromatic film for aerial photography is being progressively superseded by natural colour film which is widely used in urban applications because of ease of interpretation: the colour balance parallels the interpreter's real world experience (Jensen 1983). Colour infrared film is particularly valuable because of the increased contrast between

vegetation, water and man-made structures, but is less suitable for general purpose interpretation.

The photogrammetric stereo plotter is the instrument most commonly used to transfer aerial photography information to planimetric and topographical maps. In a stereo plotter, the operator views two mutually overlapping aerial photographs taken from different positions to form a three-dimensional model. By moving a floating point around in the stereo model, the operator can draw in roads, rivers, contours and other features. If the plotter has a digital transmitter, all movements used in drawing in the photogrammetric model can be numerically encoded and entered in computer storage.

The main advantage of collecting data by digitising from photogrammetric stereo models rather than from existing maps is that greater positional accuracy of the features should be possible. In addition, digital data collection using photogrammetry is flexible in terms of scale and choice of features to be digitised. It can also be used to update an existing data base. Elevation data acquisition by points, profiles and contour lines can be performed according to distance, time or coordinate increment.

In addition to measurements of size, perimeter, area, volume, and elevation of objects based on a reconstituted stereo model, a substantial amount of thematic information can also be obtained from aerial photographs by applying photo interpretation techniques. Aerial photo interpretation is a process by which information on site, association, tone, geometry, texture, pattern, size and shadow is utilised by an interpreter to obtain reliable identification of entities on the ground (Jensen 1983). Further details of the nature of aerial photo interpretation will be assumed.

In order to obtain a general but comparative understanding of photogrammetric techniques with respect to the historical development of photogrammetric plotters in the modern world, a summary of the features of different photogrammetric plotters is

shown in Table 2.1.

Table 2.1 A classification and comparison of stereoplotters

Characteristics	Stereoplotters		
	analogue	analytical	digital
Instruments	optical-mechanical	optical-mechanical	computer
Date introduced	c. 1920s	c. 1970	c. 1990
Input	photographs	photographs	scanned images
Stereomodel	true	mathematical	mathematical
Orientations	manual	fast	automated
Restrictions on photo/plot scales	yes	no	no
DEMs	manual	enhanced	automated
Orthophoto	optical-mechanical	enhanced	digital
Interpretation	manual	manual	increasingly automated
Output	generally graphical	digital, graphical	digital, graphical
Database function	generally no, some with interface to computers	yes	yes
Examples	Kern PG2, Galileo G6	Kern DSR 14, Carto. Eng. AP190	Kern DSP1, Leica DVP

As shown in Table 2.1, the traditional instruments are analogue plotters which use space rods and an optical-mechanical system to facilitate the reconstitution of the stereo models. In analytical plotters, software replaces the rods and all instrument movements are digitally controlled. Transformation parameters are used to transform continuously the image coordinates to ground coordinates. Known lens distortions can be corrected. There are no restrictions regarding the photoscales, principal distances and plot scales. There are additional data edit and management functions. Output can be digital data bases or maps in graphical forms.

For analytical plotters, however, human control is still needed to identify and track entities to be digitised and keep the light dot exactly on the ground surface. On the other hand, digital photogrammetric stations are based on a stereo-reconstitution on a high resolution computer monitor driven by a PC or a workstation. Digital plotters can fully automate the generation of DEMs and orthophotos, and may have automatic feature extraction depending on the machine's capability and function. They can process mono, stereo, and multi-images; terrestrial, aerial and satellite images; different kinds and combinations of imaging sensors and nonimaging sensor data; and digitised photography and digital scenes.

However, despite the wide-spread application of these different photogrammetric techniques for the provision of geometric and thematic information, various uncertainties occur in all steps of photogrammetric data acquisition and processing. These uncertainties occur from aerial photography to the final data output: camera lens distortion, atmospheric refraction, displacement due to tilt and relief, instruments limitation, and operator's bias. The accuracy of digitising from a correctly orientated stereo model will depend not only on the stereoscopic acuity of the operator at work but also his/her interpretation. Often, the interpretation bias has greater effects on the digitising results than the accuracy of measurement (Kirby 1988).

In urban areas, the uncertainties involved in aerial photo interpretation are more substantial than in rural areas for the following reasons. Urban land cover, consisting of residential and industrial buildings, transportation networks, parks and a variety of mixed cover types, represents a large variety of fine detail (Kirby and Zhang 1993). The highly varied land use and land cover in urban areas is complicated by the often limited contrast between different land use and land cover types. As a result, it is not uncommon for land cover mapping in urban areas to be of limited accuracies, even when high quality aerial photographs are used (Jensen 1983; Wolf 1988).



It is clear from the account above that geometric and thematic information obtained by measurement and interpretation based on aerial photographs suffers various uncertainties. Although the development of digital photogrammetry with automated feature extraction may lessen the problems of uncertainties in delineation and interpretation on aerial photographs, uncertainties will not be eliminated. Because automated feature extraction is far from satisfactory, various uncertainties will continue to exist in the data products. Nevertheless, it is important to recognise that the advancement of digital plotters will speed the production of topographic and thematic data and at lower cost. Extra amounts of raw data will be of benefit for handling uncertainties.

### **2.3.3 Remote sensing**

As described above, recent developments in photogrammetry have significantly improved the efficiency and currency in spatial data acquisition, with increased accuracy. This is especially apparent in digital photogrammetry, implying great opportunity for automated feature identification and image understanding as well as geometric measurement. In terms of digital image processing, there is another closely related branch of techniques: remote sensing, as established below.

According to the definition by Lillesand and Kiefer (1994), remote sensing is the art and science of obtaining information about an object, area, or phenomenon through the analysis of data acquired by a device that is not in contact with the object, area or phenomenon. Remote sensing technology is more complex than either land surveying or photogrammetry, and therefore there are more possibilities for errors, as will be made clear subsequently.

Remote sensors may operate on airborne or spaceborne platforms. An example of the former is the airborne thematic mapper (ATM) sensor on board aeroplanes, providing

data with fine resolutions (pixel sizes ranging from sub-metre to several metres). Foody and Cox (1994) used ATM data for land cover mapping in an urbanised area, a university campus. On the other hand, spaceborne remote sensors on a number of satellites in orbit, such as Landsat MSS (multispectral scanner), Landsat TM, SPOT HRV (high resolution visible), JERS OPS (optical sensor), are providing update and regular coverages useful for deriving natural resources and environmental data from regional to global scales, as summarised in Table 2.2 from Neto (1996).

Table 2.2 Optical and radar sensors on spaceborne platforms

Platform	Sensor	bands	ground resolution (m)	scale recom.	applications
Landsat	MSS	4	79	1:200,000	land use, mapping, geology
Landsat	TM	7	30 120 (ch. 6)	1:100,000	land use, mapping, geology land use, DEMs,
SPOT	HRV	4	10 (P) 20 (XS)	1:50,000	cartography, regional planning
IRS-1A	LISS I	4	73	1:200,000	land use,
	LISS II	4	36.5	1:100,000	cartography
JERS-1	OPS	1	18.3 × 24.2	1:50,000	geology, land use
JERS-1	SAR	L-band	18	1:100,000 ~1:600 000	geology
ERS-1	AMI	C-band	30	1:100,000 ~ 1:600,000	ecology geology
RADARSAT	SAR	C-band	25 × 28 11 × 9	1:100,000 ~ 1:600,000	agriculture, forestry, topography

(source: Neto 1996)

As shown in Table 2.2, two kinds of remote sensors are relevant for mapping: optical

and microwave. The optical sensors such as Landsat and SPOT sensors detect reflected sunlight and re-emitted thermal radiation; while microwave sensors deal with transmission and reflection of energy in the microwave portion of the radio frequency spectrum, such as the active microwave instrumentation (AMI) on board ERS-1, and the synthetic aperture radar (SAR) on board JERS-1 and RADARSAT.

In some situations, remote sensing data are used to provide direct estimates of certain physical variables independently of other data. Operational examples are found in the studies related to vegetation or forestry, whose spatial extent is frequently comparable to the scale of the satellite data, as promoted by Harris (1987). Although the continuum of urban land cover classes cannot be divided readily into discrete classes with a conventional image classification approach, satellite sensors with a relatively fine spatial resolution, such as those carried on the SPOT or Landsat satellites, have considerable potential in urban studies (Campbell 1987). In fact, Weber (1994) argued that the repetitiveness of remote sensing data makes them attractive for many applications, including updating information on urban areas, for monitoring urban growth.

Some variables may be inferred from the measurement of a related variable based on remote sensing data. For example, by relating population density to building characteristics such as building type and coverage, remote sensing data may also be used to derive estimates of demographic variables (refer to Weber 1994 for technical detail). Another example may be found in Langford and Unwin (1994), where population density surfaces were derived from estimates of built up areas based on classified Landsat TM data.

However, uncertainties exist in remotely sensed data, both in their acquisition and in their subsequent handling. Firstly, there are radiometric and geometric distortions present, which occur for various reasons: the measurement magnitude, or radiometric

value, of each pixel may be degraded by noise caused by sensor inaccuracy, variation of reflected energy due to change of terrain and solar incidence; the geometric distortions are created by rotation of the Earth, instability of satellite in orbit or aircraft in flight, atmospheric aberrations and variations in surface elevation.

Secondly, uncertainties occur in the classification of remotely sensed images, of which there are two main approaches, supervised and unsupervised. A supervised method uses ground truth data, that is, observations of conditions on the ground. It begins with a compilation of training data for each class to be mapped. A skilled analyst works on the screen, selecting areas known to contain the classes named. Spectral means and standard deviations are calculated for each class. Each pixel is classified by such simple methods as the parallelepiped or minimum distance rule, or by following more probabilistically sophisticated methods such as maximum likelihood classifier (MLC) or Bayes' decision rule (when *a priori* knowledge of the likely relative frequency of classes is available). In particular, the MLC method assumes that the training sample of each class has a multi-variate normal probability density distribution on spectral space so that the pattern of each can be defined by the position of the centroid and the variance and covariance matrix of its spectral distribution (Campbell 1987; Townshend 1981; Wood and Foody 1993).

However, there are limitations to the supervised classification method. Because the supervised method, in particular the MLC approach, is based on the assumption that the classes have multi-variate normal probability functions, which is not usually held in the real world, especially in suburban land cover mapping, as a result, the classified data are often far from perfect. In addition, the classification rule tends to discriminate against classes with a low probability of occurrence.

Unsupervised classification, as an alternative to supervised classification, exploits the inherent structure in the data. It starts with a cluster training procedure based on

spectral characteristics. The clustering is a way of ordering data by sorting pixels into classes according to a distance measure. The analytical procedure for clustering may be interactive and includes iterative steps of initialising cluster number, assigning pixels to the nearest cluster according to a method of distance measuring, calculating updated clusters means. The process stops when the clusters no longer change between successive iterations. The character of the clusters is determined *a posteriori* by looking at ground properties of samples from each cluster (Swain and Davis 1978).

The unsupervised method is specially useful when working in a new area, because the clustered classes should reveal what land cover or terrain types can most successfully be distinguished by using the image data. It is strongly recommended within regions of natural, non-agricultural terrain, because of the difficulty of obtaining unique spectral signatures or homogeneous training samples. The difficulty may be due to spectral differences in the cover types themselves or variations in the terrain in the use of supervised methods. However, the resultant classes from an unsupervised method are not guaranteed to be useful. Some of the clusters may be meaningless because they may contain too wide a variety of ground conditions. Moreover, the interpreter will make less use of ground and ancillary data that are usually available for interpretation, implying limited accuracies in classified data.

In overview, the accuracy of classification of remote sensing images, either by supervised or unsupervised methods, is affected by many factors. First of all, the classification has relied on the assumption that the area of study is composed of a number of unique internally homogeneous classes and that classification analysis based on reflectance data and ancillary data can be used to identify these unique classes by means of ground truth areas, as discussed in Townshend (1981). In other words, both supervised and unsupervised classification assume that image data form separate groups in feature space, and the groups can be associated with ground observations. These groups are described by some boundaries such as the rectangular parallelepiped

shape and the hyper-ellipsoid in MLC. According to Skidmore and Turner (1988), it is frequently observed that clusters in feature space hardly find themselves exhibiting distinctive patterns which may be approximated by a rectangular or hyper-ellipsoid shape. Therefore, classification uncertainties are bound to exist in remote sensing data.

In another view, uncertainties occur due to the presence of mixels (mixed pixels, that are not completely occupied by a single, homogeneous category). For example, the heterogeneous nature of urban areas (i.e., the varying mixture of urban surface) produces a mixed pixel response, as shown in Forster (1983). As argued by Fisher and Pathirana (1989), the assumption that pure pixels comprise a homogeneous area representative of one land cover class does not usually hold. Campbell (1987) showed that mixed pixels are common especially at the edge of large parcels, or along long linear features, or among scattered occurrences of small parcels. Besides, mixed pixels also occur where the land cover elements are continuous and gradual. Wood and Foody (1993) provided empirical results about the gradual changing land cover in a natural environment.

Though the mixed pixel contains more than one class, it may only be allocated to one class. Furthermore, as the mixed pixel displays a composite spectral response which may be dissimilar to each of its component classes, the pixels tend not to be allocated to one of the component classes. Error is therefore present in the classified image containing mixels. The estimation of the areal extent of the land cover classes may thus also be prone to error. In short, mixed pixels degrade the accuracy of image classification.

Despite growing research into new methods for classification of remotely sensed images, such as contextual, neural network, rule-based approaches, which may overcome some of the shortcomings in supervised and unsupervised approaches and

adapt to the uncertainties occurring in remote sensing process, no method can claim to be perfect (Davis and Simonett 1991). And, given the fact that remotely sensed images have undergone just approximate corrections for radiometric and geometric distortions, any appropriate use of remote sensing data must address the error issues.

In summary, the multi-spectral data of air or space borne sensors are used for natural resources mapping and environment monitoring on a variety of image processing systems. However, remotely sensed data continue to suffer radiometric and geometric distortions. Besides, the relatively large pixel size of the sensors implies that there is a high probability of more than one cover type contributing to reflectance values as recorded. The varied nature of urban areas produces mixed pixels (termed mixels) even with higher spatial resolution data, as shown by Mather (1987). Moreover, individual patches of the same category may have different spectral signatures (see Bailey 1988 for ecological applications). The net result of these problems of sampling resolution and complex relationships between ground conditions and recorded signatures is that it is difficult to obtain classification accuracy better than 70% for any class on an image other than water, as consistently shown by various studies (Stuart 1996, pers. comm.)

#### **2.3.4 Secondary data acquisition**

Land surveying, photogrammetric and remote sensing techniques perform measurements and collect data directly from the field, from photographs (terrestrial and aerial) or from remotely sensed images. They are considered as primary data acquisition methods, and they lend themselves to current, reliable and direct measurements (Thapa and Burtch 1991).

On the other hand, depending on the specific requirements and projects, data may be acquired with sufficient accuracy from existing sources, such as maps, charts or

graphs, which are secondary products derived from the primary products. Map digitising is probably one of the most widely used secondary data acquisition methods. The most commonly used map digitising method is manual digitisation, performed by locating a cursor at the point or moving a cursor along the line to be digitised. The method also requires software to display and store the digitised data. Automated line-following techniques are devised to automate the manual digitising to some extent. They are widely used to digitise linear features, e.g., rivers, roads and contour lines. To minimise the uncertainties in the line-following process, different map separates may be used. Map or image scanning involves a computer-controlled instrument equipped with optics and detectors that can create the digital data from the document. The scanners may be divided into drum scanners, laser beam scanners, and video camera scanners. The scanning of a document may take only a few seconds, but the subsequent vectorisation and editing will take longer to complete, as reviewed by Fisher (1991).

Digitised data from maps will have uncertainties both derived from the maps themselves and introduced by the digitising process (Bolstad, Gessler and Lillesand 1990). Published maps have usually been subjected to a map generalisation process, which implies a variety of abstraction, selection, simplification and approximation (Joao et al. 1992). Thus, in some sense, maps portray selected features of reality in a highly abstracted and generalised form, as described by Harley (1975). Some maps can be strictly free of impurity, e.g., property parcel maps. In spatial data on natural resources such as vegetation, soil and land use, a major cause of errors is the omission of heterogeneous mapping units within areas delineated on the map as uniform (see Bailey 1988 and Chrisman 1989 for examples). This problem is caused by the oversimplification of a more complex real world.

The digitising of maps, either by manual or scanning methods, suffers further uncertainties imposed by machine and operator limitations (Thapa and Bossler 1992),



though scanning methods are more consistent and controllable in regard to the uncertainties during digitising than the manual methods. Dunn, Harrison and White (1990) discussed the uncertainties during their digitising of land use map data. Comparatively, uncertainties “locked in” during map generalisation are not easily identified and measured, unless more information is provided with respect to the specifications and procedures applied in map generalisation as a prerequisite. Therefore, the distinction between primary and secondary data acquisition is useful, because uncertainties in primary data are easier to identify and quantify than those in secondary data. Furthermore, categorical maps may have varying levels of detail mapped, resulting in the “slivers” problem encountered in map overlay processing. Further discussion on the characteristics of maps and map data can be found in Fisher (1991) and Flowerdew (1991).

## **2.4 Discussion**

This chapter began by introducing different forms of geographical abstraction: objects and fields. An object-based model with concepts of spatial entities, objects and attributes seems to be largely shaped by the conventional cartographic view, by which real geographical phenomena are modelled via a collection of discrete points, lines and areas (Even continuously varying surfaces are commonly displayed by means of contour lines!). Object-based models are suitable for well-defined phenomena. However, this suitability should be understood on relative terms, because geographical variations and details can rarely be duplicated by discrete objects. Alternatively, a set of fields is commonly used to model, usually continuously varying, geographical phenomena, and is suitable for physical geography and environmental studies. But, again, field models are by no means perfect representations of a highly complex real world.

As a result of complex interaction of human and machine factors, there exist various

uncertainties of different patterns and magnitudes during the acquisition of spatial data, no matter how advanced are the methods and techniques for data acquisition. The usefulness of the data will depend on their reliability and quality in respect of the specific applications or purposes. The assessment of data quality will be difficult without the necessary information about the raw data used and the procedures applied in generating the data. So, this is one of the reasons why further studies on spatial data themselves are crucial to appropriate applications of GISs. The uncertainty issues will be discussed in more detail in the next chapter. Furthermore, it is due to the uncertainties residing in spatial data that it is not straightforward jointly to represent and analyse disparate spatial data that may vary in structure, currency, resolution and level of human interpretation. A similar observation was made by Dahlberg (1986) in the context of a land information system combining data from different sources. It is thus of vital importance in GIS to address the data integration issues from a fresh recognition of the differences existing between the real world being investigated and the models being utilised.

As a further consideration, there are many geographical phenomena such as land cover, geology, soil and vegetation, which cannot be suitably modelled because they do not consist of well-defined and discrete entities. For example, Mark and Csillag (1989) identified that there are various types of lines: some refer to real entities such as rivers and roads, and thus can be recorded at high accuracy; others are abstracted objects, such as boundary lines on soil maps, which are only meaningful in their specific contexts. Though categorical maps are traditionally employed to represent such kind of phenomena as discrete polygons or contiguous patches of cells, these "objects" are actually abstracted objects, which do not exist in the real world. This is because there are spatial variabilities present in individual polygons or pixels and the boundaries of adjacent polygons are transition zones of varying patterns. In this case, real world phenomena are considered "fuzzy". Fuzzy phenomena such as fuzzy boundaries are recurrent topic of concerns (Burrough and Frank 1996; Wang and Hall

1996). Obviously, representing fuzzy phenomena by using discrete points, lines and areas will be a severe simplification, thus leading to significant uncertainties. In this case, a field-based model may be more suitable, because field models, especially those represented by grids, provide a mechanism by which spatial variabilities and heterogeneities are effectively captured, as developed below.

As seen in Section 2.2, polygons and grids are the two fully field-based models that are used for both categorical and numerical variables. A further distinction can be made here. By a grid model, geometric forms of the objects created are independent of the spatial distribution of the variable under study. But, in polygon models, geometries of objects are determined by the underlying phenomenon under study. For example, uncertainties in deciding soil types or soil pH values can move the boundary of a polygon. So, polygon models are less useful because of non-separability of position and attributes. In other words, when a numerical variable is under study or when a categorical variable can be transformed to continuous data, it is the grid model that is chosen as the most suitable for modelling a continuously varying variable, which is often perceived as a surface (Goodchild 1989b).

A surface contains an infinite number of locations and values. For instance, elevation can be measured at any location on the terrain; each x,y location has a surface value. Surface data are able to capture and represent spatial variabilities and heterogeneity intrinsic to many geographical phenomena. For example, in most natural resources applications, elevation data represent an important component, both directly and indirectly by the means of derivatives such as slopes and aspects (Mark, Lauzon and Cebrian 1989). There are other highly varied surface data such as vegetation continua data (see Wood and Foody 1989 for their application in lowland heaths), rainfall, snowfall and temperature data (Atkinson 1996, pers. comm.). The spatial variabilities and heterogeneities seen in these types of surface data are the basis for deriving the uncertainties in each specific data set, as will be established in the next two chapters.

## **Chapter 3**

### **Object and field perspectives of uncertainties**

#### **3.1 Introduction**

As was established in the previous chapter, spatial data are distinguished from other kinds of data in terms of their underlying spatial complexities, dependencies, variations and their alternative data modelling methods. It has also been shown that spatial data will always contain certain elements of compromise in the process of geographical abstraction and data acquisition, and are thus liable to various uncertainties.

In handling uncertainties in GISs, the separate components are their identification, measurement and representation. As the identification issue has been discussed in the previous chapters, this chapter will consider the measurements of uncertainties of spatial data, followed by a discussion of representations of uncertainties in spatial databases. As there are object-based and field-based data models, this chapter will discuss how uncertainties are described, measured and represented under object and field perspectives of spatial data modelling. Such a strategy may lead to an understanding that object-based methods, though often adopted in land cover mapping, find themselves unable to cope with spatial variabilities and heterogeneities deemed important for modelling uncertainties in land cover mapping, whereas field-based data methods lend themselves to effective uncertainty modelling.

#### **3.2 Uncertainties in objects**

It has been seen that an object-based model perceives the real world as consisting of

discrete points, lines and areas, which have positions and attributes. Because of the generalisation and approximation implied in the process of geographical abstraction and limitation unavoidable during spatial data acquisition, there are various uncertainties concerning discrete objects.

Uncertainties of different kinds need to be properly classified before being further discussed with respect to their measures, estimations and representations. Aronoff (1989), again, provides a good account of the common sources of uncertainties encountered in GIS applications, from the data sources, through data storage and manipulation, to data output and uses of GIS data results. Drawing upon the discussions presented in Bedard (1987), Chrisman (1991), Flowerdew (1991) and Veregin (1989), this thesis adopts the definitions as explained in the text below, some of which have been previewed in Chapter 1. Three types of uncertainties have been recognised earlier: conceptual uncertainties, measurement uncertainties and processing uncertainties. Conceptual uncertainties are associated with the process of abstraction or generalisation about the real world. Measurement uncertainties are uncertainties in position, height and attributes of objects. Processing uncertainties are primarily due to geo-processing such as data transformation from a vector structure to a raster structure or *vice versa*.

Measurement uncertainty can be differentiated into positional (or locational, or cartographic) uncertainty, uncertainty in the positions of points, lines and polygons, and attribute (or thematic) uncertainty, uncertainty in the values given to the attributes of objects. Attribute uncertainty can itself be subdivided according to the scale at which the attribute is measured: numerical uncertainty, if it is measured on a continuous scale such as elevation or land value, and categorical uncertainty, if it is measured on a classification system such as the ownership of a land parcel. Based on the account established above, Table 3.1 below summarises the different types of uncertainties in objects, using examples from land cover mapping.

Table 3.1 Definitions of different types of uncertainties in spatial data  
with examples in digital land cover mapping

Uncertainties and their subdivisions	Occurrence	Examples
1. Conceptual	geographical abstraction	whether to isolate small patches of trees from the surrounding grass fields
2. Measurement	object position and characterisation	
(a) Positional	position measurements	uncertainties in digitising the boundaries of specific land parcels from maps or aerial photographs
(b) Attribute	attribute measurements	
(i) Numerical	numerical data	uncertainties in measuring the spot heights for certain critical points from aerial photographs
(ii) Categorical	categorical data	uncertainties in land cover classification from aerial photographic interpretation or remotely sensed images
3. Processing	geo-processing	uncertainties due to data transformation from a vector to a raster structure <i>or vice versa</i> .

Now that uncertainties have been classified, the following text concentrates on a consideration of the uncertainties of measurement. Conceptual uncertainties and processing uncertainties are not discussed further in this thesis. While uncertainty is generally concerned with the deviation between the measured or computed value and the true value or the value accepted as being true, accuracy measures the closeness

between them. Therefore, it is suitable to say that the measure of uncertainty is opposite in meaning to that of accuracy. As such, measures of uncertainties are interchangeable with measures of accuracies.

Further, it is useful to clarify the concepts of accuracy and precision. Precision measures the degree of conformity of measurements among themselves. For example, the average difference between the surveyed data and their corresponding map data would estimate the accuracy of map data, while frequency distribution could be used to establish the precision of map data, as explained in Bolstad, Gessler and Lillesand (1990). Thus, high precision may not guarantee good accuracy because there are no cause and result relationships between precision and accuracy. Moreover, the confusion and misunderstanding about precision and accuracy may be dangerous, as spurious precision is at the core of the “slivers” problem (Goodchild 1991).

According to NCDCCDS (1988), there are six fundamental components to the accuracy of digital cartographic data: lineage, positional accuracy, attribute accuracy, logical consistency, completeness, and temporal accuracy. Lineage refers to a description of the source material from which the data were derived, and the methods of derivation involved in producing the final digital data, including details such as the specific control points used and the computational steps taken. Lanter and Veregin (1990) developed a program for propagating measures of spatial and thematic error, in which lineage was used to document the quality of derived GIS data products. Maintenance of logical consistency may be tested with respect to topology. Completeness of the data quality report would include such information as selection criterion, definitions, and relevant mapping rules, as well as geometric thresholds such as minimum width and minimum area. Temporal accuracy refers to the currency of data. Lineage, logical consistency, completeness and temporal accuracy have been discussed in detail in the textbook edited by Guptill and Morrison (1995); thus are not dealt with further in this thesis. Instead, attention is concentrated on positional and

attributes accuracies and how they are described, depending on the types of objects under study.

The simplest objects are points. The accuracy of the position, in coordinates, of a point can be represented in the form of an ellipse centred on the point, if the uncertainties in x and y coordinates are represented by Gaussian distributions (Goodchild 1991). When the uncertainties in both coordinates are the same and independent, the ellipse becomes a circle, again centred on the point. This is the circular normal model of positional uncertainty (ASPRS 1989).

Using this model, it is possible to compute the probability that the true position lies within any given distance of the measured position, with the average distortion expressed in the form of a standard deviation. When the standard deviations in x and y are not equal, the circular standard deviation is no longer the same as the standard deviations in the two coordinates, but is approximated as the mean of the two. The probability that a point's true position lies somewhere within the circle of radius equal to the circular standard deviation is 39.35 %. Under the circular normal model, a radius of 2.146 times the circular standard deviation will contain 90 % of the distribution (ASPRS 1989; Goodchild 1991).

In vector databases, lines (polygon boundary lines) are represented as sequences of digitised points connected by straight segments. It would seem plausible to model the uncertainty of a line by modelling each point's accuracy, assuming that uncertainty in the line is entirely due to uncertainties in the points. However, the points are captured subjectively during digitising, and the uncertainties in digitising the points comprising the line are not independent, but tend to be correlated (Goodchild 1991). The relationship between true and digitised lines cannot therefore be modelled as a series of independent uncertainties in points.



Blakemore (1984) suggested that an epsilon error band can be used to indicate the accuracy level about a digitised line, so the true line will occur at some displacement from the measured position (see also Chrisman 1989). This model has been used in both deterministic and probabilistic forms (Goodchild 1991). In the deterministic form, it is considered that the true line lies within the band and never deviates outside it. In the probabilistic form, on the other hand, the band is understood as one standard deviation in width, so that one might assume that a randomly chosen point on the observed line had a probability of 68 % of lying within the band.

The assessment of accuracy for a numerical attribute can be based on the calculation of mean and standard deviation from a set of samples. The samples are often repetitive measurements of the attribute in or on the relevant object. It is then interpreted that the true value of the attribute will not deviate more than twice the standard deviation from the mean value with a probability of about 95 %. For categorical attributes, on the other hand, it is only suitable to judge whether a labelling or classification is correct or wrong. It is thus often the practice to calculate probability of a classification being correct or wrong based on adequate sampling.

After the description of a variety of measures for uncertainties, attention is now given to their estimation. For estimating positional and attribute uncertainties, there is a hierarchy of methods. Deductive estimates may be made in the light of the appropriate lineage information. At a higher level, repeated measurement can provide some internal evidence, which works by using the redundancy designed into many data collection systems (e.g., closure of traverse, contours expected to meet at sheet edges).

The most reliable method is by using accuracy assessment via an independent source of higher accuracy, though there are many test results on uncertainty estimation using either simulation (Dunn, Harrison and White 1990; Dutton 1992), or other methods



such as multiple delineation and interpretation of objects and their attributes (Middelkoop 1990). As specified by the American Society for Photogrammetry and Remote Sensing (ASPRS, 1989), an empirical, site-specific estimate of positional uncertainties can be produced via tests against an independent source of higher accuracy (see also Vonderohe and Chrisman 1985 for a case study). Depending on the specific requirements, the independent source of higher accuracy may be obtained through land surveying or derived from aerial photography. It is specified that the 'nominal positional accuracy' of the check survey be three times that required of the product to be tested. In all tests, it is important to achieve the reliability and adequate number of test data sampled in a well-distributed way.

After the check survey, there is a set of statistical measures to calculate for positional accuracy. Defining that discrepancies ( $d_i$ ) are the differences in coordinate as derived from the data sets to be tested and as determined by a check survey, the sample statistics used for assessing accuracy of map data are explained by the following formulae (applied in x, y coordinates):

$$m = \frac{1}{n} \sum_{i=1}^n d_i \quad (3-1)$$

$$s = \left[ \frac{1}{n} \sum_{i=1}^n (d_i)^2 \right]^{1/2} \quad (3-2)$$

$$sd = \left[ \frac{1}{(n-1)} \sum_{i=1}^n (d_i - m)^2 \right]^{1/2} \quad (3-3)$$

where:

n = sample size

m = sample mean of test point discrepancies

s = the square root of the averaged squared discrepancies

sd = sample standard deviation of these discrepancies.

The sample mean of test point discrepancies will be 0 if the discrepancies are of random nature. If they are not random, i.e., the value of m is not 0, there is an overall

shift in the discrepancies. Therefore, the sample standard deviation  $sd$  is introduced to take the systematic error ( $m$ ) from the individual discrepancy at each check point. Harley (1975) described how the accuracies of OS maps are tested by using this set of equations.

As an estimate of positional uncertainty for a line, the epsilon band width may be measured by the gross misfit between the observed line and the assumed true line. However, for an individual point, it is not so simple to measure the epsilon distance, as there is no obvious basis to select a point on the true line as representing a point on the distorted line (Goodchild 1991).

Accuracy for numerical attributes can be assessed by using similar techniques as in estimating positional accuracy of points. Accuracy assessment for categorical attributes such as land cover data, on the other hand, is carried out by first deciding on the suitable comparing units such as pixels in the case of remote sensing image (Campbell 1987). Reference data are usually provided by means of field survey, map data or aerial photography. It is necessary that the two data sets register to one another and they use mutually compatible classification systems with respect to number and identities of parcels and mapping detail.

Comparisons are then made based on superimposed overlaying or via sampling. A superimposed overlaying produces an overlaid map with combination of classes from the test data and the reference data. A sampling process, on the other hand, generates a set of samples according to pre-selected sampling schemes. It is straightforward to tabulate for each pixel or sample the predominant category shown on the reference map, and the category as shown on the map or image to be evaluated. A summation of this tabulation forms the basis for the construction of the uncertainty matrix. For the purpose of illustration, an error matrix is provided in Table 3.2 with symbolic entries ( $c_{ij}$ ,  $i, j = 1, 2, 3$ ), while Table 3.3 provides a numerical example.

Table 3.2 An error matrix for classification accuracy assessment

		Ground or reference data					
		1	2	3	row total	row marginal	user's accuracy
Classified or mapped data	1	$c_{11}$	$c_{12}$	$c_{13}$	$t_r(1)$	$p_r(1)$	$ua(1)$
	2	$c_{21}$	$c_{22}$	$c_{23}$	$t_r(2)$	$p_r(2)$	$ua(2)$
	3	$c_{31}$	$c_{32}$	$c_{33}$	$t_r(3)$	$p_r(3)$	$ua(3)$
	column total	$t_c(1)$	$t_c(2)$	$t_c(3)$	$t$		
	column marginal	$p_c(1)$	$p_c(2)$	$p_c(3)$			
	producer's accuracy	$pa(1)$	$pa(2)$	$pa(3)$			

$$t_r(i) = \sum_{j=1}^3 c_{ij}, \quad t_c(j) = \sum_{i=1}^3 c_{ij}, \quad t = \sum_{i=1}^3 \sum_{j=1}^3 c_{ij}$$

$$p_r(i) = t_r(i)/t, \quad ua(i) = c_{ii} / t_r(i), \quad p_c(j) = t_c(j)/t, \quad pa(j) = c_{jj} / t_c(j).$$

Table 3.3 A numerical example of Table 3.2

		Ground or reference data					
		soil	grass	forest	row total	row marginal	user's accuracy
Classified or mapped data	soil	24	9	0	33	0.33	0.73
	grass	10	25	5	40	0.40	0.63
	forest	2	5	20	27	0.27	0.74
	column total	36	39	25	100		
	column marginal	0.36	0.39	0.25			
	producer's accuracy	0.67	0.64	0.80			

$$p_o = (24 + 25 + 20) / 100 = 69\%$$

$$\text{chance agreement} = 0.33 \times 0.36 + 0.40 \times 0.39 + 0.27 \times 0.25 = 0.34$$

$$K = (0.69 - 0.34) / (1 - 0.34) = 53\%$$

The user's accuracy is the probability that a location labelled as category  $k$  ( $k = 1, 2, 3$ ) actually belongs to category  $k$ , and is a measure of commission error ( $= 1 - ua$ ); producer's accuracy is the probability that a location known to belong to category  $k$  is accurately labelled as category  $k$ , and is a measure of omission error ( $= 1 - pa$ ). The overall classification accuracy (i.e., the observed proportion of agreement or percent correctly classified) is calculated by dividing the sum of the diagonals in the error matrix by the total number of all matrix elements (i.e.,  $\sum(c_{ii}) / t$ ). This overall classification accuracy is denoted by  $p_o$ , which is usually reported as the percent correct classification. For example in Table 3.3,  $p_o = 69\%$ .

The overall classification accuracy represents the probability that a randomly selected location is correctly classified. But it takes no account of chance agreement, because even a purely random assignment of class labels will result in a positive value. In other words, the overall classification accuracy tends to give an inflated index about the classification accuracy, as Veregin (1995) described.

To remedy this shortcoming, the Kappa coefficient of agreement,  $K$ , has been recommended. As a quantitative measure of classification accuracy,  $K$  is calculated as following:

$$K = (p_o - p_e) / (1 - p_e), \quad (3-4)$$

where  $p_o$  is the observed proportion of agreement as explained previously, and  $p_e$  is the proportion of agreement that may be expected to occur by chance, which is calculated from the row and column marginals of the error matrix from summation of  $p_r(i)p_c(j)$  (refer to Campbell 1987 for detail). For example, in Table 3.3, the chance agreement is 0.34, while the Kappa coefficient is 0.53. Kappa coefficient indicates that the accuracy of a classification is  $K * 100\%$  better than the accuracy that would result from a random assignment (Janssen and van der Wel 1994). It has become a widely used measure of classification accuracy, because all elements in the classification error matrix contribute to its calculation, and because it compensates for

chance agreement. See Foody (1992) for some examples.

So far, a variety of uncertainty measures and their estimates in object-based models have been described. It is necessary to check how useful this set of models and methods is for the representation and handling of uncertainties in the context of land cover mapping. Land cover data are usually mapped as categorical maps, on which a discrete collection of polygons are labelled to represent the different land cover types according to a chosen classification system. The accuracies of boundaries are described by epsilon band models, which are useful to address the issue of transition zones. However, epsilon band models are unable to represent within-polygon heterogeneities.

Though meant to assess classification accuracy, error matrixes are not suitable to describe accuracy below the level of individual classes. This is because an invariant accuracy of classification is implied for individual classes or polygons. As a result, within each polygon, spatial heterogeneities are suppressed, though they are sometimes implied in a textual report specifying inclusion and minimum mapping unit.

When dealing with propagation of uncertainties during geo-processing such as map overlay, the problems confronting object models are even worse. A lot of slivers are created as a result of slight differences in positional accuracies of the map layers involved, or as a result of different interpretations recorded on different map layers (Chrisman 1989; Dougenik 1979). Though Chrisman (1989) attempted to isolate the effects due to positional accuracy and attribute accuracy, his work, as any other object-based methods, fell short of an integrated approach to modelling positional and attribute uncertainties.

In an attempt to remove the conflicting effects of positional uncertainties, Veregin (1989 and 1995) adopted a field-like strategy, and used probabilistic methods to

investigate the propagation of uncertainties in map overlay. However, his work was based on error matrixes, whose general use was hence rather limited, because no spatial heterogeneities are incorporated. Besides, in his work, spatial independence in uncertainties across map layers is assumed. Therefore, field-based approaches need to be further explored in the next section.

### **3.3 Uncertainties in fields**

As implied in the previous section, ground truth exists for all objects, that is, objects have true values of positions and attributes, which can, in principle, be determined. Therefore, it is possible to measure positional and attribute uncertainties of objects by testing against ground truth data. This assumption is suitable for real, well-defined entities such as buildings and land parcels, and invoked in many of the underlying principles in object-based handling of uncertainties.

As seen in Section 2.2, field models consider the real world as a set of single-valued functions defined everywhere. Depending on the scales of measurements, there are numerical variables and categorical variables. For a numerical variable such as elevation  $z$ , it is possible to measure the uncertainty of  $z$  at a point by its root mean square error (RMSE). Suppose a RMSE of 50 cm is recorded for a spot height of 100 m; it is then interpreted that the true height will be within the range 99 m to 101 m with a probability of about 95 %. Hunter and Goodchild (1995) employed similar methods in their research on uncertainties involved in applications of DEMs of a flood plain.

Uncertainty measures such as RMSE may be derived from multiple sampling, which, though very costly, serves to provide raw data to calculate means and standard deviations. This is only possible for topographic variables like elevation, which, as a parameter of a visible and well-known landscape, can be measured several times. For

other variables such as urban noise level, atmospheric pressure and population density, however, it is impossible to go back to collect a few more samples once a specific sampling is done. Thus, multiple sampling can not be used to measure uncertainties involved in these variables. Besides, such variables are usually invisible and changeable, and true data are known, to some extent, only by extra sampling (Carter 1988). These extra samples are used to detect gross aberrations, but provide no basis to argue over small detail. Further, interpolation of samples of such variables is subjective, because it is only possible to judge whether an interpolated surface is plausible or absurd, rather than right or wrong.

In this case, kriging is often used to generate a variance surface, as a by-product, for an estimated surface (Oliver and Webster 1990). So, for every point in an estimated surface, a measure of uncertainty is readily available from the variance surface. The interpretation of a variance surface may follow similar rules as RMSE, on the assumption of normal distribution (Bregt 1991).

For categorical variables such as land cover, a suitable method might be to consider the situation in which every point of a field is concerned with a discrete outcome such as a nominal or an ordinal label in a classification system. Thus, an appropriate measure of uncertainty for a categorical variable at a point is the probability that a classification is not correct. Such a probabilistic measure may be obtained from repetitive sampling as described previously for elevation data. It is also possible to use an error matrix to derive a variety of measures such as the overall classification accuracy and Kappa coefficient.

The measure of uncertainty as mentioned above might suggest a spatially invariant level of uncertainty for each class. However, it is rarely the case that a field-wide classification can be reached with equal accuracy. This is because many variables, as properties of natural phenomena, exhibit significant spatial variations, which will exist



at all scales (Burrough 1987; Goodchild 1989a). Moreover, spatial variations arise as a result of ambiguities involved in the process of definition and classification necessary for incorporating categorical variables in spatial databases (Goodchild 1989a). Suppose a class of grassland is found for two patches A and B. It is not unusual to find that the growth of grass at these two patches is distinctively different. Even within each patch, the condition is not quite the same.

The spatial variations mentioned above suggest that many categorical variables are continuously varying, and so are the uncertainties. Thus, suppose  $c$  classes are possible, a suitable approach might be to view a categorical variable as multi-nominal field  $p_i(x)$ , where  $p_i(x)$  represents the probability of point  $x$  belonging to a candidate class  $i$  ( $i = 1, \dots, c$ ) (Goodchild, Lin and Leung 1994). So, instead of a uniform level of uncertainties at individual class level as suggested by an error matrix, the use of  $p_i(x)$  facilitates a continuously varying representation of uncertainties.

While probabilistic measures are used above, many variables are often fuzzy, as fuzziness is seen as an inherent feature of categorical or qualitative measurements in classifications such as for land cover and land use. This idea has also been suggested by Altman (1994) in an introduction to fuzzy spatial relations. There are also situations as in the case of spatial decisions based on multiple criteria such as land suitability assessment, where classical set theory may not be suitable, as discussed by Drummond and Ramlal (1992). Therefore, fuzzy approaches are needed for describing and modelling fuzzy phenomena.

Fuzzy set theory has been developed specifically to deal with fuzzy phenomena (Kaufmann 1975; Zadeh 1965). Let  $X$  be a universe of discourse, whose generic elements are denoted  $x$ :  $X = \{x\}$ . The membership in a classical set  $B$  of  $X$  is often viewed as a characteristic function  $\chi_B$  from  $X$  to  $\{0,1\}$  such that  $\chi_B(x) = 1$  if and only if  $x \in B$ . A fuzzy set  $A$  in  $X$ , on the other hand, is characterised by a fuzzy membership

function  $\mu_A(x)$ , which associates with each  $x$  a real number ranging from 0.0 to 1.0 (Wang 1990). The fuzzy set  $A$  can be designated as:

$$A = \{(x, \mu_A(x)) \mid x \in X\} \quad (3-5)$$

where the value of  $\mu_A(x)$  at  $x$  represents the grade of membership of  $x$  in  $A$ , and is commonly termed as a fuzzy membership value (FMV). The closer the FMV is to 1.0, the higher the grade of membership of  $x$  in  $A$ . The fuzzy membership function  $\mu_A(x)$  may be probabilistic, but may also be based on unrepeatable expertise (Drummond and Ramlal 1992).

Fuzzy set theory is versatile, because it allows many otherwise non-quantifiable phenomena to be usefully dealt with. In other words, fuzzy set theory provides a quantitative way to measure and model qualitative variables such as categorical land cover data, which is useful in decision-making (Gopal and Woodcock 1994). Besides, given uncertainties surrounding the measures of uncertainties and uncertainties, it is probably more sensible to use fuzzy set theory than rigorous statistical and probabilistic means (Drummond and Ramlal 1992). Therefore, FMVs will be used as a superset combining all kinds of fuzzy membership measures, be they fuzzy or probabilistic. Such a generalisation is useful because of the need to synthesise the past work concerning fuzziness in spatial databases, as will be developed in the following chapter.

So far, it has been seen that both probabilistic and fuzzy measures are useful for describing uncertainties in categorical variables. A question remains of how to estimate the required probabilistic and fuzzy measures. Because of the diversity of data incorporated in urban and suburban land cover mapping, the methods for estimating fuzzy land cover data seem to be rather complicated. A detailed discussion on the estimation of probabilistic and fuzzy measures of uncertainties in fields will be developed as Chapter 4, after a further section below describing representations of uncertainties.

### 3.4 Representing uncertainties in objects and fields

As shown in Sections 3.2 and 3.3, possible measures of uncertainties include qualitative descriptions such as a lineage report, and quantitative parameters such as standard deviations, probabilities and fuzzy membership values, depending on specific data models. After discussing the measures of uncertainties, attention is given to their representations in spatial databases. As predicted by Lunetta et al. (1991), the integration of spatial data seems to be complicated by the need to cope with different uncertainties in addition to the diverse data themselves. Therefore, it is very crucial to consider effective and flexible approaches to the representations of spatial data and their uncertainties. In practice, representations of uncertainties in spatial databases have followed some logical extensions of the underlying data models for the data themselves. These will be reviewed before going on to propose further synthesis and refinement of the techniques that are needed for spatial data in land cover mapping.

As seen above, spatial data modelling has been seen as being either object-based or field-based. There is another interconnected concept: data structure. It has been seen previously that spatial data modelling is the process of abstracting and generalising real geographical variations, thus expressing it in a suitable form for representation in a discrete digital store. A data structure, on the other hand, is the specific representation scheme within a computer system based on a certain data model (Peuquet 1984).

In general, there are two distinct data structures in GIS: vector and tessellation (raster or grid), as shown in Figure 3.1. Vector structure represents spatial variation using points and lines located in continuous coordinate space. Vectors may be unlinked, that is, their object boundaries are encoded without reference to neighbours, or topologically linked, that is, their arcs are referenced by their end points, orientation and the attributes of adjoining regions, as listed in Table 3.4. Table 3.4 provides an

example for the topologically linked vector structures, where each line has an ID, a from-node, a to-node, a left-polygon and a right-polygon. The vector structure similar to that shown in Figure 3.1 (a) in connection with Table 3.4 is known as a georelational data model in the ARC/INFO GIS package.

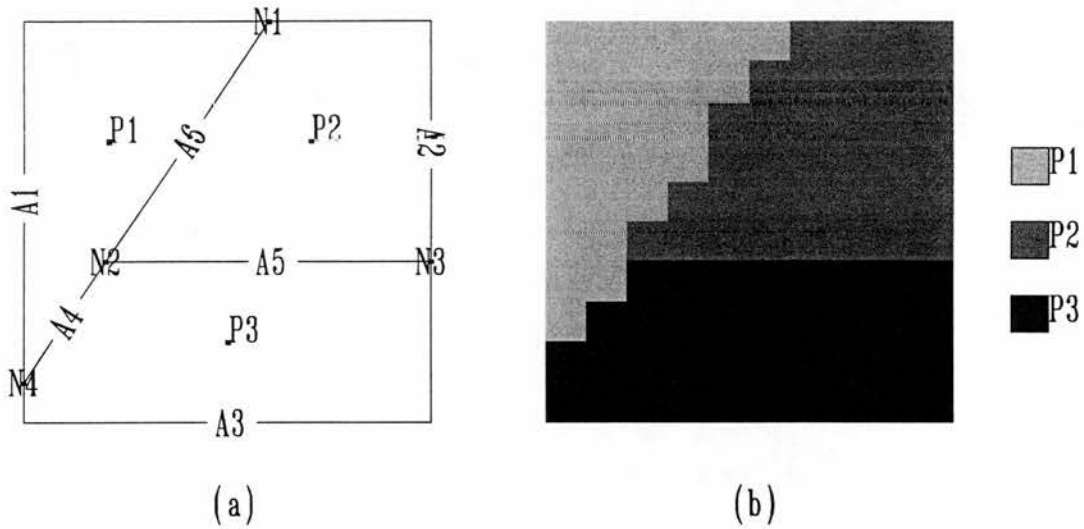


Figure 3.1 Data structures in GIS: (a) vector structure, (b) raster structure.

Table 3.4 A topological data structure for lines as drawn in Figure 3.1(a)  
(the 0s denote outer areas)

Rec#	ID	FNODE#	TNODE#	LPOLY#	RPOLY#
1	a1	n4	n1	0	p1
2	a2	n1	n3	0	p2
3	a3	n3	n4	0	p3
4	a4	n4	n2	p1	p3
5	a5	n2	n3	p2	p3
6	a6	n2	n1	p1	p2

A raster structure, on the other hand, tessellates space and assigns a unique identifier to each spatial element. The position of objects is defined implicitly by the row and

column position of the cells they occupy. The value stored for each cell indicates the type or condition of the object found at that position, as shown in Figure 3.1(b) (Aronoff 1989; Burrough 1986).

In a vector structure, the entity's identity is preserved, thus lending itself to storage of spatial data in a object-based data model, as argued by Davis and Simonett (1991). Because object-based data models describe the geographical phenomena as discrete collections of points, lines and polygons, the corresponding representation of uncertainty measures in databases is also object-based. It incorporates information on positional and attributes accuracy into vector-based spatial databases by attaching the accuracy parameters (e.g., standard deviations or probability indicators, as discussed in Section 3.2) to individual points, lines and polygons or individual attributes (e.g., district population counts) (Drummond and Ramlal 1992). A similar approach may be found in Guptill (1989), where data on the uncertainties are accommodated in an extended geo-relational data model similar to that illustrated in Table 3.5. This table includes hypothetical epsilon band widths for incorporating uncertainty of lines.

Table 3.5 An extended relational data structure for incorporating uncertainties of lines as drawn in Figure 3.1(a).

Rec#	ID	FNODE#	TNODE#	LPOLY#	RPOLY#	$\epsilon$ -WIDTH
1	a1	n4	n1	0	p1	1.0
2	a2	n1	n3	0	p2	1.0
3	a3	n3	n4	0	p3	1.0
4	a4	n4	n2	p1	p3	1.0
5	a5	n2	n3	p2	p3	1.0
6	a6	n2	n1	p1	p2	1.0

In vector structures, objects are defined by their positions and attributes. As a result, uncertainties in positions and attributes are usually addressed separately. The separate

handling of positional and attribute uncertainties has, unfortunately, been the core for weakness of object-based methods in modelling uncertainties (Goodchild 1992).

On the other hand, raster structures are favoured by researchers (Goodchild, Lin and Leung 1994; Veregin 1989). This is because attributes in a raster data structure are stored explicitly cell by cell, whose positions can be calculated by referring to the origin and cell size of the raster structure. The square grid, which is the most common, has additional advantages such as simplicity of structure and ease of processing. For these reasons, though often criticised for data redundancy due to a fixed resolution or the “pixel” size adopted by a raster structure, which sets the upper bound for its accuracy, raster structures are ranked highest in terms of suitability for modelling uncertainties in fields (Goodchild 1989b).

Therefore, one may represent the uncertainties in land cover data as a set of fuzzy surfaces using rasters with FMVs defined for every grid cell. Each surface represents the contribution from an individual class. The value on a given surface represents the strength or certainty of having that class at that given position. As an example, a set of fuzzy surfaces adopting raster structures is shown in Figure 3.2, where the value labelled for each grid cell stands for its strength of membership belonging to the named class.

As shown in Figure 3.2, using a raster-based structure, fuzzy surfaces enable FMVs to be explicitly recorded for each grid cell. Such a scheme facilitates the maintenance of spatially-varying class memberships. Thus, the uncertainties associated with class allocation are easily accessible during data analysis based on the fuzzy surfaces. Such a strategy is superior to a vector-based structure (Table 3.5) where uncertainties are recorded for individual objects and hence it is difficult to infer spatially-varying uncertainties at arbitrary locations unless a suitable model is assumed. Further discussion is provided in the next section.

100	90	80	70	70	60	40	30	20	10
100	100	80	70	60	40	30	20	10	0
100	100	80	60	40	30	20	10	0	0
100	90	70	60	40	30	20	10	0	0
80	70	60	40	30	20	10	0	0	0
70	60	30	30	20	10	0	0	0	0
70	50	30	20	10	0	0	0	0	0
60	40	30	20	10	0	0	0	0	0
40	30	20	10	0	0	0	0	0	0
30	20	10	0	0	0	0	0	0	0

(a)

0	10	20	30	30	40	60	70	80	90
0	0	20	30	40	60	70	80	90	100
0	0	20	40	60	70	80	90	100	100
0	10	30	40	50	60	70	80	100	100
20	30	40	50	50	60	70	70	70	70
10	20	40	50	60	60	60	60	60	60
0	20	30	30	30	30	30	30	30	30
0	10	20	10	10	10	10	20	20	20
0	0	10	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

(b)

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	10	10	10	10	0	0
0	0	0	10	20	20	20	30	30	30
20	20	30	20	20	30	40	40	40	40
30	30	40	50	60	70	70	70	70	70
40	50	50	70	80	90	90	80	80	80
60	70	70	90	100	100	100	100	100	100
70	80	90	100	100	100	100	100	100	100

(c)

Figure 3.2 A set of fuzzy surfaces with hypothetical data:

(a) class P1, (b) class P2, and (c) class P3, distributed as in Figure 3.1.

### 3.5 Discussion

There are many advantages to be gained from using surface models in the handling of uncertainties in spatial databases. The first advantage is that surface models are able

to capture the spatial variabilities of uncertainties, which are lacking in object-based data models. For example, derived fuzzy values indicate the relative strength of class membership across the classes being considered and can be used therefore to map the spatial variation in class especially when the study area displays gradual change in classes, as described in Wood and Foody (1993).

Categorical maps usually contain nominal data which have been stripped of spatial heterogeneities and fuzziness, whereas remotely sensed data inherently involve some continuous sampling in spatial, spectral and temporal domains. Differences in their data structures and also in the nature of the underlying data pose impediments to the effective integration of GIS and remote sensing data (Hutchinson 1982). As an alternative, by using surface methods, various data are represented in a unified format irrespective of their original formats, making spatial data integration more efficient and meaningful. This is the second advantage of surface methods.

In terms of the visualisation of data quality, the traditional approach on topographic maps is to use reliability diagrams produced to assist the potential users in determining the maps' fitness or suitability to specific applications or projects. There are also textual means to indicate overall levels of data accuracy: data resolution ("the smallest map unit represented is about 10 m X 10 m"), error ("map unit boundaries should be accurate within +/- 15 m; in areas labelled as specific land cover type, this cover type occupies at least 75% of the areas"). However, representing the overall spatial variability of data accuracy is more worthwhile than simple diagrams or texts, as such a strategy will greatly facilitate such processing as error propagation (Berry 1993). Three dimensional views built straightforwardly from surface data will give the spatial variability of the uncertainties in spatial data in an inherently visually appealing and convincing way, as will be demonstrated in the empirical tests. This constitutes the third advantage of the surface-based approaches.



The fourth and the most important advantage is that surface approaches provide an integral strategy to give more direct access to the uncertainties in the whole process of spatial databases, as outlined in Chapter 1. Clearly, uncertainty surfaces will provide the extra information necessary to determine the accuracy in a composite map. As mentioned in Section 3.2, Veregin (1995) presented empirical results for error propagation using probability values of individual map layers. Now with a continuously varying surface of uncertainty, this research would be able to be extended to a point basis. This is the main driving force behind the development of surface-based approaches as presented in this thesis. Moreover, the combined use of probabilistic and fuzzy measures will allow for a suit of quantitative analyses for handling uncertainties in spatial databases, as will be developed in the next chapter.

In summary, in line with the two alternative forms of spatial data modelling as covered in Chapter 2, uncertainties have been discussed with respect to their measures and representations in object-based and field-based models. Under object-based models, uncertainties were dealt with separately in positional and attribute domains. These uncertainties are appropriately viewed as special attributes of objects, and can thus be accommodated in extended geo-relational data structures. Under field-based models, on the other hand, fuzzy surfaces were proposed, which can give more direct access to the continuously varying nature of uncertainties during data input, update, geo-processing and production of output data.

While methods for estimating uncertainties in objects have been discussed, drawing on the past work, the estimation of fuzzy surfaces in categorical fields has not been examined. Though there have been useful proposals or practical tests aimed for the generation of fuzzy surfaces, the research has been largely non-systematic, in particular for urban and suburban land cover mapping. Therefore, the next chapter will endeavour to fill in this gap.

## Chapter 4

### Deriving and analysing fuzzy surfaces

#### 4.1 Introduction

As indicated in the last chapter, methods for estimating uncertainties in objects are relatively well-established. On the other hand, much work remains to be done to derive fuzzy surfaces as relevant for categorical fields. Moreover, as also discussed in the previous chapter, the feasibility of surface-based models and methods does not imply that deriving fuzzy surfaces will be easy or cheap in any sense. The reasons are discussed below.

For modelling fuzzy bands of polygon boundaries, Drummond (1987), Lowell (1994) and Middelkoop (1990) employed multiple interpretation techniques in order to calculate and model the magnitudes and distributions of uncertainties occurring in polygon boundaries. They were fortunate to have a number of students to perform the interpretation and delineation of polygon boundaries. In practice, probably due to limited budgets and time, spatial data acquisitions tend to be concerned more with practical and economic factors including constraints than about how redundant measurements might be used to check the data quality internally. Thus oversampling is likely to be something exceptional rather than usual. This fact implies that there are usually very limited amount of raw data, which may not be sufficient to derive fuzzy surfaces accurately. Therefore, some cost-effective methods are required to generate fuzzy surfaces in situations where only limited sets of data for checking are available.

Furthermore, empirical accuracy tests implicitly assume that the databases are populated by well-defined objects with exactly valued numerical or categorical attribute data. However, most naturally occurring phenomena such as land cover or

vegetation may not be as well depicted as generally homogeneous regular or irregular units. As a result, there are no ground truth data in any absolute sense for poorly or fuzzily defined phenomena. Therefore, it is necessary to address the problems in constructing fuzzy surfaces from a relatively new perspective.

Because land cover, as a special categorical variable, is the focus of this research, attention is given to the derivation of fuzzy surfaces for land cover data. Chapter 2 described a range of spatial data acquisition methods including photogrammetric and remote sensing techniques. It is now useful to make some further distinction among these commonly used methods of land cover data acquisition, in order to provide initial insights into how different fuzzy surfaces can be derived, depending on the specific data types being acquired.

The first group of spatial data acquisition methods is represented by digital image processing. It has been shown previously that remote sensing images, largely in the form of digital multi-spectral reflectance data, readily lend themselves to a variety of digital processes (Richards 1993). Increasingly computerised image processing techniques enable remote sensing analysts to perform a variety of image processing procedures and to derive a wide range of data products. For remote sensing classification, such highly computerised image processing tends to result in the analysts rushing to the classified data after playing with the classifiers provided in an image processing system, implying that the intermediate procedures and outputs are not likely to be studied at any length once a satisfactory classification has been achieved. Such a kind of classification is called “hard” by other workers (Foody 1995a and 1995b), because only the final labelled product is generated, and is also called hard in this thesis. “Hard” is used as the opposite of “fuzzy” and in preference to “discrete”.

For example, a maximum likelihood classification involves the computation and

comparison of probabilities of a pixel belonging to all candidate classes, and assigns the pixel to the class to which it has the biggest probability of belonging. These probability data are rarely explored further by the analysts and are seldom transparent to the end users of the classified data. These intermediate data, however, often directly indicate or are indirectly related to the relative strengths of membership of a pixel to all candidate classes, and thus are vital to the derivation of spatially varying uncertainties for the classified data (Foody and Cox 1994). Clearly, it is useful if the relative strengths of class membership are maintained throughout the classification processing. Such a strategy leads to a fuzzy classification, as opposed to a hard classification. Usually, a fuzzy classification can be carried out by defining a suitable fuzzy membership function or by clustering in the spectral domain, as will be described in the next section.

The second group of methods for categorical mapping is exemplified by aerial photo interpretation, which constitutes a substantial element in photogrammetry. As seen in Chapter 2, in the process of aerial photo interpretations, interpreters simultaneously make use of the information on association, tone, geometry, texture and pattern as well as a variety of knowledge available from field survey and existing maps (Jensen 1983). The classification is carried out in accordance with certain hierarchical classification schemes. Traditionally, the information on categorical maps is shown in the form of discrete polygons, each of which is internally homogeneous and can be differentiated from adjacent polygons by sharp boundaries. The information on the heterogeneity and spatial variability present in the reality may be apparent to data acquisition personnel, but is rarely further exploited for assessing the spatial variations of the classification accuracy. By conventional methods, classification accuracies are assessed on a global basis, or are spatially or thematically aggregated, implying that no information is provided of the spatially varying accuracies below the level of individual classes.

To provide further detail, accuracy assessment for individual polygons can be achieved by noting the proportion of the component classes in each regular or irregular mapping unit (Foody 1995a). This method can be applied in situations where the mixture of classes is significant in mapping units due to spatial aggregation. For example, it is often the case that pixel sizes of remote sensing images are relatively large in relation to the compressed features in urban areas which need to be resolved. When photogrammetric data, often in the form of vectors, are used as either ancillary data or ground truth data, different land cover types are expected to be present in many pixel-equivalent grid cells transformed from the photogrammetric data, usually stored as polygonal data. Therefore, one method for constructing fuzzy surfaces is to maintain the proportions of sub-pixel component land cover types. Under such a strategy, information is made available not only for the dominant land cover types but also for the whole set of component land cover types at the individual pixel level.

However, this method implicitly assumes that boundaries separating component land covers can be accurately delineated. This assumption is usually unattainable in situations where the mapped phenomena are inherently fuzzy and spatial variations occur at all scales. In this case, other methods have to be devised.

A possible way is to go back to the heterogeneity and spatial variability central to many naturally occurring phenomena. Goodchild (1989a) recognised that, for cartographical map compilation in general, field observations are firstly focused on areas with relatively homogeneous characteristics, and are then extended to heterogeneous areas using aerial photographs or remotely sensed images. Obviously, homogeneous locations can be identified more accurately than heterogeneous, transitional zones, suggesting the plausibility of using interpolation to generate uncertainty surfaces, as will be described in Section 4.3. To some extent, the derivation of fuzzy surfaces from photogrammetric data lies in the data acquisition process itself.

The way that each type of fuzzy surface is calculated depends on the nature of the underlying data sources. Table 4.1 shows the two types of source data and suggests the means for deriving fuzzy surfaces from them. Further details on the variety of digital and graphical images available may be found in Neto (1996). As the first aim, this chapter will define the ways in which fuzzy surfaces are derived for each of the common data types exemplified in Table 4.1.

Table 4.1 A classification of methods for deriving fuzzy surfaces

Source data	Data format	Derivations
1. Digital images derived from: <ul style="list-style-type: none"> <li>• Landsat MSS, TM</li> <li>• SPOT HRV</li> <li>• NOAA AVHRR</li> <li>• RADARSAT SAR</li> <li>• Space shuttle MOMS, ESC</li> <li>• AMSS, ATM</li> <li>• scanned aerial photographs</li> </ul>	raster	fuzzy classification: <ul style="list-style-type: none"> <li>a. supervised - fuzzy membership functions</li> <li>b. unsupervised - fuzzy <i>c</i>-means clustering</li> </ul>
2. Graphical images derived from: <ul style="list-style-type: none"> <li>• Space shuttle MC, LFC</li> <li>• Russian MIR KAP-350</li> <li>• RESURS-F KFA-1000</li> <li>• Aerial photographs</li> </ul>	vector	<ul style="list-style-type: none"> <li>c. proportional component land cover types in mapping units</li> <li>d. interpolation based on distances</li> <li>e. interpolation by indicator kriging</li> </ul>

Representing uncertainties by fuzzy surfaces is different from using object-based methods, as demonstrated by Lowell (1992). As the second aim, some useful methods for the analysis of fuzzy surfaces, following the discussion of their derivations, will be introduced.

As implied in Table 4.1, fuzzy surfaces need to be understood differently, depending on the underlying data and the ways the fuzzy surfaces are derived. Thus, the final section of this chapter will clarify the distinction between different types of fuzzy surfaces.

## 4.2 Deriving fuzzy surfaces from remotely sensed images

### 4.2.1 Supervised methods

As indicated in Table 4.1, fuzzy surfaces can be derived from fuzzy classifications, which can be derived by both the supervised and the unsupervised methods. By using a supervised method, a fuzzy membership function needs to be defined. Depending on the specific classification techniques being used, there is a variety of ways of defining fuzzy membership functions required to derive supervised fuzzy classifications.

The maximum likelihood classification is a widely used supervised classification method. Assuming normality, the maximum likelihood classification is based on an estimated probability density function,  $p(x/i)$ , as expressed by equation (4-1) for each of the reference classes under consideration:

$$p(x/i) = p(i)(2\pi)^{-(n/2)} |\text{Det}(\text{cov})|^{-1/2} \exp(-d^2/2) \quad (4-1)$$

where  $n$  is the number of bands,  $p(i)$  is the *a priori* probability of class  $i$ ,  $\text{cov}$  is the variance and covariance matrix of class  $i$ ,  $\text{Det}$  denotes the determinant, and  $d^2$  is the squared Mahalanobis distance from pixel  $x$  to the centre of  $i$  (Campbell 1987). The squared Mahalanobis distance of a pixel (denoted by vector  $x$ ) to the centre of a class  $i$  (denoted by mean vector  $v_i$ ) is calculated by:

$$d(x, v_i)^2 = (x - v_i)^T \text{cov}^{-1} (x - v_i) \quad (4-2)$$

where superscript  $T$  stands for a transpose operation, and superscript  $-1$  denotes the inverse of a matrix.

In the conventional maximum likelihood classification, pixels are sorted into classes

with which they have the highest probabilities, measured by the probability density function. In a fuzzy classification, on the other hand, the probability data computed before the final labelling process are kept for further uses. This technique is illustrated in Figure 4.1. Specifically, Figure 4.1(a) shows scattered data points representing the pixels (a particular pixel is indicated by  $x$ ) in a two-dimension multispectral space, where classes 1, 2 and 3 are defined by their bounding ellipses centred at point  $v_1$ ,  $v_2$  and  $v_3$ . Figure 4.1(b) shows the normal distribution curves along the transects passing the pixel  $x$  and each of the centres.

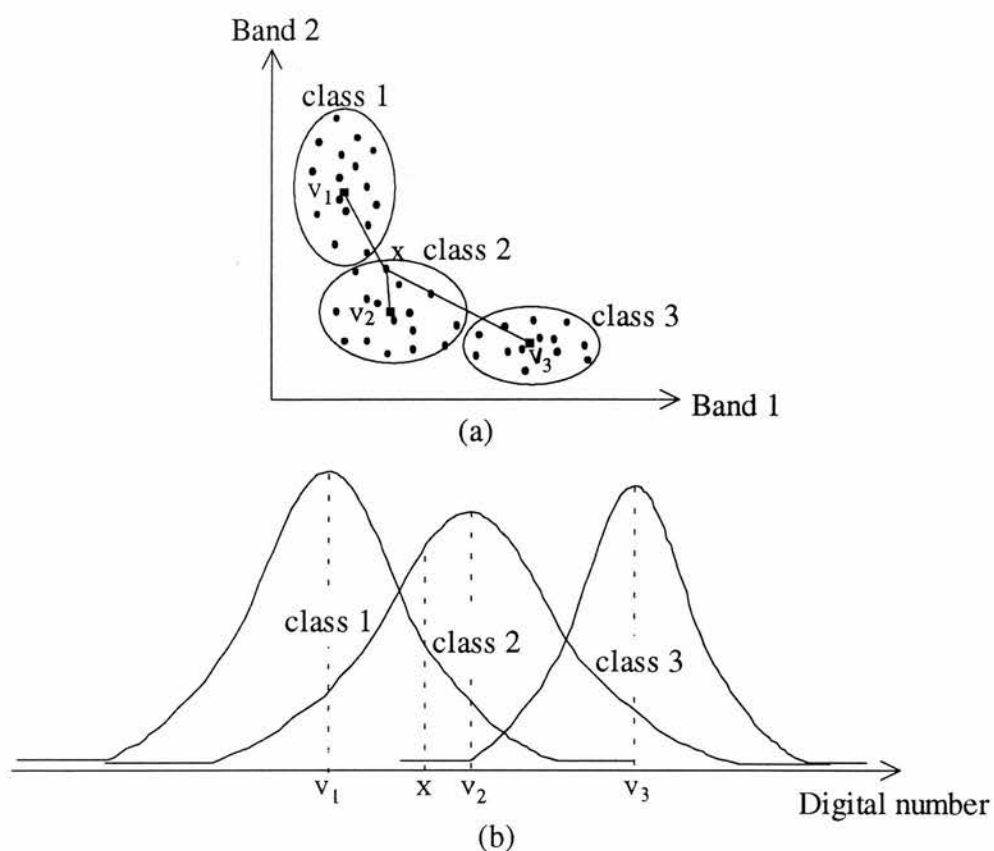


Figure 4.1 The principle of a fuzzy classification.

The probability density function values of pixel  $x$ , measured by the three normal distribution curves as shown by Figure 4.1(b), indicate the relative strengths of membership of pixel  $x$  belonging to class 1, 2 and 3, respectively. Thus, a suitable fuzzy membership function can be defined as follows:



$$\mu_i(x) = p(x/i) / \sum_{j=1}^c p(x/j) \quad (4-3)$$

where  $\mu_i(x)$  is the fuzzy membership function of pixel  $x$  belonging to class  $i$ ,  $c$  is the total number of classes, and  $p(i)$  and  $p(x/i)$  are as defined in equation (4-1). Further detail on this fuzzy membership function is provided in Wang (1990), with application of a supervised fuzzy classification expert system in a land cover change detection.

In essence, FMVs as calculated by equation (4-3) are probability density values normalised by their sum across all the pre-defined classes. Thus, for each pixel  $x$ , fuzzy membership values across all classes (i.e.,  $\mu_1(x)$ ,  $\mu_2(x)$ , ...,  $\mu_c(x)$ ) can be derived using equation (4-3). For each class  $i$ ,  $\mu_i(x)$  values are assembled pixel by pixel, thus forming a fuzzy surface. If there are  $c$  classes,  $c$  fuzzy surfaces can be derived, as exemplified in Figure 3.2.

In a maximum likelihood classification, the labelling can also be done by assessing *a posteriori* probability of membership on the assumption that the pixels belong to one of the pre-defined classes. The *a posteriori* probability of a pixel  $x$  belonging to class  $i$ ,  $L(i/x)$ , may be determined from equation (4-4) below:

$$L(i/x) = p(i) p(x/i) / \sum_{j=1}^c p(j) p(x/j) \quad (4-4)$$

where  $c$  is the total number of classes,  $p(i)$  and  $p(x/i)$  are as defined in equation (4-1). These probabilistic measures are used as FMVs by Fisher and Pathirana (1989) and Foody, Campbell, Trodd and Wood (1992).

When the Mahalanobis distance based classifier is used, probabilities of class membership are referred from squared Mahalanobis distance ( $d^2$ ) to the  $\chi^2$  distribution with  $n$  degrees of freedom, where the value  $n$  is the number of spectral bands comprising the reflectance data. This kind of probability is understood as the proportion of pixels at a distance further than  $d^2$  from the centre of the class  $i$

belonging to the class  $i$ . Again, FMVs may be derived by normalising probability measures, as described in equation (4-3).

Nowadays, artificial neural networks are attractive for use in the classification of remotely sensed imagery and have gained increasing popularity in remote sensing (Foody 1996). When a neural network is used for classification, the strength of class membership can be measured by the activation level of the network output units. Further detail can be found in Foody (1996). Moreover, FMVs can be derived from the so-called nonparametric classification approach, such as that reported by Skidmore and Turner (1988), which was found of particular usefulness in areas with mixed spectral responses.

#### 4.2.2 Unsupervised methods

As indicated at the beginning of this section, fuzzy classification can be performed in both supervised and unsupervised modes. In an unsupervised mode, fuzzy surfaces can be derived from the fuzzy  $c$ -means clustering (Bezdek, Ehrlich and Full 1984). Fuzzy  $c$ -means clustering is a form of cluster analysis as discussed in general statistics. Cluster analysis refers to a broad spectrum of methods which try to subdivide a data set  $X$  into  $c$  subsets termed clusters.

Let  $X = \{x_1, x_2, \dots, x_n\}$  be a sample of  $n$  observations in  $n$ -dimensional Euclidean space;  $c$  is an integer not greater than  $n$ . A fuzzy clustering is represented by a fuzzy set  $M_{fc}$  with reference to  $n$  observations and  $c$  clusters, which is defined as:

$$M_{fc} = \{U_{c \times n} \mid \mu_{ik} \in [0.0, 1.0]\} \quad (4-5)$$

where  $U$  is a real  $c \times n$  matrix comprised of elements denoted by  $\mu_{ik}$ , and  $\mu_{ik}$  is a fuzzy membership function expressing the FMV of an observation  $x_k$  to the  $i$ th cluster. The value of FMV ranges between 0.0 and 1.0 and is positively related to the degree of similarity or strength of membership for a specified cluster, as shown by the closed

interval. Besides, it is required that the values of FMVs for an observation  $x_k$  sum to 1 across all clusters, because FMVs can represent either probability or a certainty factor associated with a fuzzy set.

Unlike a supervised fuzzy classification, a fuzzy  $c$ -means clustering does not work by defining a fuzzy membership function beforehand, but rather the values of  $\mu_{ik}$  are derived by exploring the coherent structure of a particular set of data points. In this sense, the fuzzy  $c$ -means clustering is similar to the clustering typically used in a conventional unsupervised classification such as the Iterative Self-Organising Data Analysis technique (ISODATA) (Campbell 1987; Richards 1993).

There is a variety of algorithms aiming for an optimal fuzzy  $c$ -means clustering. One method works by minimising a generalised least-squared errors function  $J_m$ :

$$J_m = \sum_{k=1}^n \sum_{i=1}^c (\mu_{ik})^m (d_{ik})^2 \quad (4-6)$$

where  $m$  is the weighting exponent which controls the degree of fuzziness (increasing  $m$  tends to increase fuzziness; usually, the value of  $m$  is set between 1.5 and 3.0),  $d_{ik}$  is the distance between each observation  $x_k$  and a fuzzy cluster centre (Bezdek, Ehrlich and Full 1984). Usually, the Mahalanobis distance is used in remote sensing.

Suppose  $c$  clusters are aimed for, with a particular weighting exponent being equal to  $m$ . To optimise the function  $J_m$  as defined in equation (4-6), an algorithm for a fuzzy  $c$ -means clustering can be programmed as follows:

(1) Set an initial matrix  $U^0$  (often randomly, i.e., each  $\mu_{ik}$  is given a random value between 0.0 and 1.0)

(2) Calculate the  $c$  fuzzy cluster centres,  $v_i$ , for  $i=1,2, \dots, c$ :

$$v_i = \frac{\sum_{k=1}^n (\mu_{ik})^m x_k}{\sum_{k=1}^n (\mu_{ik})^m} \quad (4-7)$$

(3) Update  $U^l$ , for  $k=1,2,\dots, n$ ;  $i=1,2, \dots, c$ :

$$\mu_{ik} = \left( \sum_{j=1}^c (d_{ik}/d_{jk})^{2/(m-1)} \right)^{-1} \quad (4-8)$$

where  $d_{ik}$  is usually the Mahalanobis distance for remote sensing data.

(4) Compare  $U^{l+1}$  to  $U^l$ . If the difference between all corresponding pixels is less than or equal to a predetermined iteration convergence criterion, denoted by  $\epsilon$ , then stop, with  $U^{l+1}$  being the final fuzzy  $c$ -partition; otherwise, set  $U^l = U^{l+1}$  and return to (2).

Table 4.2 and Figure 4.2 in combination provide an illustration of the process of fuzzy  $c$ -means clustering, where three clusters are aimed for, and the value of parameter  $m$  is 2.0, using a simple example with 18 pixels in a two-dimensional multispectral space. Pixels are in the order from left to right and bottom to top and the values for FMV are in percentages.

Table 4.2 Initial and final U matrix (i.e., FMVs) for the pixels as shown in Figure 4.2

Pixels	Initial U matrix			Final U matrix		
	cluster 1	cluster 2	cluster 3	cluster 1	cluster 2	cluster 3
1	87	91	85	14	14	72
2	99	85	21	23	17	60
3	11	83	89	17	32	51
4	8	43	23	9	13	78
5	99	51	82	12	9	79
6	19	16	51	17	54	29
7	27	71	39	15	58	27
8	6	78	80	23	42	35
9	68	18	35	52	20	28
10	45	3	86	74	11	15
11	98	67	45	63	16	21
12	11	68	44	13	73	14
13	93	92	44	4	93	3
14	78	47	28	34	49	17
15	26	68	7	73	16	11
16	22	15	14	76	13	11
17	74	79	9	26	58	16
18	17	47	77	41	41	18

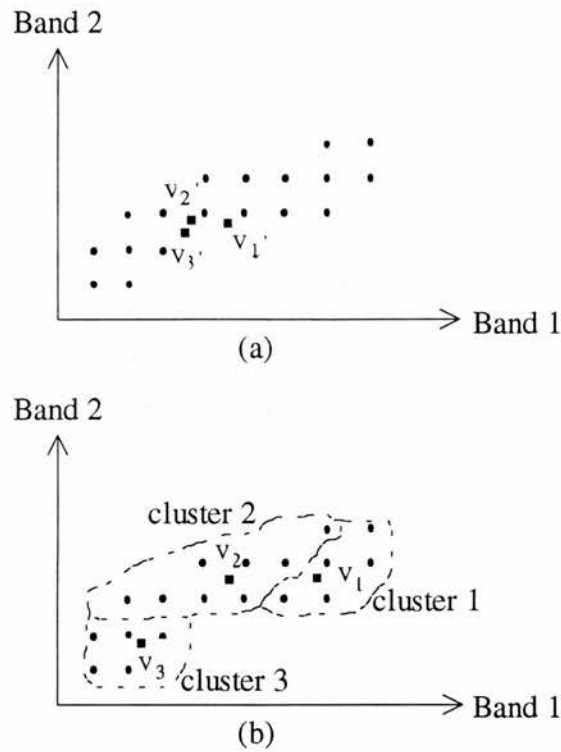


Figure 4.2 The process of a fuzzy 3-means clustering.

Initial fuzzy cluster centres  $v_1'$ ,  $v_2'$  and  $v_3'$  are indicated by solid squares shown in Figure 4.2 (a), while final fuzzy cluster centres  $v_1$ ,  $v_2$  and  $v_3$  are indicated by solid squares, with three clusters being formed by dashed boundaries, as shown in Figure 4.2 (b).

The initial U matrix (i.e.,  $U^0$ ) is generated randomly. In other words, each  $\mu_{ik}$  in the fuzzy set U is given a random value in the range from 0 to 100 %, implying that FMVs may not sum to 100 % for a pixel. In the final U matrix, however, FMVs sum to 100 % across all clusters, as shown in Table 4.2.

Accordingly, as shown in Figure 4.2, like a conventional clustering process in a unsupervised mode, the clusters in terms of their centres are not defined *a priori*, but rather formed iteratively from the process of optimisation of  $\mu_{ik}$ 's in relation to the

distance measure defined in equation (4-6).

A supervised approach may be taken if class means are known. In other words, if  $v_i$ 's are available by consulting training data, the fuzzy  $c$ -means clustering algorithm described previously becomes simply a one step calculation, by which the FMV for each pixel in each of the known classes can be derived from equation (4-8) straightforwardly. Thus, fuzzy  $c$ -means clustering can be applied in both an unsupervised and a supervised mode, making itself particularly attractive for fuzzy classification of remote sensing images.

Key, Maslanik and Barry (1989) applied the fuzzy  $c$ -means clustering method to Advanced Very High Resolution Radiometer (AVHRR) data for the purpose of classifying polar clouds and surfaces. They produced fuzzy classification results with optimal number of clusters after comparative tests. They also generated fuzzy sets obtained from supervised approach in a polar case study. Analysis of the fuzzy sets provides information on which spectral channels are best suited to the classification of particular features and can help determine areas of potential misclassification.

### **4.3 Generating fuzzy surfaces from photogrammetric data**

Following discussion of the methods for deriving fuzzy surfaces, which are relevant to digital image processing, as outlined in Table 4.1, attention is now given to the issues of generating fuzzy surfaces from categorical data derived from aerial photographs. This section refers, in particular, to photogrammetric data, but does not exclude any other types of visually interpreted categorical data such as those interpreted from remotely sensed images and any other photographic data (Table 2.2), as discussed further in the final section of this chapter.

### 4.3.1 Resolving proportions of component land cover types in mapping units

As described in chapter 2, land cover information may be derived from photogrammetry as well as remote sensing techniques. Often, spatial variabilities are suppressed, and are forced into predetermined mapping categories in traditional categorical map productions, as Campbell (1987) and Jensen (1983) described. For example, by means of analytical plotters, land cover information extracted from rigorously orientated stereo models reconstituted as aerial photographic pairs will end up in discrete polygons with the boundary fuzziness and interior heterogeneity filtered out (Kirby 1988).

The mapping accuracy is assessed by additional sampling, which is used to calculate some commonly adopted measures of classification accuracy such as Kappa coefficients. Such kinds of accuracy measures are often global indicators, and have limited uses for generating fuzzy surfaces.

As indicated in the introductory section, at a lower level component class proportions provide information on the uncertainty per polygon or any mapping unit. Figure 4.3 shows a regular polygon dominated by the class represented by rectangles (say, occupying 85% of the polygon) with inclusion of classes as represented by circles and triangles (say, occupying 8% and 7% of the polygon, respectively). One may use these component classes proportions as probabilities, with which each component class is expected to be found within the polygon (Foody 1995a).

It has been described above that a simple type of fuzzy surfaces may be derived from photogrammetric data by resolving proportions of component land cover in individual mapping units such as polygons. This method is based on the assumption that component land cover types within each mapping unit can be accurately delineated. However, it is more often the case that great fuzziness is present in the subject area,

implying that this method is flawed and that the fuzziness is only pushed to a lower level. Therefore, there seems to be merit in a back-to-the-reality approach, by which spatial variations are recovered and maintained through the whole mapping process, as developed in the next sub-section.

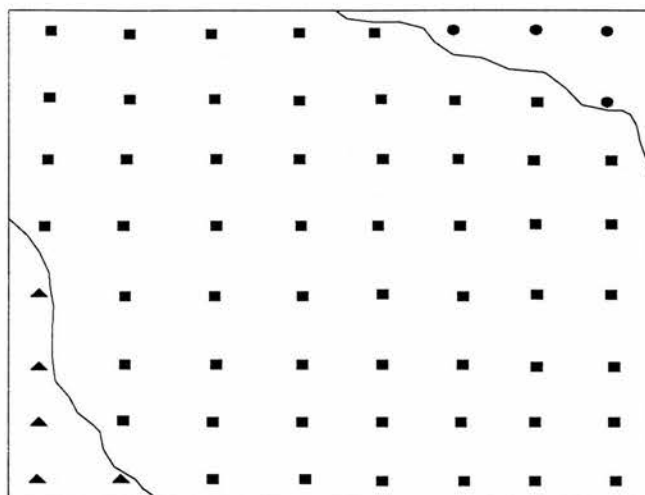
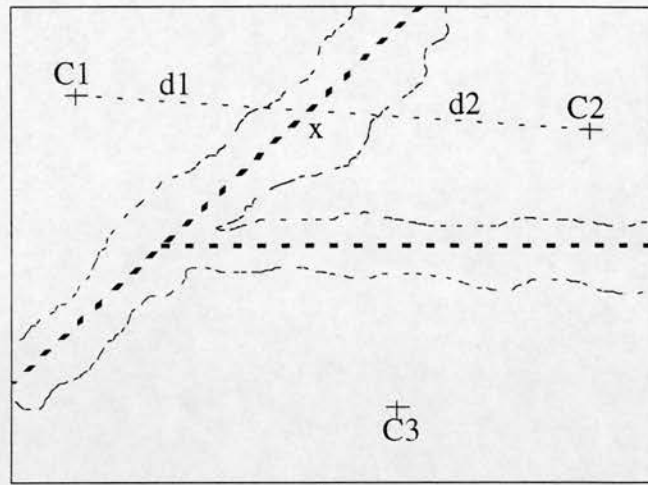


Figure 4.3 A mapping unit dominated by class denoted by rectangles, with other classes denoted by circles and triangles.

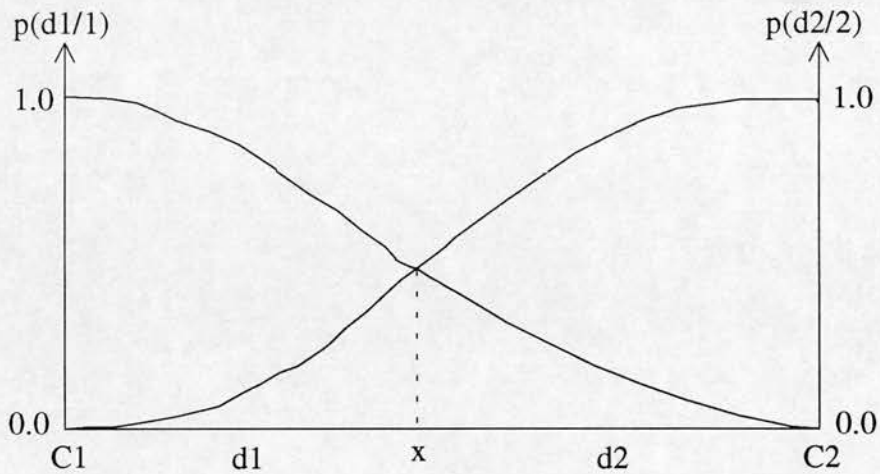
#### 4.3.2 Interpolation approaches

As seen in the previous sub-section, the probabilistic measures obtained are still spatially invariant within one map unit. In reality, the uncertainty should be larger if the surrounding samples or observations all belong to different classes, while this uncertainty would become smaller if the surrounding samples belong to the same class. One could decrease the uncertainty in classification by placing more samples around the location of prediction, as shown in Bierkens and Burrough (1993a and 1993b). Thus, a better method for generating fuzzy surfaces is by using interpolation. This is a two-step process: first sampling relatively well defined points and then inferring those unsampled and fuzzier points based on some distance functions. Figure 4.4 illustrates an interpolation process.





(a)



(b)

Figure 4.4 The process of interpolating fuzzy surfaces

As shown in Figure 4.4(a), one can label the centres of individual polygons representing class 1, 2 and 3, that is, C1, C2 and C3, with a probability of 100 % or 1.0 (virtually no uncertainty). At x (Figure 4.4 (a)) the probability will reflect the component probability of each adjacent polygon (as in Figure 4.4 (b)). Probability will decrease when moving towards the boundaries until it reaches 0 % or 0.0 in the centre of each adjacent polygon. Polygon boundaries are seen somewhere within the transitional zones indicated by dashed lines in Figure 4.4 (a). The changing pattern of class probabilities along a transect may be modelled by some function. This function may well be based on the distances between the points with known class probabilities

and the points whose probability with respect to the candidate classes are to be interpolated. Probability functions of finding 1 or 2 along the transect of C1 to C2 are shown in Figure 4.4(b), where  $p(d1/1)$  stands for the probability of finding class 1 at a distance of  $d1$  away from the centre C1, while  $p(d2/2)$  stands for the probability of finding class 2 at a distance of  $d2$  away from the centre C2. It is worth noting that, in the case of a relatively discrete class such as a lake, the probability function of class membership does not vary continuously but abruptly across boundaries with adjacent classes, and is reasonably stable within the polygon of the class (Foody 1996, pers. comm.).

However, there are problems with such a simple interpolation approach. Firstly, it is likely that different transects with different mixture and transition patterns may have to be modelled by different probability functions (Edwards and Lowell 1996; Wang and Hall 1996). This implies, in turn, that, despite its logical simplicity at first instance, an interpolation process may be difficult if it is to be based on properly defined and modelled probability functions. Moreover, there is lack of both theoretical and empirical evidence that interpolation methods based solely on distances can be suitably applied in situations where the attributes under study are of a categorical nature in the common sense, such as soil types or forest species. More satisfactory interpolation approaches exist within geostatistics.

Geostatistics grew from the need to interpolate certain properties (or attributes) of interest from sparse data, which are modelled as spatially correlated random variables. A random variable is a variable that can take a variety of outcome values according to some probability (frequency) distribution (Deutsch and Journel 1992). Apart from being location-dependent, a random variable is also information-dependent in the sense that its probability distribution changes as more data become available. Examples include atmospheric pressure and population densities.

According to Burrough (1987), the basic model in geostatistics is:

$$z(x) = u(x) + \delta(x) + \varepsilon, \quad (4-9)$$

where  $z(x)$  is the value of variable  $Z$  at point  $x$ ,  $u(x)$  is a deterministic function describing the general structural component of variation,  $\delta(x)$  describes the spatially correlated local random variation, and  $\varepsilon$  is the random (or noise) term.

Kriging, as a specific geostatistical method, assumes a constant local mean and a stationary variance of the differences between places separated by a given distance and direction. The semi-variance of difference (usually denoted by  $\gamma$ ) is half of the expected (E) squared difference between two values:

$$\gamma(h) = E[\{z(x)-z(x+h)\}^2] / 2 \quad (4-10)$$

where  $z(x)$  is the value of variable  $z$  at position  $x$  and  $z(x+h)$  is the value at position  $x+h$ ;  $h$  is a vector called the lag which describes a separation in both distance and direction between two positions. The function that relates  $\gamma$  to  $h$  is called the semivariogram. The semivariogram is the function which is used to quantify the spatial autocorrelation and to guide interpolation.

An experimental semivariogram can be estimated from sample data. The formula is:

$$\hat{\gamma}(h) = (1/2M(h)) \sum_{i=1}^{M(h)} \{z(x_i)-z(x_{i+h})\}^2, \quad (4-11)$$

where  $M(h)$  is the number of points of observations separated by lag  $h$ . By changing  $h$ , a set of values is obtained, from which the experimental semivariogram is constructed. The experimental semivariogram depends on the scale of survey, and relates to the size of sampling area and the support of the sample. The precision depends to some extent on the sample size.

An example of a semivariogram created from a data set published in Deutsch and Journel (1992) is shown as Figure 4.5, where the scatter plot is the sample data (representing experimental semivariogram), and the solid line is the fitted

semivariogram. As commonly understood, the experimental semivariogram may be subject to various errors. However, in principle, the true or theoretical semivariogram should be continuous. Thus, the experimental semivariogram is usually fitted by a mathematical model. In fact, Figure 4.5 shows a semivariogram model (the solid line) fitted mathematically from the experimental semivariogram (the scatter plot), which is expressed by a spherical model,  $0.04 + 0.23 (1.5 (h/25) - 0.5 (h/25)^3)$  (Deutsch and Journel 1992).

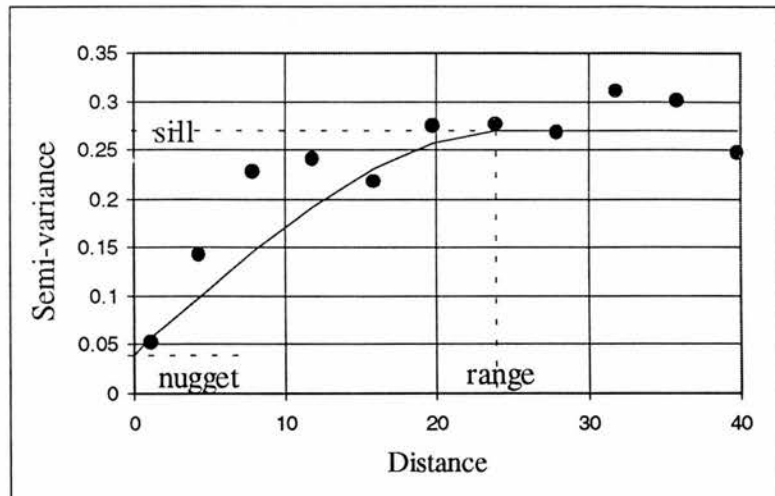


Figure 4.5 The features of an idealised semivariogram (unit of distance: miles).

There are several features to note in the semivariogram. At relatively short lag distances of  $h$ , the semi-variance is small but increases as the distance between the pairs of points increases. At a distance referred to as the range, the semi-variance reaches a relatively constant value referred to as the sill, as illustrated in Figure 4.5. This implies that beyond this range distance, the variation in  $z$  values is no longer spatially correlated. The experimental semivariogram often cuts the ordinate at some positive value, termed the nugget variance, as shown in Figure 4.5. This might arise from measurement errors, spatially dependent variation over distance much shorter than the smallest sampling interval or from spatially uncorrelated variation. One can check the existence and magnitude of the measurement errors by carrying out repeated measurements or sampling of the same variable at the same location. As

usual, this is done for several locations to reduce possible bias. Measurement errors affect the estimation of semivariograms. However, as measurement errors are assumed to be non-correlated and should be the same for all data points, they should not affect the overall form unless it is large relative to the total variance and the number of sample points is limited (Burrough 1995, pers. comm.).

Kriging is commonly used as a special kind of spatial interpolation method for numerical variables, and is often implemented as ordinary kriging. Indicator kriging, as a variant of kriging, estimates the conditional (*a posteriori*) probability distribution that a variable does not exceed a certain threshold (also known as a cutoff value), and is usefully explored in environmental studies (Cressie 1993). Indicator kriging performs the estimation of conditional probability distribution without making assumptions about the form of the prior distribution functions. This is an attractive feature for indicator kriging (Bierkens and Burrough 1993a). More importantly, when a categorical variable, such as land cover or soil type, is concerned, indicator kriging directly estimates the probabilities of finding individual classes at an unsampled location, given a set of classified samples.

Indicator kriging is therefore seen as being superior to other methods such as the distance-based weighting interpolation method outlined in Lowell (1994). This is because indicator kriging, in addition to its well-grounded theory, is guided by a semivariogram model that quantifies the spatial correlation intrinsic to the underlying variable, while simple distance-based weighting interpolation seems to be weakly grounded. Moreover, indicator kriging is worth exploring as an initial step in the conditional simulation in geostatistics, which is particularly attractive to researchers on GIS uncertainties (Englund 1993). The concluding chapter will elaborate on this.

Attention is now given to how the indicator kriging is performed. Suppose  $c$  mutually exclusive classes can be found over a field and  $n$  observations (e.g., classified samples)

are available, as shown in Figure 4.6, where solid rectangles represent observations. Each observation is classified as a member of one of the possible classes  $\{c_1, \dots, c_c\}$ . For each class  $c_i$  ( $i = 1, \dots, c$ ) under consideration, the  $n$  observations are transformed into binary data (i.e., an observation is transferred to 1 if it is classified as class  $c_i$ , 0 otherwise). These binary data are denoted as  $z_i(x_k)$ ,  $k=1, 2, \dots, n$ , for a particular class  $c_i$ .

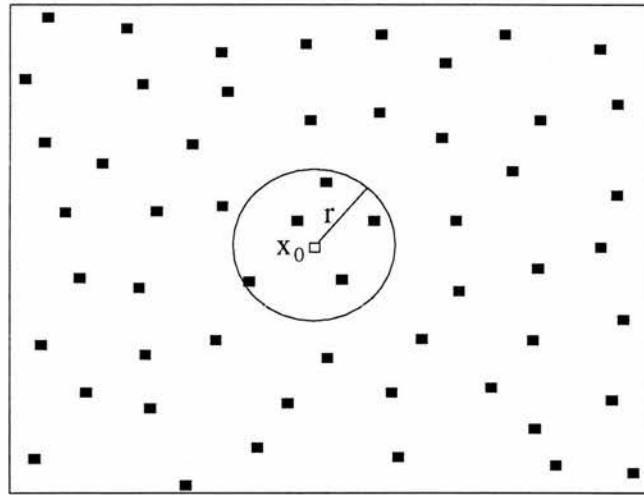


Figure 4.6 The process of estimating the occurrence of classes at a point  $x_0$  in the neighbourhood within the search radius  $r$ .

Suppose that an experimental semivariogram has been calculated by using equation (4-11), and has been subsequently fitted by a suitable model. Indicator kriging performs the estimation of probabilities of finding individual classes  $c_i$  ( $i = 1, \dots, c$ ) at a point  $x_0$  using:

$$\hat{z}_i(x_0) = \sum_{k=1}^n \lambda_k z_i(x_k), \quad (4-12)$$

where  $z_i(x_k)$  represents a binary variable at an observation  $x_k$  ( $k = 1, 2, \dots, n$ ) as described above,  $\lambda_k$  is the weight associated with the observation  $x_k$  (again,  $k = 1, 2, \dots, n$ ). These weights (i.e.,  $\lambda_k$  s) depend on the semi-variances between pairs of sampling points and the average semi-variances between the point  $x_0$  and a sampling point (Bierkens and Burrough 1993a). Clearly, indicator kriging is carried out by

using an ordinary kriging algorithm directly, when the variable under study is transformed into binary data.

When a similar process is applied to every cell on a pre-defined grid, a fuzzy surface is generated. As there are  $c$  classes, each class should be dealt with separately. In other words, for each class, a set of binary data is transformed from classified samples. A suitable semivariogram model is fitted on the corresponding experimental semivariogram, calculated using equation (4-11). In the end,  $c$  fuzzy surfaces are generated for  $c$  classes, on each of which the value at a point represents the probability of finding a given class at that point.

The probability scores as calculated above may not always sum to 1.0 across all candidate classes for each point. In accordance with the definition for FMVs, the probability scores need to be normalised by the actual sums, similar to the transform as expressed by equation (4-3), in case of non-unit sums.

It is worth mentioning that the classified samples used in indicator kriging should aim to be observations or interpretations without significant fuzziness. This may best be achieved by locating samples in the centres of certain land cover patches, which will be far away from the boundaries and so avoiding the transition or mixture zones between homogeneous land cover patches in the spatial domain. The uncertainty may be high near all types of boundaries, while it should be low within inner parts of individual polygons for dominant land cover types.

#### **4.4 Analysing fuzzy surfaces**

Once fuzzy surfaces have been derived, by using the methods discussed previously, it is possible to perform a variety of analyses to exploit the potential held by the wealth of information maintained in these fuzzy surfaces.

The previous two sections have described the derivation of fuzzy surfaces from remote sensing images and photogrammetric data separately. However, it is often the case that the two different types of fuzzy surfaces are analysed in combination, as developed below.

Suppose that two types of fuzzy surfaces are derived for the same area, and further assume that fuzzy surfaces generated from photogrammetric data are to be used as reference data layers. Thus, denote vectors:

$$\begin{aligned}
 U(x) &= (\mu_1(x), \mu_2(x), \dots, \mu_c(x)), \text{ and} \\
 P(x) &= (p_1(x), p_2(x), \dots, p_c(x)) \qquad (4-13)
 \end{aligned}$$

where  $\mu_i(x)$  and  $p_i(x)$  ( $i = 1, 2, \dots, c$ ) are the FMVs of a pixel  $x$  belonging to class  $i$ , as derived by the methods described in Sections 4.2 and 4.3, respectively.

The degree of fuzziness for a particular set of fuzzy surfaces may be measured by using entropy. Entropy is a measure of uncertainty and information formulated in terms of probability theory (Foody 1995a). Measures of entropy express the way in which the probability of class membership is partitioned between the classes. It is based on the assumption that in an accurate classification each pixel will have a high probability of membership in only one class. Specifically, entropy is maximised in the situation when the probability of class membership is partitioned evenly between all defined classes and minimised when it is associated entirely with one class. So large values indicate low accuracy in classification, while small values indicate high accuracy in classification. This can be more easily understood by referring to Table 4.2, where a list of FMVs is provided. When two or more alternative classes have non-zero probabilities associated with them then each probability is in conflict with the others. This is because the probabilities must sum to 1.0 or 100 %, and a gain of probabilities for one or more classes must involve a loss of probabilities for one or more of the other alternative classes. The expected value of conflict is given by



entropy H measured using equation (4-14) shown below,

$$H = -\sum_{i=1}^c \mu_i(x) \log_2 \mu_i(x) \quad (4-14)$$

where  $\mu_i(x)$  is the FMV of pixel x belonging to class i, where the index i ranges from 1 to c ( the total number of classes). A measure of entropy for probabilities, that is,  $p_i(x)$ , can also be calculated.

The interpretation of measures of entropy on an individual pixel or point basis is, however, not that straightforward. In situations where both classified data and reference data are fuzzy, entropy measures will not be suitable. For example, when reference data are fuzzy, any entropy could be associated with an accurate representation; the interpretation of entropy values is therefore difficult. Foody (1995a) suggested that cross-entropy can be used to illustrate how closely a fuzzy classification represents the geographical reality when multiple and partial class membership is a feature of the reference data as well as the classified data: the smaller the measure of cross-entropy, the closer the classified data are to the reference data. Cross-entropy  $H_c$  is measured using equation (4-15) shown below,

$$H_c = -\sum_{i=1}^c \mu_i(x) \log_2 p_i(x) + \sum_{i=1}^c \mu_i(x) \log_2 \mu_i(x) \quad (4-15)$$

where  $\mu_i(x)$  is the FMV of a pixel x belonging to class i, and  $p_i(x)$  is the probability of finding class i at pixel x as defined on the reference data layer. The index i again ranges from 1 to c.

For assessing the closeness of a fuzzy classification with a fuzzy reference data set, say based on vectors  $U(x)$  and  $P(x)$ , it is possible to compute the sum of all pixels  $\mu_i(x) * p_i(x)$  as a measure of the degree of agreement between the two uncertainty surfaces with respect to class i, when  $\mu_i(x)$  and  $p_i(x)$  are normalised. This could be called a vector dot product. Based on this, one would wish to go further to calculate the average correlation of coefficients over all pixels by dividing the vector dot

product by the number of pixels in a field. One could also sum  $\mu_i(x) * p_i(x)$  over  $i$  ( $i = 1, \dots, c$ ) as a measure of the agreement between the two vectors  $U(x)$  and  $P(x)$  for a pixel  $x$  (Foody 1995b; Goodchild 1996, pers. comm.).

The correlation analysis is especially meaningful when the probability scores (i.e., the values of  $P(x)$ ) are component land cover proportions, because FMVs as relevant to pixels can then be used to resolve the sub-pixel land cover information, if strong correlations are found between FMVs and probability scores (Foody and Cox 1994). Such a correlation analysis also provides a practical interpretation for FMVs, which, originally attached to spectral domains, would otherwise be difficult to relate to spatial domains.

To generate a conventional maximum likelihood classification, vectors  $U(x)$  and  $P(x)$  are subjected to a “maximisation” process, by which  $x$  and  $y$  are labelled as the classes having the maximum values. For example, pixel  $x$  is to be classified into class  $j$  on the condition as expressed in equation (4-16) below:

$$\mu_j(x) = \text{maximum} (\mu_1(x), \mu_2(x), \dots, \mu_c(x)),$$
$$\text{for } j = 1, 2, \dots, c \quad (4-16)$$

where the maximum FMV, i.e.  $u_j$ , of pixel  $x$  can later be used to compose a companion band of image to depict the spatial variability of confidence a user can place in the classified data layer. The classified data layer consists in this case of class labels  $j$ 's satisfying equation (4-16) at pixel level.

In order to illustrate such a process, the data from Table 4.2 in Section 4.2 are used to derive Table 4.3 below, where both labelled classes and their probabilities of being correctly classified are listed.

As in a conventional hard classification, the maximisation will produce categorical maps comprised of contiguous patches of labelled pixels. These categorical maps can

be assessed with respect to their classification accuracies as measured by overall percentages of agreement or Kappa coefficients using the procedures described in Section 3.3. To do this requires an independent source of reference data.

Table 4.3 Labelled pixels and their probabilities (based on Table 4.2 in Section 4.2)

Pixels	Labels	Probabilities (%)
1	3	72
2	3	60
3	3	51
4	3	78
5	3	79
6	2	54
7	2	58
8	2	41
9	1	52
10	1	74
11	1	63
12	2	73
13	2	93
14	2	49
15	1	73
16	1	76
17	2	58
18	1	41

Table 4.4 Pure pixels selected by slicing at a threshold of 60 %

(based on Table 4.3)

Pixels	Labels
1	3
2	3
4	3
5	3
10	1
11	1
12	2
13	2
15	1
16	1

Sometimes, it is necessary to assess the accuracy of a particular classification, which

requires “pure” pixels to be identified and selected in order to provide reference samples. This selection process can be done via a “slicing” operation, by which the maximum FMVs (denoted by  $\mu_{\max}$  and  $p_{\max}$ ) from vector  $F(x)$  and  $U(x)$  are examined with reference to a pre-determined threshold  $\tau$ . Specifically, this processing is performed such that a pixel  $x$  is selected if the value of  $\mu_{\max}$  is not less than value  $\tau$ . As an example, the data from Table 4.3 are sliced to produce selected pure pixels with FMVs not less than 60% (Table 4.4).

Such processing is particularly useful in situations where the mapped phenomena are inherently fuzzy or exist as mixtures, in particular for remote sensing data at coarse spatial resolutions. This is because the identification of representative classes can be performed on a quantitative basis, which would be highly subjective when no information on class membership was available (Lowell 1994).

More importantly, a slicing operation will generate a special kind of classified map. On this kind of map, selected pixels or points belong to their nominal classes with probabilities of not less than the threshold  $\tau$ . Those pixels or points left unclassified constitute uncertainty zones, which can be used to derive epsilon band widths.

To illustrate such an operation, the example based on Figure 4.4 is developed. The transect between C1 and C2 is sliced at the threshold  $t$ , leading to an epsilon band of width  $w$  (Figure 4.7).

As shown in Figure 4.7, the threshold can be set between the minimum probability indicated by “min” and the maximum probability 1.0. Increasing the threshold will widen the epsilon band width, and vice versa. Thus, a variety of epsilon band models can be derived, based on different thresholds.

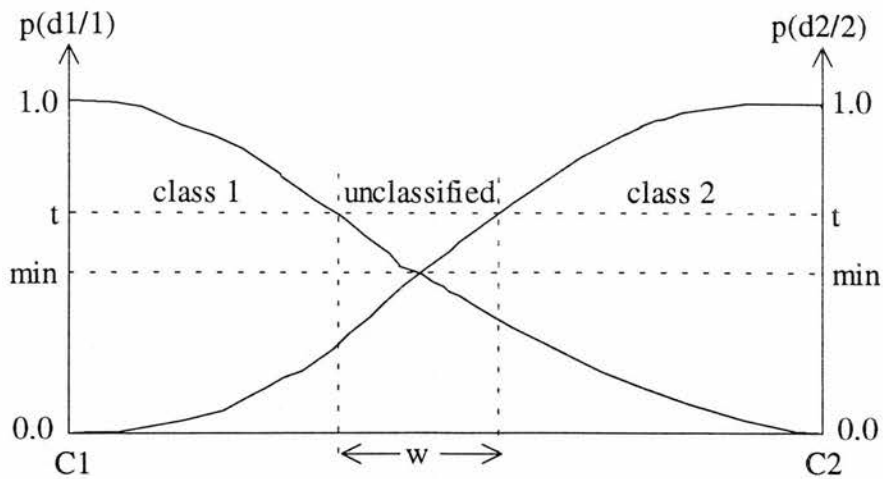


Figure 4.7 The process of analysing fuzzy surfaces by slicing along the transect

Furthermore,  $U(x)$  and  $P(x)$  can be scrutinised by relative magnitudes of their elements. For example,  $\mu_i$ 's ( $i = 1, 2, \dots, c$ ) of a pixel  $x$  can be arranged in descending order, that is, actually sorting out the most likely down to the least likely classes for each pixel. Firstly the sorted classes are denoted by a vector  $o(x) = (c_1, c_2, \dots, c_c)$ . Similar processing can be applied to vector  $P(y)$ . Then the resulting sorted sequence for  $P(y)$  is denoted by vector  $O(y) = (C_1, C_2, \dots, C_c)$ . A "soft" comparison is carried out by comparing between the most likely, the second most likely, down to the least likely classes as indicated by the vectors  $o$  and  $O$ , that is, by examining if  $c_a = C_b$ ,  $a, b = 1, 2, \dots, \tau$ , where  $\tau$  specifies the tolerance set for a soft comparison ( $\tau \leq c$ ). The greater the tolerance, the more likely a match is obtained, and hence the larger the error. When  $\tau$  is set 1 and if  $c_1 = C_1$ , it is said that a perfect match is achieved because the most likely classes are found to be the same for the test data and the ground data. This is what happens when an agreement is reached during a hard classification. When the tolerance is set to the second most likely class, a soft comparison is performed based on a set of conditions such as  $\{c_1 = C_1 \text{ or } c_1 = C_2 \text{ or } c_2 = C_1 \text{ or } c_2 = C_2\}$ .

## 4.5 Discussion

Some of the methods for deriving and analysing fuzzy surfaces have now been described, by which fuzzy surfaces, as special kinds of representation of uncertainties in spatial databases, can be explored. This chapter looks forward to the case study, which endeavours to demonstrate the advantages held by surface-based methods, following a further section discussing the methods established previously in this chapter.

Based on Table 4.1 which lists the two types of fuzzy surfaces and their derivations, this chapter has discussed how different fuzzy surfaces are produced. It was seen that fuzzy surfaces can be derived from remote sensing multispectral data by defining a suitable fuzzy membership function in a supervised mode or by using the fuzzy *c*-means clustering in an unsupervised mode. It was found that the fuzzy *c*-means clustering is particularly versatile with both unsupervised and supervised alternated as required, making itself an attractive technique for fuzzy classification.

For visually interpreted photogrammetric data, a few different methods for generating fuzzy surfaces have also been discussed. Though sub-pixel component land cover proportions may be used as probabilities to construct corresponding fuzzy surfaces in a relatively easy way, it may not be suitable in situations where significant fuzziness exists. Indicator kriging, on the other hand, estimates directly the *a posteriori* probability of finding a category at a location, given a set of classified samples, without making any assumptions about any probabilistic distributions. This property is very beneficial for the mapping of fuzzy surfaces in place of otherwise categorical land cover data.

As seen in this chapter, the process of a fuzzy classification based on remote sensing digital images seems to be quite automated, and relatively easy to implement, because

the digital image data are already in grid form. On the other hand, as photogrammetric data are often stored in polygonal formats, where no mechanism is originally provided to maintain the spatial variations present within each polygon, the derivation of fuzzy surfaces from the photogrammetric data tends to be more difficult than that from remote sensing digital images.

The analysis techniques presented so far serve to demonstrate how fuller uses of the derived fuzzy surfaces can be made. These techniques can expose the spatially varying uncertainty levels as a final “hardened” product, identify pixels or points relatively accurately classified, enable the generation of epsilon bands, and make comparisons on a relaxed condition. More importantly, surface-based methods open many possibilities to quantitative analysis of uncertainties in spatial databases. To make the arguments more convincing, their uses will need to be demonstrated in a case study.

Differences between fuzzy and probabilistic measures used to construct fuzzy surfaces might better be clarified at this stage. Fisher (1994) was one of the relatively few researchers who has recognised the importance of differentiating between fuzzy and probabilistic approaches when dealing with uncertainties in GISs, after admitting that he himself had previously made confused use of fuzzy and probabilistic concepts. Though both are measured in the range of 0.0 to 1.0 (or 0 to 100 percent), fuzzy and probabilistic concepts are fundamentally different in theoretical and practical terms. According to Fisher (1994), probabilistic approaches assume that there is a Boolean phenomenon (existing or not), and the probability measures the accuracy with which the Boolean event occurs. Fuzzy surfaces generated using indicator kriging contain probabilistic measures, given a set of classified samples, as shown in Section 4.3.

Fuzzy set theory, on the other hand, has been developed mainly to deal with imprecisely located phenomena such as land cover and soil type, and to cope with spatial decisions based on multiple criteria such as land suitability assessment, where

crisp set theory may not be suitable (Kaufmann 1975; Zadeh 1965). In this case, the fuzzy concept is used to determine the degree to which an object is a member of a set such as a class (Fisher 1994). While crisp set theory permits only a zero or a full membership of an object belonging to a class, fuzzy set theory allows for partial and multiple memberships of an object belonging to all the candidate classes. It can thus be said that fuzzy set theory is the general case, whereas crisp set theory is a sub-set of this.

Fuzzy set theory uses certainty factors (ranging from 0.0 to 1.0), which may be probabilistic, but more often are based on educated guess, as mentioned in Chapter 3. In other words, fuzzy measures such as FMVs are usually not probabilistic in distribution, but convenient and continuous transforms (Heuvelink and Burrough 1993). As shown in Section 4.2, many FMVs are virtually conversions of probabilities defined on certain probability distributions to class memberships normalised to the range between 0.0 and 1.0, except for the *a posteriori* probabilities used as FMVs as shown in equation (4-4) (Wood and Foody 1993). Similar situations occur in the normalisation of probability scores estimated by using indicator kriging, as mentioned in Section 4.3.

Despite its relaxed mathematical basis, fuzzy set theory allows many otherwise non-quantifiable phenomena to be dealt with quantitatively. Useful examples can be found in Altman (1994), Gopal and Woodcock (1994), and Robinson (1988). Besides, given uncertainties surrounding the measures of uncertainties, it seems more sensible to use fuzzy set theory than rigorous statistical and probabilistic methods (Drummond and Ramlal 1992). Many people use fuzzy methods because they seem to make intuitive sense. Therefore, “fuzzy” may be more relevant than “probabilistic” as a term to be used in the study of uncertainties. But, in order to build anything further, it is necessary to use the concept of probability (Goodchild 1996, pers. comm.).



Finally, a few words on the surface-based approaches as a whole. Clearly, without convincing examples showing the advantages or potentials held by surface-based methods, GIS communities may hesitate to commit themselves or invest in mapping the uncertainties of spatial data, because of their preoccupation with the spatial data themselves. Unless it can be demonstrated that working with uncertainties of spatial data is not a luxury but a necessity for proper and fuller applications of GISs, fuzzy surfaces will remain something on blueprint only. This will be shown in the case study, where the most attractive aspects of surface-based models and methods are established as the simplified uncertainty analysis functionality they can offer, in particular, for quantitative analysis and visualisation of spatially varying uncertainties.

## Chapter 5

### Empirical study I: data and uncertainties in objects

#### 5.1 Introduction

As suggested towards the end of the previous chapter, a case study will be helpful and necessary to support the advocated methods. This chapter will discuss how the study area was chosen, and then describe the data sources accessed and the techniques and processing involved in the data acquisition. A second purpose of this chapter is to illustrate how uncertainties are estimated and described in an object-based approach, thus giving hints to the surface-based approach promoted in this thesis.

#### 5.2 Study area

Because the work seeks to understand the identification and combination of uncertainties in spatial data more fully, it is important, firstly, that the study area be geared to providing data diversity, so that the empirical results will be as representative as possible of data typically used in urban-orientated GISs and specifically for land use and land cover mapping in urban areas. For this reason, selection of test sites must consider the information richness of the phenomena under study.

Secondly, concerning uncertainties in spatial databases, it is important to ensure that the ground data are both adequate and readily available in order to allow for various data validity tests. As discussed in Chapter 2, spatial data are usually not cheap in the first place. It follows that spatial data of extra high accuracy, which are required to derive uncertainty measures such as a RMSE for DEMs or Kappa coefficients for

remote sensing classification data, would, therefore, not be obtained simply. Thus, selection of test sites should take into account the availability and accessibility of data sources and convenience of field work in addition to the data diversity as mentioned in the previous paragraph.

Furthermore, the approach proposed is better tested empirically with diverse spatial data of varying accuracies. A suburban area would not only be geared to enabling cost-effective data acquisition but would also cater for the construction of test data with different accuracies. This is because suburban area applications involve a range of mid-scale data products, usually incorporating base map data such as buildings, contours, street networks, land records on one hand, and environmental data such as soil, water and noise level on the other hand (Aronoff 1989). Therefore, there is a need in suburban applications to combine data sets with quite different positional and attribute accuracies. It is for this outstanding reason that a suburban area was chosen.

A local area was chosen for convenience, covering about 1.5 square kilometres, located within the city of Edinburgh, centred on Blackford Hill, where the Royal Observatory is situated, as shown in the marked rectangle on the colour aerial photography, reprinted as Figure 5.1.

Orientating oneself to the area within which the study takes place as shown in Figure 5.1, the north part of the photograph is the old town below Edinburgh Castle. Immediately north of the oval-shaped Meadows park are the central buildings of The University of Edinburgh, including the Royal Infirmary. Holyrood Park in which Arthur's Seat is located is at the north east corner of Figure 5.1. A main road runs NNW-SSE from the central buildings of the University to the King's Buildings, a modern extension to the university, which is flanked by the golf-course on the lower eastern end of Blackford Hill.



Figure 5.1 The study area as marked by the rectangle (this reprint covers the southern parts of the City of Edinburgh)

Focusing on the study area, there is a variety of urban thematic and topographic features: a wooded valley, residential, commercial and academic buildings, road networks and footpaths, recreational areas, a small lake, agricultural fields and worked allotments, hills and flat ground.

As seen on the aerial photography in Figure 5.1, the residential districts are set densely together. The roads, the pavements, the roofs, the walls and the hedges exist in complex spatial arrangements. Manual digitising on a reconstituted stereo model is bound to be a demanding job in these kinds of areas, yet these are very typical of the fringe of densely urbanised districts.

Some of the difficulties for the interpreter of suburban photography arise from the indistinct nature of boundaries between land cover types. For example, on Blackford Hill, the dispersed individual trees or groups of trees blend into adjacent land cover types. Shrubs and trees dominate the western end of Blackford Hill, and continue down into the valley, the Hermitage of Braid, where they mix with grazing land. Shrubs and grassland cover the Braid Hills in the south. Urban amenity trees and large gardens with lawns are a feature of the suburban residential districts.

In summary, the fabric of urban and suburban areas is often highly varied and compressed, where human beings have made intensive use of every possible space. The same circumstance was described by Campbell (1987) in the context of remote sensing of urban land uses and land covers.

### **5.3 Data sources**

The consideration of a varied and compressed urban fabric leads to the requirement that spatial data in urban applications should be recorded at large scale, to permit more accuracy. For example, land surveying based on a total station may sensibly be

used for the recording of detail. Indeed, for well-defined features, accurate positions are obtained by land surveying, by precisely measuring the distances and angles with reference to certain known points. An accuracy of centimetres is usually achievable, allowing for detail mapping at large scales, say 1:1,250 and 1:2,500 scales. Recent developments in GPS have greatly facilitated efficient and reliable spatial data acquisition, and are attractive for urban applications.

However, as also described in Chapter 2, using land surveying for detail mapping in urban areas is common only for relatively small areas of say 1 ~ 2 km<sup>2</sup>. In other words, detail mapping by land surveying becomes impractical over large areas, where photogrammetry and remote sensing are designed to work efficiently. Moreover, as a result of continuous and dynamic processes, urban mapping and change monitoring necessitate currency and timeliness in data acquisition and analysis, which are best met by established photogrammetric and remote sensing techniques.

Using photogrammetric techniques, for topographic mapping generally and for thematic mapping of dynamic land use and land cover in urban areas, aerial photographs at large and medium scales are normally used. For example, Lo (1971) used aerial photographs (1:10,000 scale) for a typological classification of buildings in the city centre of Glasgow. In his study, five discriminating variables were employed in the cluster analysis and discriminant function analysis: building height, building area, roof type, presence and absence of street parapet wall and degree of excrescence. Using medium scale aerial photographs (approximately 1:25,000 scale), Gautam (1976) undertook an urban land use mapping project based in Bikaner city, India, where combined use of aerial photo interpretation and ground-based surveys was necessary. In addition to low and medium altitude aerial photographs, high altitude aerial photographs are also useful media to provide spatial data at low spatial resolutions for urban applications, as described in Jensen (1983). As mentioned in Chapter 2, colour infrared aerial photographs are particularly attractive for enhanced

distinction between vegetation, water and man-made structures. Baker, DeSteiguer, Grant and Newton (1979) employed colour infrared aerial photographs at a range of scales (1:6,000, 1:12,000 and 1:30,000) for their project on land use and land cover mapping. They proposed a set of useful practical procedures for land use and land cover mapping in urban areas.

Though not as commonly used as aerial photographs, remotely sensed data acquired on airborne (such as ATM data) and spaceborne platforms (such as satellite data especially those with higher spatial resolutions, like SPOT HRV and Landsat TM imagery) are increasingly utilised in urban applications to provide valuable and cheap data with respect to the physical and cultural properties of urban areas. An example for the use of airborne remotely sensed data is provided by Jensen, Cowen and Halls et al. (1994), who used calibrated airborne multispectral scanner (CAMS) data with 5 by 5 metres spatial resolution to inventory and predict residential and commercial land-use development for capital investment forecasting based in BellSouth, South Carolina, the United States. For satellite remote sensing data, an early example by Carter and Gardner (1977) described the extraction of urban growth information in the UK from Landsat data for urban planners, although with unacceptably low quality results. A recent test was carried out by Langford and Unwin (1994), who made successful use of Landsat TM data for mapping the population density based on identified built-up areas in England.

It is, however, part of this thesis to re-evaluate the value of photogrammetric and remote sensing techniques. As was also explained in Chapter 2, urban phenomena often exist in mixed associations, the individual member of which may have to rely on extensive ground-based surveys and inventories for their accurate identification and measurement (Jensen 1983). Thus, it is usually the combined use of land surveying, photogrammetry and remote sensing that features in successful urban and suburban land use and land cover mapping projects.

As discussed in Chapter 2, depending on the requirement of a specific project, published maps may be useful as secondary data sources (Thapa and Burtch 1991). In general, topographic and thematic maps are particularly relevant for urban studies such as land use planning. A more specific use of published maps is for georeferencing satellite images which usually cover far greater areas than aerial photographs. The great coverages of satellite images imply that extensive land surveying would be costly, especially when it is recognised that satellite images have relatively coarse pixel sizes. Therefore, maps at large scales can sensibly be used to identify and measure certain well-defined points as GCPs for georeferencing satellite images.

For the purpose of this study, data sources including OS large scale plans, aerial photography at large and medium scales, and remotely sensed images were acquired, as shown in Table 5.1, while their spatial coverages in relation to the study area are shown in Figure 5.2, where the study area is shown by the shaded area and coverages are approximately presented.

Table 5.1 The data sources used to build the test data

Data sources	Scales/resolutions	Numbers	Dates
OS plans	1:1,250	NT2468 - NT2773	published 1966, 1968, 1971-4, 1978, 1980-1, 1986-7, 1989
OS plans	1:2,500	NT2570 - NT2571	published 1967
Aerial photographs	1:5,000	sortie 41 090 print 423-8, 458-63	flown 24th July 1990
Aerial photographs	1:24,000	sortie 51888016-7 print 6447-8	flown 14-18th June 1988
Satellite image	SPOT HRV (10 m)*	no information	imaged July 1985
Satellite image	Landsat TM (30 m)	path 204 row 21	imaged 14th May 1988

\* The spatial resolution of the SPOT simulation data used was enhanced using panchromatic data, so that the resulting spatial resolution is better than 20 m, but not necessarily as good as 10 m.



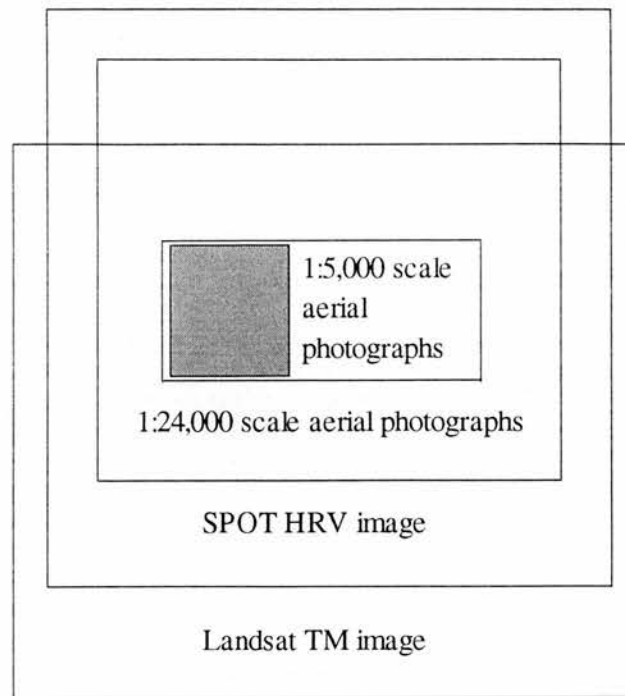


Figure 5.2 The spatial coverages of aerial photographs and satellite images used in the case study (not to scale).

The OS large scale plans were published at various dates between 1960s and 1980s. They were used to provide certain planimetric control points for geometric rectification of satellite images incorporated in the case study.

The 1:5,000 scale aerial photographs are in natural colour and are part of an experimental sortie of high resolution material flown for the Ordnance Survey. They were flown in late July 1990, using a Zeiss 630 FMC camera with a focal length of 304.77 mm. Two strips of photographs, each of five stereo-models, cover the area of interest. The aerial photographs can be used to generate reference data, against which data extracted from 1:24,000 scale aerial photographs might be tested.

The 1:24,000 scale aerial photographs (in natural colour) were flown in mid June 1988, as a part of the Scottish national aerial photographic initiative (Kirby 1992). An

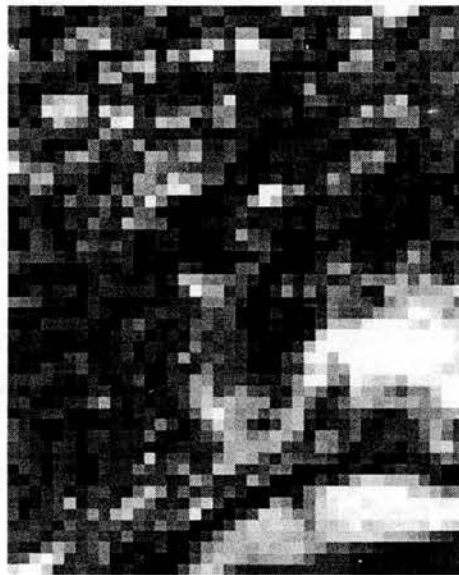
example has been shown in Figure 5.1. As seen in Figure 5.1, individual buildings, roads, hedges, footpaths and a small lake can be clearly identified and located. Existence of patches of grassland, shrub and wooded land is reasonably well depicted, though their accurate delineation is not as easy as in the case of many man-made structures. This set of aerial photographs at medium scale was used as the basis to carry out photogrammetric digitising, which provided reference data to check satellite image data.

Sub-scenes of SPOT HRV and Landsat TM data were accessed from the Department of Geography collection, and are shown as (a) and (b) of Figure 5.3, respectively. Figure 5.3 (a) is a black-and-white copy based on SPOT HRV XS bands 1,2 and 3, while Figure 5.3 (b) is a black-and-white copy based on Landsat TM bands 3, 4 and 5. The SPOT HRV image appears reasonably clear for its spatial resolution of 10 metres: contrast between vegetation, water and built-up areas is outstanding; differentiation between residential and non-residential districts is possible, as shown in (a) of Figure 5.3. For the Landsat TM image, as shown in (b) of Figure 5.3, on the other hand, distinction among land cover types become blurred, implying a classification based on such a coarse resolution image tends to be more difficult than that based on the SPOT HRV image.

Because the 1:24,000 scale aerial photographs and the Landsat TM data were acquired at about the same time in May/June 1988, it can be assumed that there are no significant differences between them in the representations of either permanent artifacts or vegetational changes. The SPOT HRV image was acquired in July 1985. Field checking re-confirmed that there had been no changes in the land use and land cover in the study area between 1985 and 1990 except for a few newly completed structures. These agreements in dates and field checks suggest that both SPOT HRV and Landsat TM data can be checked by referring to the 1:24,000 scale aerial photographs.



(a)



(b)

Figure 5.3 Sub-scenes of remotely sensed images used in this case study (scale 1:20,000 approximately): (a) SPOT HRV data, and (b) Landsat TM data. The same location is shown but with different boundaries.

#### 5.4 Methods used to acquire test data in required form

By incorporating data sources of different origins, scales and resolutions, it became possible to build a hierarchy of test data with a corresponding hierarchy of accuracies, as shown in Table 5.2. The table is followed by paragraphs explaining how the hierarchy was developed.

Table 5.2 The hierarchy of test data

Hierarchy of test data	Specifications
1. ground control points (GCPs)	a. photocontrol data combining National Grid points and field surveyed control points b. check data comprising GCPs extended by using photogrammetric block adjustment based on the 1:5,000 scale aerial photographs c. planimetric control points digitised from OS large scale plans
2. photogrammetric data digitised from the 1:24,000 scale aerial photographs	d. based on the photocontrol data and checked against the check data above
3. classified SPOT HRV and Landsat TM data	e. rectified on the planimetric control points digitised from OS large scale plans and checked against the photogrammetric data.

As is well known, apart from the most rudimentary or reconnaissance surveys, the starting point for a land cover mapping project is usually the collection of adequate ground control points (GCPs) and ground truth data, which may be supplied by land surveying and visits. Land cover mapping is much more useful and valuable in a LIS if it is geo-referenced, when it can be one of many layers in an urban-orientated GIS.

For cadastral purposes, for ownership, and for taxation, a high precision is often required, and this is enabled by properly setting up ground control data combining GCPs and ground truth data.

The global reference systems can be the geographical coordinate system (as latitude and longitude  $\phi$ ,  $\lambda$ ) or in accordance with map grid systems (as Easting and Northing E, N). In the UK, the Ordnance Survey National Grid is the standard coordinate system used by most LIS and urban GIS projects and is the easiest for data conversion.

Clearly, the set of ground control data shown in Table 5.2 is a prerequisite not only for statistically based analysis of uncertainties, but also for routine photocontrol (aerial photography) and is needed to create geometrically corrected land cover maps from remote sensing images. In other words, ground control data serve two purposes: (1) experimental need for higher-order checking, and (2) routine use for photocontrol and geometric corrections.

The National Grid coordinates of two second-order pillar triangulation stations were supplied free by OS. These two locations, at Arthur's Seat and Blackford Hill, provide the basis for the land surveying, which in turn generated ground control data: one set for the photogrammetric block adjustment based on the 1:5,000 scale aerial photographs and the other set for photogrammetric digitising based on the 1:24,000 scale aerial photographs, as will be described below.

Because of the cost and time that would have been necessary to obtain a dense ground control network of higher order by ground surveying alone, cost-effective photogrammetric block adjustment was used to densify the control network. For this purpose, aerial photographs at the scale of 1:5,000 were used, as described in the previous section. The densified control data were then used to check the

photogrammetric processing of aerial photographs at 1:24,000 scale, and are included as check data shown in Table 5.2.

Photogrammetric digitising was carried out on the 1:24,000 scale aerial photographs, which were absolutely orientated by part of the ground surveyed control data, and checked by the densified control data mentioned above. As this study was set in the context of suburban land cover mapping, which was based on photogrammetric and remote sensing techniques, it is necessary to introduce a land cover classification system appropriate for photogrammetric and remote sensing data. Considering the compatibility of photogrammetric and remote sensing data, the USGS land use and land cover classification system for use with remote sensing data (Campbell 1987) was used with the following classes appropriate to the scene:

- (1) grass (park and grassland),
- (2) built (built-up and barren land),
- (3) wood (woodland, no distinction made between deciduous and coniferous woodland),
- (4) shrub (shrubland, including open wooded land), and
- (5) water (water bodies such as lakes; (Scottish lochs)).

The five classes are all of areal coverage; there are no classes of linear features such as hedges and small roads. Major roads are included in the built-up classes.

Because each remote sensing image has far greater spatial coverage than the 1:24,000 scale aerial photographs, SPOT HRV and Landsat TM images required extra planimetric points to control the additional ground area. These points were digitised from OS large scale plans. Similar circumstances may also be found described by Janssen and van der Wel (1994). Subsequently, the 1:24,000 scale aerial photographs can be used as reference data in the evaluation of the classification of remote sensing images.

In overview, the scheme for constructing test data is illustrated in Figure 5.4. The next sub-sections will describe in more detail the test data acquired and the techniques applied to derive these data. The techniques include land surveying, control densification by using photogrammetric block adjustment, map digitising, photogrammetric digitising and classification of remote sensing images.

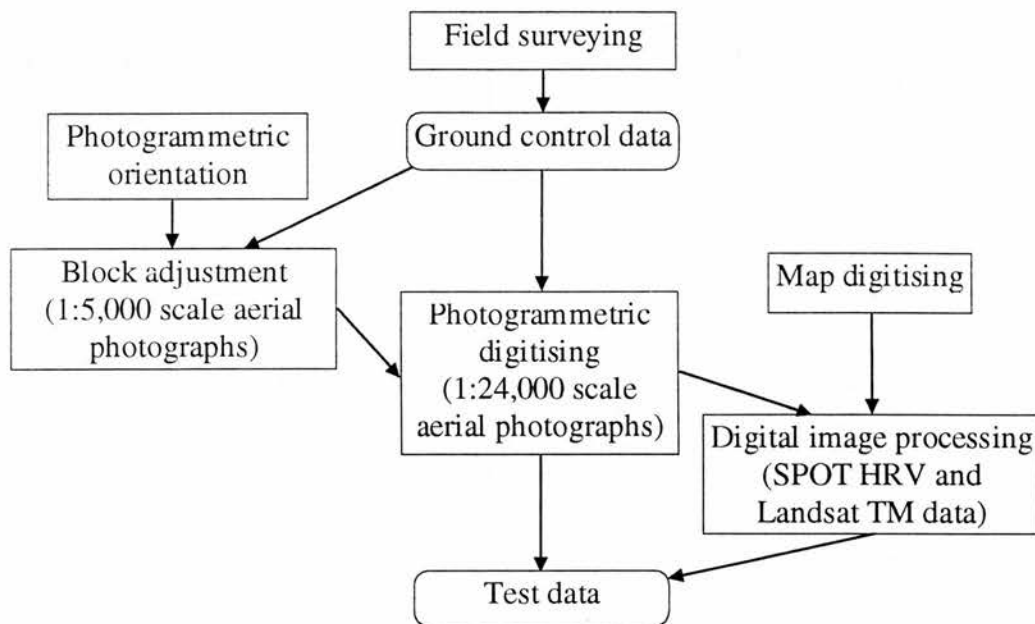


Figure 5.4 The scheme to build test data for the case study

The equipment utilised in the case study included the Summagraphics tablet for map digitising, the AP190 analytical plotter for photogrammetric digitising based on aerial photographs, and the ERDAS (Earth Resources Data Analysis System) software for remotely sensed digital image processing, as shown in Table 5.3. This set of equipment was available in house at the University of Edinburgh.

Table 5.3 The equipment for test data automation

Data sources	Equipment
OS plans	Summagraphics tablet
Aerial photographs	AP190 plotter
Satellite images	ERDAS workstation

#### 5.4.1 Land surveying

As mentioned previously in this section, land surveying was carried out to provide adequate GCPs as photocontrol for photogrammetric block adjustment based on the 1:5,000 scale aerial photographs and for photogrammetric digitising based on the 1:24,000 scale aerial photographs. Some of these GCPs were for height only, while others were for planimetric only.

Height control points were established by levelling, based on third order bench marks (Newlyn Datum). The bench mark heights for the central part of the study area were obtained directly from Ordnance Survey bench mark lists, while those for peripheral locations were read from OS 1:2,500 scale plans. A total of 46 bench marks were utilised, as shown located in Figure 5.5. The levelling procedure was strictly followed, with closing errors less than 1 cm. The resulting levelled height points are shown in Figure 5.5. The relatively denser distribution of height points along roads was mainly due to the convenience of levelling in these areas.

Planimetric control points were derived based on the observation of angles and distances made by theodolite (Wild T2) and EDM (Wild Distomat), respectively. The locational control network was tied to OS National Grid coordinates provided for



Arthur's Seat and Blackford Hill. In the light of the requirements of photocontrol and the special "setting" of the study area, a point on the roof of the 10-storey Darwin Building on the King's Buildings site was established, based on the known stations on Arthur's Seat and Blackford Hill with redundant observations, as shown in Figure 5.5.

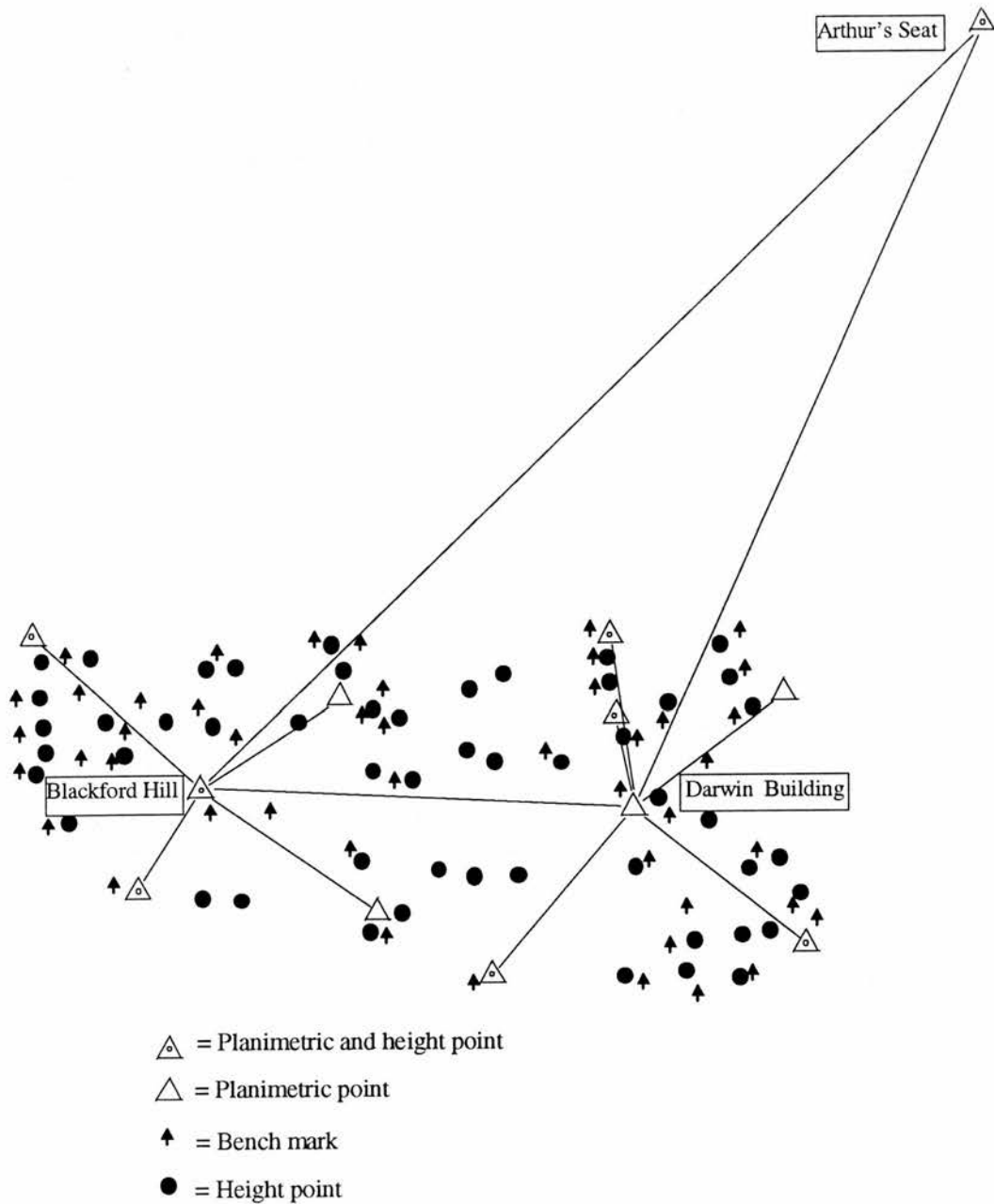


Figure 5.5 The distribution of control points produced from land surveying in the local Edinburgh area (points are only approximately located).

The known stations on Arthur's Seat and Blackford Hill and the established point on the Darwin Building were then used as the three major reference points to measure distances and azimuths to another 9 points. During the field work, repeated measurements were taken, together with careful booking and sketching on site; these are extremely important in order to eliminate the risk of gross errors and to ensure awareness of sufficient accuracy.

Adjustment computation was accomplished for levelled points and planimetric points with redundant measurements using the condition equation method, while other points without additional distance or angular constraints were calculated directly using trigonometry. Wolf (1987) provided practical reference to surveying adjustment. It was found that adjustment by "condition equation" is much easier than by "observation equation" method. The surveyed planimetric points have accuracies ranging from 8 to 18 centimetres. These ground control points are sufficiently accurate for the block adjustment, photogrammetric orientation and digitising, given the limitations imposed by surveying instruments, time and manpower costs.

In summary, this field work to provide photocontrol for photogrammetric block adjustment produced coordinates for 10 planimetric points and 58 height points, as shown in Figure 5.5. Their coordinates are listed in Appendix 1.

#### **5.4.2 Control densification using block adjustment based on 1:5,000 scale aerial photographs**

The number of field surveyed control points was hardly sufficient for individual photogrammetric stereo models to be rigorously checked, although height control was relatively strong. Given the requirement for this test, extensified control was obtained by photogrammetric block adjustment, as described below.

To perform block adjustment, an analytical plotter AP190 produced by Cartographic Engineering A/C was used. This instrument is characterised by modular design, and menu driven interactivensess. Table 5.4 summarises the characteristic features of the AP190. An interesting feature of the AP190 is that data collection from aerial photographs by means of pre-orientated models is nearly as straightforward as digitising from maps. The orientated aerial photographic pairs provide a precise three dimensional model from which X, Y, Z coordinates of any visible feature can be measured. The AP190 can also store, display, edit and manage digitised data with a high degree of operator-instrument interaction.

A commercial photogrammetric block adjustment package BLOKK based on independent model approach is installed on the AP190, and was utilised to densify the locational control network, thus allowing for photogrammetric digitising.

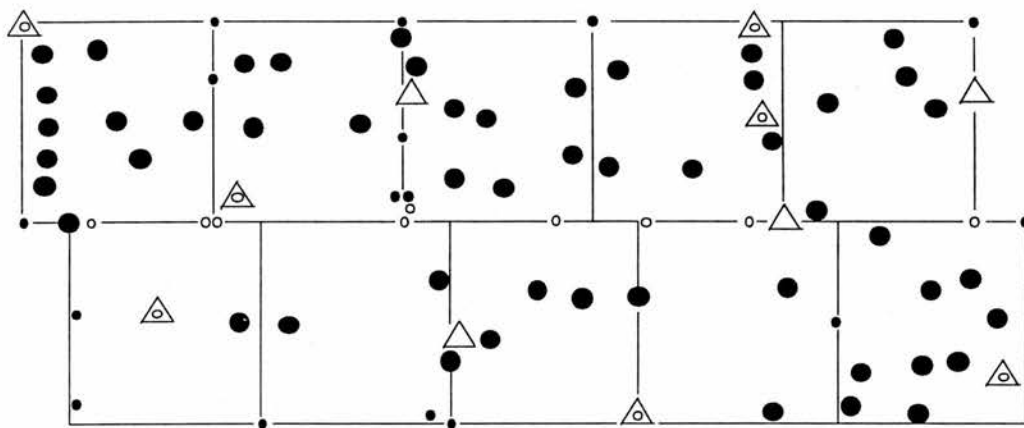
Table 5.4 The main features of the analytical plotter AP190

Mirror stereoscope	- 4X or 6X or 8X magnification - 50 micron measurement light dot
Photo-carrier	- 5 micron resolution linear encoder - all format photos (35 mm to 230 mm)
Parallax adjustment	- stepper drive (2.5 micron resolution)
Controls	- rotary encoder - record button - function keys
Interface	- Z-80 microprocessor with RS232 port

First, pass points and tie points should be carefully located in order to achieve strong controls for all stereo models within the block. Using the definition of Wolf (1988), pass points are points of extended horizontal control, used for controlling subsequent photogrammetric procedures such as planimetric mapping or mosaic construction.

They are also necessary to continue the triangulation through one strip. They must be located in desirable positions on three successive overlapping photographs, ideally, opposite to the principal points and conjugate principal points. Pass points common to two strips should be chosen in the centre of the side lap area. These special pass points are often known as tie points (i.e., to tie strips together). Pass or tie points may be natural well-defined objects or artificially marked points. A total number of 25 pass points (including 9 tie points) were planned as in Figure 5.6.

In accordance with BLOKK procedures provided, all the control points, pass points and tie points as shown in Figure 5.6 were measured from 10 stereo models on the AP190 analytical plotter. Before measurements were taken, the models were accurately set in relative and absolute orientations.



- △ = Ground control point with known X, Y, Z
- △ = Ground control point with known X, Y
- = Ground control point with known Z
- = Tie point
- = Pass point

Figure 5.6 Block configuration, Blackford Hill, Edinburgh.

Measurements for each model were recorded in its relevant job file. It is worth noting that relative orientation is usually performed in an iterative way, with each iteration

indicating the points with big residuals, thus guiding the operator to do re-measurements as necessary. It is recommended that relative orientations of all the stereo models be completed with overall standard deviation below 20 microns.

Prior to adjustment computation, model coordinates for control and pass points were derived from corresponding job files. With ground control point data established and proper parameters set as required, BLOKK was performed to do the block adjustment until the adjustment converges to a satisfactory result. In order to ensure the reliability of the adjustment, certain check points were set aside to assess the adjustment. In order to ensure the reliability of this block adjustment, two tests were carried out. The two tests used the same model coordinate measurements but slightly different setting of controls in the interior of the block. The adjustment results are tabulated in Table 5.5.

Table 5.5 Photogrammetric block adjustment results (unit: metres)

Test Results	I		II	
	XY	Z	XY	Z
RMS residual in photogrammetric measurements	0.233	0.289	0.163	0.282
RMS residual in check points	0.350	0.456	0.294	0.419
$S_0$	0.262	0.429	0.181	0.427

As shown in Table 5.5, the results conform to the prediction by Andersen (1993, pers. comm.) that standard deviation ( $S_0$ ) in Z will be 2 to 3 time that in XY. Specifically, the adjusted Z outputs from both tests are not significantly different, while consistently better results in XY are achieved from test I. The ratio of  $S_0$  in Z against that in XY is about 1.6 (less than 2) for test I, while the ratio for test II is

approximately 2.4 (between 2 and 3). Better accuracy was achieved from test II, and the results from test II were those actually used in subsequent processing, as listed in Appendix 2.

#### **5.4.3 Map digitising to supply control points for remote sensing data**

As mentioned previously, due to the larger spatial extent of SPOT HRV and Landsat TM images beyond the aerial photographic block configuration, and in order to ensure data independence between photogrammetric data and satellite data, control points different from field surveyed and densified GCPs had to be provided. These control points were digitised from OS large scale plans, i.e., 1:1,250 and 1:2,500 scale plans for areas covering sub-scenes of SPOT HRV and Landsat TM images, as shown previously in Figure 5.2.

It has been discussed previously that there are a number of factors contributing to the uncertainties encountered in paper map digitising. For instance, machine imprecisions and operator biases introduce a variety of uncertainties during the digitising process. Moreover, maps are usually produced via some map generalisations, which are abstractive, selective and approximate in nature. Depending on specific purposes and applications, different maps may have undergone different generalisation processes with different effects. Therefore, there are complex uncertainties associated with maps themselves and map digitisations. Because of that, some special techniques are required to figure out various uncertainties occurring in map digitising as a compounded result of map generalisations, machine and operator limitations.

Given the condition that only well-defined points (e.g., pillar stations, buildings and other identifiable or ground surveyable man-made structures) will be considered, it is therefore assumed that there is no uncertainty "locked in" during map generalisation and only those uncertainties imposed by operator and machine limitations will be

relevant in the case of OS larger scale plans. Some tests may then be carried out to find out the uncertainties during digitisation of well-defined points on maps. However, adequate and accurate samples of field surveyed points are needed in this case, and, when it is difficult to get sufficient field surveyed points due to timing or economic constraints, the accuracy of the digitised coordinates can be alternatively evaluated by using the OS National Grid grid intersections on the OS plans, which are of known coordinates by definition.

Following this idea, a test has been performed to derive some empirical estimates about the accuracy obtainable from OS plans at larger scales. The results are tabulated in Table 5.6, using equations (3-1), (3-2) and (3-3), where  $sd_x$  and  $sd_y$  stand for the standard deviations in X and Y respectively.

Table 5.6 Test results of manual digitising on OS plans (unit: metres)

Tests			Results	
map sheet	scale	number of points	$sd_x$	$sd_y$
1	1:1,250	21	0.130	0.093
2	1:2,500	56	0.186	0.229

Unsurprisingly, accuracy tends to decrease as map scale becomes smaller. It is interesting to note that the accuracy in x, measured by the standard deviation  $sd_x$ , is only slightly degraded when map scale becomes smaller (from 1:1,250 to 1:2,500), while the accuracy in y, measured by the standard deviation  $sd_y$ , is more than halved when map scale is reduced by half.

#### **5.4.4 Photogrammetric digitising of aerial photographs at 1:24,000 scale**

As sufficient control and check points had been obtained jointly by the land surveying and the control densification, photogrammetric digitising was then performed for aerial photographs at 1:24,000 scale, based on properly mounted and orientated stereo models on the AP190 photogrammetric plotter.

Photogrammetric digitising was performed in accordance with the classification system set previously. For built-up land cover, most of the building blocks can be clearly identified, and their outlines can be precisely followed and digitised. As an example, a map of land cover digitised from 1:24,000 scale aerial photographs is shown in Figure 5.7. The boundary of the lake (Blackford Pond) can also be accurately tracked. However, for other objects such as boundaries separating a patch of trees from bordering grass, the positioning of boundaries clearly involves great subjectivity. Similar fuzziness is also encountered when delineating a patch of shrub land.

Clearly, different operators would produce different land cover maps, even from the same data sources and using the same techniques. By the same token, it would be impossible for the same operator to generate exactly the same land cover maps even using the same data sources and the same techniques at different times. For this reason, an evaluation of the accuracy of digitising is very important.

As mentioned previously in this section, data digitised from the 1:24,000 scale aerial photographs were used as reference data, against which classified data based on the SPOT HRV and the Landsat TM images were tested. For this reason, the land cover data derived by interpreting the 1:24,000 scale aerial photographs are assumed to be accurate, that is, 100 % correct, in terms of attribute accuracy (classification), and only positional accuracy is relevant for this case study.



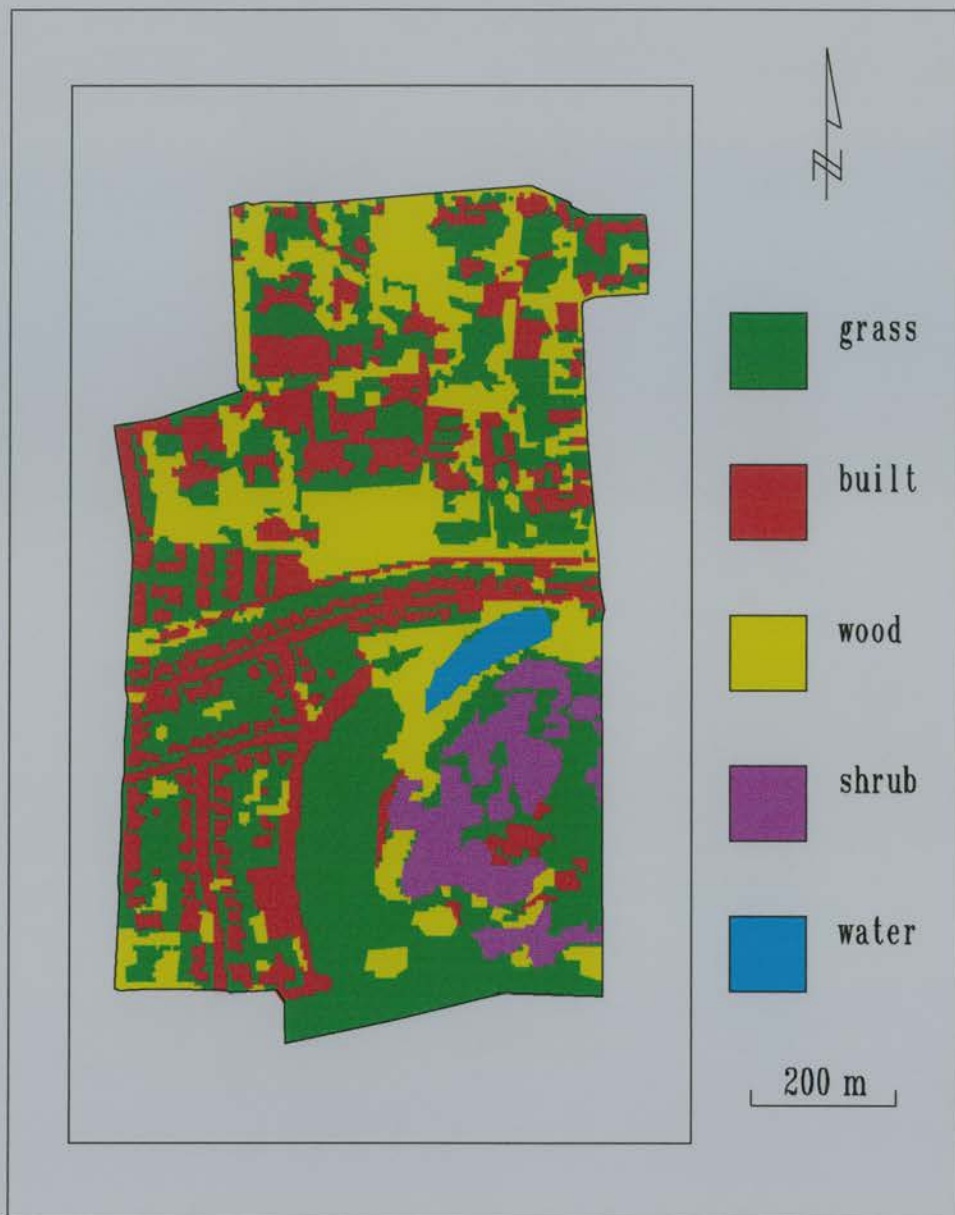


Figure 5.7 Land cover map derived from the 1:24,000 scale aerial photographs

Thus, in order to provide initial estimation of the positional accuracy of photogrammetric digitising, tests have been carried out, in which a few measured three-dimensional coordinates of well-defined objects such as outstanding landmarks, bounding corners, centres of manholes and road junctions are checked against the block densified data. The checking points must be withheld from use in controlling the stereo model, so that independence in checking is not violated. Results are listed in Table 5.7, again, using equations (3-1), (3-2) and (3-3), but with standard deviations of both planimetric (X, Y) and height (Z) data.

Table 5.7 Photogrammetric point digitising accuracy estimate (unit: metres)

Test		Results		
model scale	number of points	sd <sub>x</sub>	sd <sub>y</sub>	sd <sub>z</sub>
1:24,000	15	0.595	0.569	0.616

As shown in Table 5.7, there is sub-metre accuracy when digitising from medium scale aerial photographs. In particular, the planimetric (X, Y) accuracy is just slightly higher than the accuracy in height (Z). This confirms the values of large and medium scale aerial photographs in providing accurate three-dimensional measurements for urban applications, where accuracy is an important consideration for land use planning and management.

The task of evaluating accuracy becomes less straightforward when dealing with lines and polygons. A usual way of doing this is by overlaying the test data set with an assumed reference data set so that an estimate of digitising accuracy (the epsilon error band width) can be obtained (Chrisman 1991; Edwards and Lowell 1996). For example, the land cover map digitised from the 1:24,000 scale aerial photographs might be overlaid onto a map digitised from the 1:5,000 scale aerial photographs. Such a test was not carried out in this research for the reasons developed below.

The epsilon band width mentioned previously would stand as a global estimate of the accuracy obtainable in digitising linear objects. However, it is very likely that different types of objects will have different epsilon band widths. For example, the boundaries of building blocks can be more accurately identified and delineated than boundaries separating continuous cover types such as grassland, shrub and woodland around Blackford Hill, as shown in Figure 5.7.

More seriously, even though the measure of epsilon band width is unique for individual lines, it would still not possible to provide any indication about the spatially varying nature of attribute accuracy beyond simply global indicators such as the correctly classified percentage of land cover types (Goodchild, Lin and Leung 1994). For example, suppose an epsilon band width were estimated for each arc shown in Figure 5.7. One would only be able to know how accurately each arc were digitised. But, it would not be possible to know how far homogeneity is applicable within each polygon. The reason is as follows.

Referring back to the aerial photograph shown in Figure 5.1, it is common to find two patches of grass, which may exist in different stages of growth, with different underlying soil types, and with different inclusions of other land cover types such as bare ground and shrub, but it may not be possible to depict each patch appropriately using polygon models as employed in the production of land cover maps such as that shown in Figure 5.7, unless sub-division of classes is adopted. Even if classes were subdivided, there would still be a problem of spatial variations within each subdivided classes. In other words, the uncertainties present in the interpretation of land cover types have not been deleted, but only pushed into a lower level. Similar problems were discussed by Wang (1990). Therefore, the epsilon band model would not be adequate to describe the heterogeneities so evident in the mapping of urban and suburban land cover.

Indeed, this issue of spatial heterogeneities is completely different from the problem of positional accuracy in discrete objects, which is relatively well researched, and spatial heterogeneity may not be effectively handled by an object-based perspective of geographical reality (Burrough 1994). A sensible way to address the issue of spatial heterogeneities encountered in aerial photo interpretation would be by using field-based models where fuzziness can be accommodated. This is what the surface-based approach pursued in this thesis is designed for.

#### **5.4.5 Remote sensing image classification**

As described in Section 2.3.3 and Section 5.3, an important function of remote sensing processing is classification, which provides efficient and current information for a variety of purposes such as urban and suburban land cover mapping. For the purpose of this research, remote sensing images were also incorporated in the test data hierarchy, as mentioned in Section 5.3. This sub-section will describe how land cover data were derived from classification of remote sensing images, followed by an assessment of the classification accuracy.

Because the test data need to be derived from remote sensing images for the purpose of comparison testing of data accuracies, remotely sensed data must be properly classified and geometrically rectified to the same coordinate system as in photogrammetric digitising.

Unsupervised classification methods such as clustering were found to be unsuitable for both SPOT HRV and Landsat TM data in this research. For both images, large number of clustering classes were shown overlapping in spectral feature space and could not easily discriminated. Moreover, classes found by unsupervised classification methods could not be related easily to any particular land cover types recognisable on the ground, as also described in Chapter 2.

This conclusion did not come as a surprise, because urban and suburban phenomena tend to be characterised by much complexity within small areas, resulting in highly variable reflectance values over short distances, which may not be detected by the fairly broad spectral bands, on which SPOT HRV and Landsat TM sensors are designed to work. This situation gives rise to a common problem in urban applications of remote sensing data: mixed pixels, as discussed in Chapter 2. Forster (1983) met similar problem in his study in a municipal area based in Sydney, Australia.

Because of these problems with unsupervised classification methods, a supervised classification, specifically the Maximum Likelihood Classification (MLC), was performed. It was necessary to use knowledge of the study area and to inspect the satellite image on display using visual interpretation skills. Since the study area was within relatively easy access, several field visits were made for this purpose, with sketches or pictures created as necessary. Then, it was possible to proceed to classify certain areas (training areas) of the image as being characteristic of a certain land cover type. Following the selection of training areas, it is possible to generate spectral signatures based on the pixels of the image falling within each designated training area.

An iterative approach is usually necessary to generate the most suitable signature files for MLC supervised classification. For example, it is important, especially when dealing with urban areas, to revise training areas perhaps to remove individual or groups of unrepresentative pixels. As a single class of objects may not be represented by a single spectral class, a number of spectral classes are often created, and then are merged into object classes. For example, both a vegetation-dominated garden and a soil-dominated worked allotment may belong to a grassland cover type. However, it was rare to find both sub-classes exhibit identical or even similar spectral signatures. Thus, two different spectral classes were formed, and needed to be merged into the

grassland cover type in the end. For the SPOT HRV image, probably due to its finer spatial resolution than that of Landsat TM data, built-up areas had to be subdivided into residential (i.e., detached or terraced houses) and non-residential areas such as commercial, institutional complexes (including roads and bare grounds) for satisfactory classification.

In addition to the creation of appropriate signature files, it was also necessary to select the optimal band combination. That was done by running an ERDAS module named **Diverge**. This module provides a mechanism to select those combinations of bands with largest measures of divergence among target classes in order to increase the classification accuracy (Campbell 1987). Such a process led to the selection of bands 1, 2 and 3 for the SPOT HRV image, and bands 3, 4 and 5 for the Landsat TM data. The SPOT HRV and Landsat TM images were then classified using the MLC algorithm in the ERDAS software system.

Following the process of image classification, geometric rectification was performed using ground control points obtained from map digitising. The first order transformation was found sufficient to gain sub-pixel accuracies for both SPOT HRV and Landsat TM images. Resampling was performed using a nearest neighbour resampling scheme, based on classified images in order to maintain (rather than "spoil") the original image's integrity. The processes of classification and geometric rectification are tabulated in Table 5.8. The classified SPOT HRV and Landsat TM images are shown in Figure 5.8 and Figure 5.9, respectively.

Attention is now given to the assessment of classification accuracies for both SPOT HRV and Landsat TM data. The classification accuracy was tested firstly by simple random sampling throughout the image and secondly by superimposed comparison (i.e., exhaustive pixel-by-pixel comparison) made between classified SPOT HRV or Landsat TM data and reference photogrammetric data.

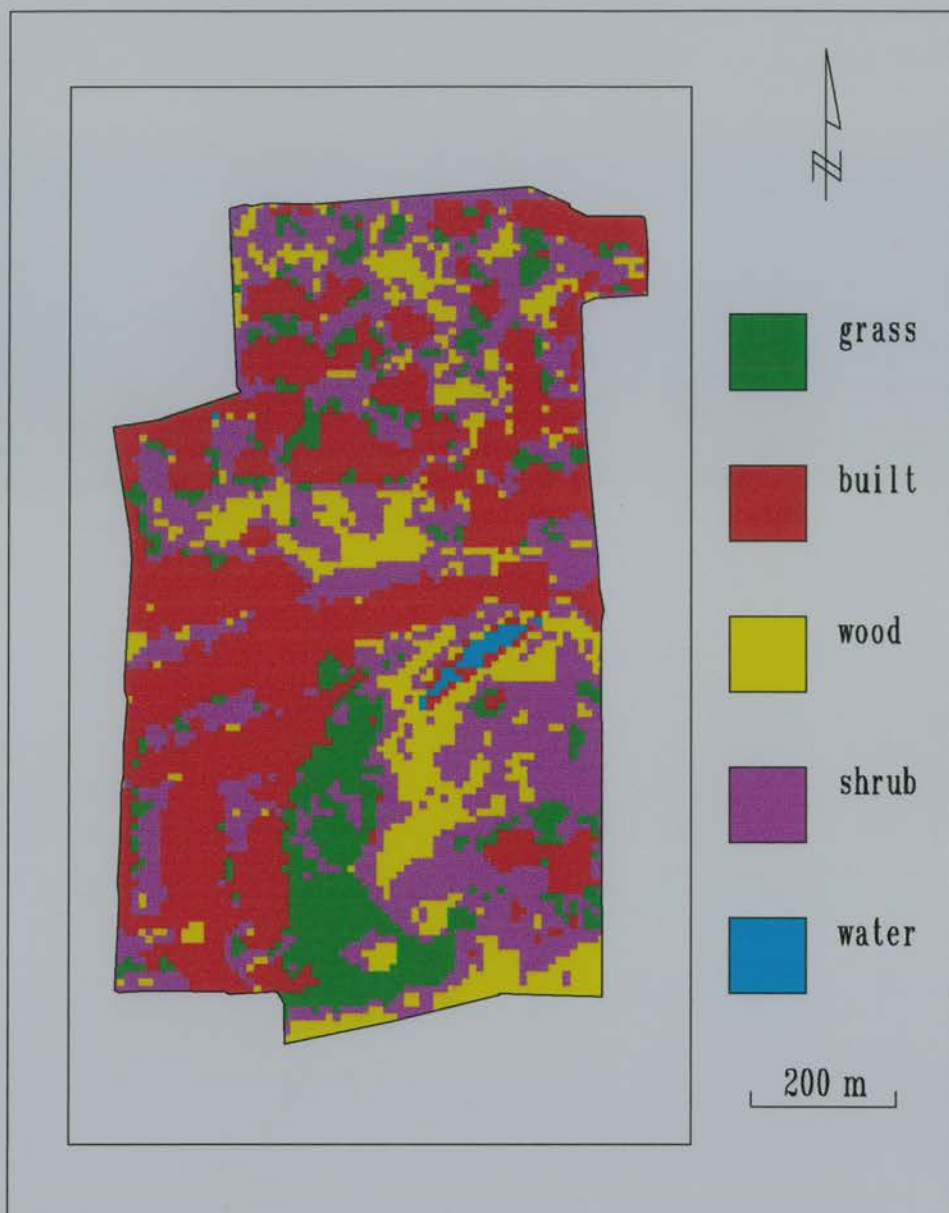


Figure 5.8 Land cover classification derived from the SPOT HRV image

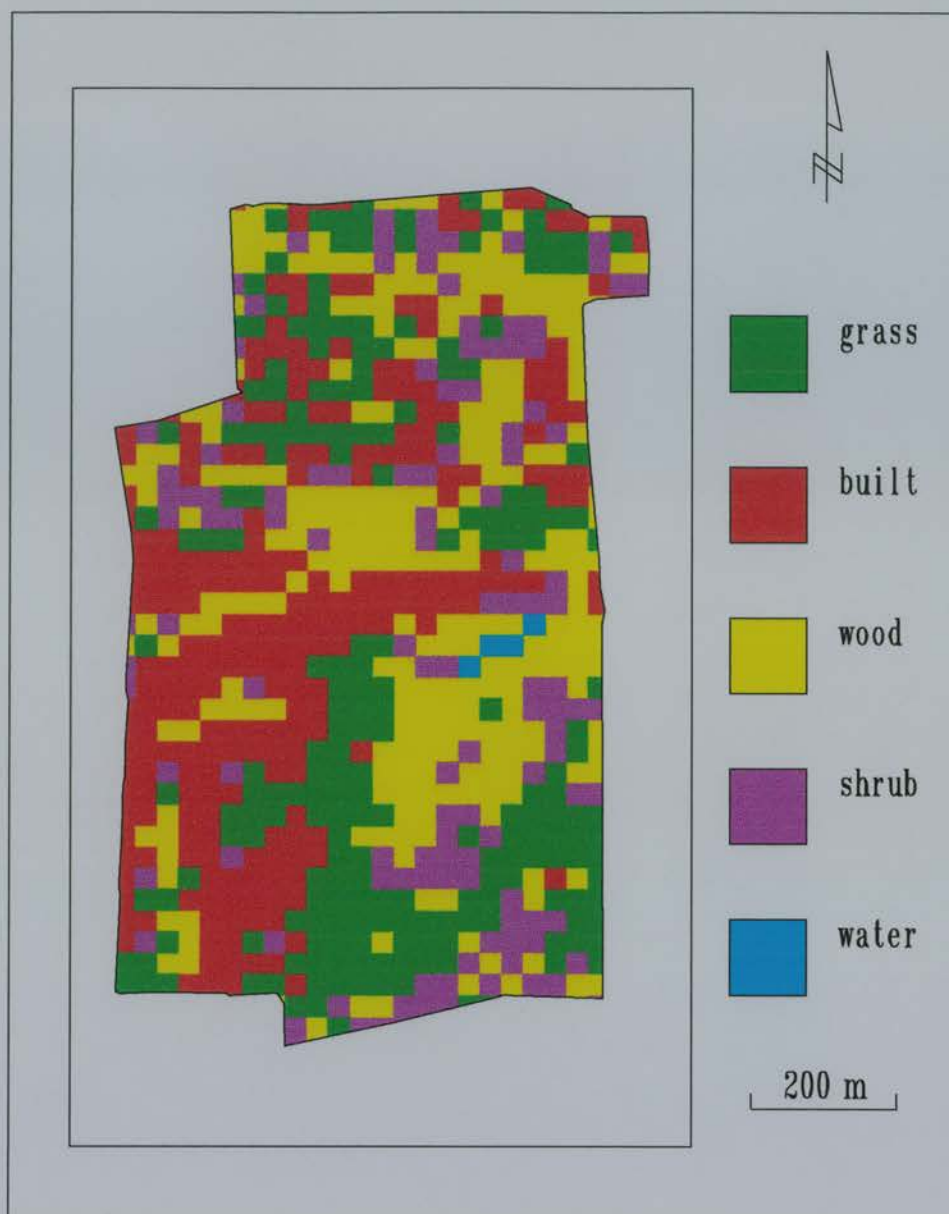


Figure 5.9 Land cover classification derived from the Landsat TM image



Table 5.8 A summary of remote sensing data processing

Type of processing	SPOT image	TM image
Classification		
Classifier	MLC	MLC
No. of spectral classes	7	6
No. of object classes	5	5
Selection of bands	bands 1, 2 and 3	bands 3, 4 and 5
Rectification		
No. of GCPs used	28	27
Order of transformation	1	1
RMS of x, y	0.95 pixel	0.89 pixel
Resampling	nearest neighbour	nearest neighbour

The simple random sampling was facilitated by a built-in module named **Randcat** in the ERDAS software system, which produced 379 and 196 random samples for SPOT HRV and Landsat TM image, respectively. This module prompted the analyst for the ground land cover types at the randomly chosen sample pixels. Land cover types for each sample pixel were supplied from consulting the classified land cover map shown in Figure 5.7. This land cover map was the assumed reference data, though with uncertainties present as discussed in the previous sub-section. The results were corresponding misclassification matrixes, from which various statistics could be calculated by another module named **Claserr** in the ERDAS software system. Results are shown in Tables 5.9 and 5.10 for SPOT HRV and Landsat TM data respectively. The derivations of overall agreement, chance agreement and kappa coefficient are explained in Chapter 3.

Table 5.9 Error matrix for SPOT HRV data obtained by simple random sampling

		Reference data					row total	row marginal
		grass	built	wood	shrub	water		
Classified data	grass	49	2	0	3	0	54	0.14
	built	3	182	4	3	0	192	0.51
	wood	1	0	12	4	0	17	0.04
	shrub	18	18	28	51	0	115	0.30
	water	0	0	0	0	1	1	0.00
	column total	71	202	44	61	1	379	
	column marginal	0.19	0.53	0.12	0.16	0.00		

Overall agreement = 78%, chance agreement = 35%, Kappa = 66%

Table 5.10 Error matrix for Landsat TM data obtained by simple random sampling

		Reference data					row total	row marginal
		grass	built	wood	shrub	water		
Classified data	grass	48	18	1	5	0	72	0.37
	built	8	51	2	14	0	75	0.38
	wood	0	1	5	6	0	12	0.06
	shrub	6	1	0	30	0	37	0.19
	water	0	0	0	0	0	0	0.00
	column total	62	71	55	8	0	196	
	column marginal	0.32	0.36	0.28	0.04	0.00		

Overall agreement = 68%, chance agreement = 28%, Kappa = 55%

As indicated in Table 5.9, significant number of pixels of the SPOT HRV image were classified as “shrub” when they were “wood”, “grass” or “built” on the reference data. This is probably because firstly “shrub” is the least well-defined type as to the SPOT HRV sensor and secondly “shrub” is heavily blended with bordering grass, wood and built land cover types, as evident from the aerial photograph shown in Figure 5.1. Moreover, it must be admitted that the assumed “ground truth”, i.e., the photogrammetric data, is by no means perfectly “accurate”, as discussed in the previous sub-section. Therefore, disagreements between classified remote sensing data and photogrammetric data are inevitable.

For Landsat TM data, on the other hand, the pattern of misclassification was notably different. As shown in Table 5.10, “built” land cover was quite often classified as “grass”. Misclassification of “shrub” as “built” was also obvious. In addition to the reasons mentioned above in the case of SPOT HRV data, another reason may be that decreased spatial resolution tends to “secure” more mixed pixels, leading to more misclassification. More outstanding is the confusion of “wood” with “shrub”, as more pixels with “shrub” land cover type were classified as “wood” than “wood” pixels themselves on the reference data, conforming to the poorest discrimination between “wood” and “shrub” as seen in Table 5.10.

In general, SPOT HRV data were more accurately classified than Landsat TM data, as evident from their respective overall classification accuracies and Kappa coefficients. Moreover, it becomes clear from Tables 5.9 and 5.10 that it is very important to use Kappa coefficients to give a fair assessment of the accuracies in classified remote sensing images. This is because classification accuracies, without chance agreement being taken into account, might be significantly inflated: 78% opposed to 66% for the SPOT HRV image, and 68% opposed to 55% for the Landsat TM image, as shown in Tables 5.9 and 5.10. Similar observations were made recently by Veregin (1995).

Attention is now given to the superimposed comparison method, because it is believed that a superimposed comparison can facilitate the examination of the pattern and hence the revelation of the sources of misclassification. Chrisman (1991) recommended the superimposed comparison method as the most reliable and comprehensive evaluation of classification accuracy relevant to categorical maps as a whole.

To facilitate the superimposed comparisons, photogrammetric data in vector format needed to be transformed into raster data with pixel sizes equal to that of SPOT HRV (10 m) and Landsat TM (30 m) respectively. Each pixel on the resulting raster data sets was assigned a single class. The superimposed comparisons were made by using a **Matrix** module in the ERDAS software system, which generated the 25 different combinations of classified and reference land cover types. It is worth noting that the test data and reference data must be co-registered before superimposing comparisons. The statistics for combined classes were then used to form misclassification matrices (Campbell 1987; Congalton 1991). Results are shown in Tables 5.11 and 5.12 for SPOT HRV and Landsat TM data respectively.

Examination of Table 5.11 indicates that significant number of “grass” pixels were misclassified as “built”, “wood” and “shrub” land cover types. As the majority of the study area is occupied by the “grass” type, which is not “pure”, “grass” type is at fault for much of the misclassification. Secondly, “wood” type was also seriously confused with “shrub”. Reasons as discussed in the case of simple random sampling may also apply in this situation.

Table 5.11 Error matrix for SPOT HRV data obtained by superimposed comparison

		Reference data					row total	row marginal
		grass	built	wood	shrub	water		
Classified data	grass	749	85	48	5	0	887	0.12
	built	1382	1137	124	3	23	2669	0.38
	wood	357	37	597	125	22	1138	0.16
	shrub	1171	137	743	323	5	2379	0.33
	water	1	0	0	0	39	40	0.01
column total		3660	1396	1512	456	89	7113	
column marginal		0.51	0.20	0.21	0.06	0.01		

Overall agreement = 40%, chance agreement = 19%, Kappa = 26%

Table 5.12 Error matrix for Landsat TM data obtained by superimposed comparison

		Reference data					row total	row marginal
		grass	built	wood	shrub	water		
Classified data	grass	138	19	34	14	0	205	0.26
	built	125	71	22	0	0	218	0.28
	wood	97	18	86	25	7	233	0.30
	shrub	60	16	22	18	1	117	0.15
	water	1	0	0	0	3	4	0.01
column total		421	124	164	57	11	777	
column marginal		0.54	0.16	0.21	0.07	0.01		

Overall agreement = 41 %, chance agreement = 26%, Kappa = 20%

For Landsat TM data, the confusions of “grass” with “built”, “wood” and “shrub”, and between “wood” and “shrub”, as found with SPOT HRV data, were further emphasised, as shown in Table 5.12. In terms of the corresponding Kappa coefficients rather than simply the overall classification accuracies, the Landsat TM image was less accurately classified than the SPOT HRV image.

So far in this sub-section, two different methods of evaluating classification accuracy have been discussed. The results are summarised in Table 5.13 to see if there are any reasons for the different levels of accuracy indicators obtained there, such as Kappa coefficients, to assist interpretation.

Table 5.13 The summary result of evaluation of classification accuracies

Classification accuracy	SPOT HRV image	Landsat TM image
Overall accuracy, Kappa		
- by simple random sampling (sample size)	78%, 66% (379)	68%, 55% (196)
- by superimposed comparison (sample size)	40%, 26% (7113)	41%, 20% (777)

As can be seen from Table 5.13, the superimposed comparison results in significantly lower accuracy than simple random sampling (in particular, less than half of the Kappa coefficients than by simple random sampling), although both methods are assessing the same classified images. The first reason that simple random sampling tends to overestimate the classification accuracy may be the effects of sampling size. Another reason is probably because the simple random sampling by human visual interpretation may somehow compensate the disagreements caused by geometric distortion present even in rectified remote sensed image, while machine-forced comparison forces all discrepancies to be labelled as misclassification.

#### 5.4.6 Discussion of the results

So far in this section, a hierarchy of test data has been acquired in the required form as set in Table 5.2. It was seen, firstly, that spatial data acquisition is a demanding job: it requires trained and versatile personnel, who can carry out field work, and use computers and specialised field and laboratory equipment; the data are obtained only through time-consuming processes. Under normal procedures, choices between using aerial photographs or satellite imagery are made based on criteria such as cost, time, availability and accuracy. Secondly, because one of the aims of this research is to highlight the uncertainties in spatial data, a detailed description about the data sources used and the data acquisition techniques employed has been made in order to identify the error sources, measure their magnitudes and assess the accuracy level attainable in various data products. It is useful here to provide a summary of the varying accuracies of the hierarchy of test data. This is presented as Table 5.14, where classification accuracies derived by superimposed comparisons are reported.

Firstly, the accuracies of GCPs are very high (in centimetres). For planimetric accuracies, the field surveyed GCPs have the highest accuracies, the GCPs digitised from OS 1:1,250 scale plans have the second highest accuracies, densified GCPs from the 1:5,000 scale aerial photographs have the third highest accuracies, and GCPs digitised from 1:2,500 scale OS plans have the lowest accuracies. For accuracies in height, field surveyed height points are far more accurate than those densified points using the 1:5,000 scale aerial photographs.

Secondly, lower accuracies are obtained by photogrammetric digitising based on the 1:24,000 scale aerial photographs. The positional accuracy of photogrammetric data digitised from 1:24,000 scale aerial photographs was evaluated by measuring a few well-defined objects, and then checking with reference data (i.e., densified control points based on 1:5,000 scale aerial photographs). Such a method produced an

average estimate, which was merely some aspatial uncertainty measure assumed applicable homogeneously over the mapped area.

Table 5.14 The hierarchies of test data and their accuracies

Hierarchy of test data	Hierarchy of accuracies			
	Positional (metre)		Attribute (%)	
	Planimetric (X,Y)	Height (Z)	Overall classification accuracies	Kappa coefficients
1. GCPSs:				
• Field surveyed	0.08 ~0.18	0.01		N/A
• Densified (1:5,000 scale aerial photographs)	0.18	0.43		N/A
• Digitised				
1:1,250 scale plans	0.13, 0.09			N/A
1:2,500 scale plans	0.19, 0.23			N/A
2. photogrammetric data (1:24,000 scale aerial photographs)	0.60, 0.57	0.62		N/A
3. Remote sensing data:				
• SPOT HRV data	9.50		40	26
• Landsat TM data	25.70		41	20

However, as described in Sub-section 5.4.4, some well-defined objects such as identifiable pillar points established by the Ordnance Survey, road junctions, walls and fences, building corners and other landmarks can be precisely positioned and repeatedly measured, and so more accurately measured, while others objects such as boundaries separating wooded land and grassland are fuzzy and are difficult to model. Furthermore, attribute accuracy was not addressed in any depth, but assumed to be “absolutely accurate” for the purpose of checking remote sensing data. It is likely that attribute accuracy of photogrammetric data will be far more complicated than the



positional accuracy would suggest, and an object-based epsilon band model would not provide any information on the spatial heterogeneities characteristic of categorical land cover maps such as that shown in Figure 5.7.

Finally, for classified remote sensing data, the positional accuracies were implied by the RMSE measures, which indicated sub-pixel accuracies for both SPOT HRV and Landsat TM data. The attribute or classification accuracies, on the other hand, were evaluated by constructing appropriate error matrices, from which overall classification accuracies and Kappa coefficients were calculated and interpreted. It is seen that the SPOT HRV data are more accurately classified than the Landsat TM data, as measured by the Kappa coefficients, because the pixel size of SPOT HRV image is finer than that of Landsat TM image. However, these measures can provide information on accuracy obtained just at individual classes level rather than at pixel level.

In general, although the accuracy may be improved by increasing the measurement resolution and precision, reinforcing rigorous data acquisition standards, and refining the measuring procedures, spatial data acquisition is bound to be limited by the human subjectivity and machine imprecision unavoidable in any classification. Besides, there are always some compromises between the accuracy of the data and the cost incurred for better accuracy. An example would be to use one stereo model of 1:24,000 scale aerial photographs, rather than many models at 1:5,000 scale.

To assist the user or project planner, it is important that the data suppliers should provide to the users the accuracy levels expected in the data as integral components of the data themselves. As will be discussed in the next chapter, such a strategy facilitates the provision of data quality, which can in turn be argued, will ultimately benefit users as they are able to assess and understand both the opportunities and limitations of different products. The strategy also enhances the opportunity for the

acceptance of new mapping procedures exploring the intrinsic uncertainty in the GIS data and analysis by the research and application communities. This acceptance may only be won by the demonstrated advantages held by the new techniques through effective and flexible error handling in GIS functionality.

## **5.5 Discussion**

A co-registered data base of spatial information on suburban land cover has been generated using typical methods of photogrammetric digitising and supervised classification of SPOT HRV and Landsat TM images. It has been seen that a hierarchy of accuracies exists, which is critical when one data source is used as the positional or attribute control for other data sources.

The results confirm that, by using an object-based method, measures of positional accuracy are only average estimates, and applicable only for well-defined objects; while attribute accuracies are relevant, at best, to individual classes. Results also shows that much of the uncertainty in classified remote sensing data is due to the subjectivity in allocating discrete classes to continuously varying land cover types, which may be compounded by the fuzziness conspicuous even in photogrammetric data.

To address the issues of fuzziness, a new approach is needed, which should adapt to the spatially varying nature of the underlying phenomenon - land cover. This is what the surface-based approach will help with, as established in the next chapter.

## Chapter 6

### Empirical study II: the surface approach to uncertainties

#### 6.1 Introduction

Chapter 5 provided the means for evaluating the relative uncertainty that is found in common types of data derived from photogrammetry and remote sensing for urban and suburban land cover mapping. Results from the classification of the satellite data and validation against aerial photographs at 1:24,000 scale show that one of the main sources of uncertainties in all the methods assessed was attributable to the requirement to identify and delineate a single class attribute at each location. This requirement, which is enforced by many classification schemes, assumes a level of discreteness and homogeneity not usually held true in reality.

Given that the problem lies at least partly in an enforced conversion from continuous to discrete space, implicit in photogrammetric digitising as well as in maximum likelihood classification of remotely sensed images, there would seem to be merit in exploring methods which seek to maintain continuous data, whilst providing tools to analyse these data without necessarily having to pre-classify them first.

This is the central argument of the remainder of this thesis and it will be explained in this chapter by a study of how two particular methods, fuzzy classification and indicator kriging, may be utilised to overcome the problem described previously for urban land cover mapping. These two methods will be demonstrated, by applying the fuzzy classification to both the SPOT HRV and the Landsat TM data, and by applying indicator kriging to the photogrammetrically derived vector map. It will be shown that, in both cases, the continuous surface representation provides a more accurate

and realistic portrayal of the underlying data, and these surfaces can be created without major difficulties.

While these methods in themselves have been researched elsewhere, it is believed that this is one of the first attempts to evaluate both methods in the framework of a controlled experiment, where data of higher accuracy are available for checking. Furthermore, as most of the previous work on continuous spatial data has been in environmental research, this is a novel use of these methods in an urban and suburban land cover mapping context where objects are usually assumed to be “well-defined” and highly accurate. In fact, urban and suburban land cover exists as a continuum between two extremes: continuous fields and discrete objects.

After fuzzy surfaces have been derived, the analyses discussed in Chapter 4 are performed. Entropy measures are calculated for individual sets of fuzzy surfaces, so that the degree of fuzziness is evaluated. This is followed by the calculation of closeness (or agreement) measures based on cross-entropy and correlation coefficients.

As well as creating these surface representations, it is important to show that these can be translated into conventional map products, which are presently produced, such as the maps shown in Figures 5.7, 5.8 and 5.9. This can be done by using a “maximising” operation, by which categorical maps are produced via a process similar to the maximum likelihood classification, while information on the certainty/uncertainty of classification is supplied as a by-product. Thus, visualisation of uncertainty is enhanced by draping the derived categorical maps onto the certainty surfaces, which proves to be very impressive.

This chapter also shows that the conversion from surfaces to common categorical maps can be achieved by applying a thresholding or “slicing” operator. This

conversion from surfaces to categorical maps will be shown to be a more conceptually sound, flexible and data-driven method than previous representations of uncertainties in categorical maps using a virtually constant epsilon band width (Chrisman 1989; Pullar and Beard 1989).

Finally, categorical maps are evaluated with respect to their accuracies. While evaluation based on “sliced” categorical maps can greatly reduce the effect of mixels on the calculation of accuracies, comparisons on a relaxed condition, i.e., by comparing both the most likely and second most likely classes for each pixel, substantially increase the agreements between remotely sensed data and photogrammetric data. This is seen as particularly useful when fuzzy membership values (FMVs) for the most likely and second most likely classes are approximately equal.

## **6.2 Generation of fuzzy surfaces**

### **6.2.1 Fuzzy surfaces built on remotely sensed data**

As shown in Chapter 5, though the overall classification accuracy and the Kappa coefficients provide useful information about the reliability of classified remote sensing data, they contain no clue to the spatial variation of the classification accuracy at pixel level. In order to show the spatial variability of classified data, an approach based on fuzzy surfaces was adopted.

As discussed in Section 4.2, a fuzzy classification can be performed in either a supervised or an unsupervised mode. A supervised fuzzy classification is usually adopted to “soften” an otherwise “hard” supervised classification, by assuming a particular distribution and hence by requiring class statistics, that is, class means and variance-covariance matrices, to be defined *a priori*. An unsupervised fuzzy classification, on the other hand, seems to work well in situations where an

unsupervised conventional classification would be superior over a supervised alternative.

In a supervised fuzzy classification, FMVs are derived from intermediate outputs indicating relative strengths of class memberships of a pixel belonging to all the candidate classes, depending on a specific classifier that would be used in a hard classification. The ERDAS software can generate a probability data layer as a by-product to assist users in evaluating the classified remote sensing data. However, such a by-product is of limited use, because only the probability of a pixel belonging to the named class on the classified data is provided, leaving out the FMVs relevant to all the rest of the candidate classes. Therefore, a fuzzy classification needs to be carried out to derive FMVs relevant to all classes for individual pixels. This kind of fuzzy classification may have to rely on extra programming work rather than simply using existing systems such as the ERDAS software.

For an unsupervised fuzzy classification, the fuzzy *c*-means clustering is widely used. Unlike a supervised method that relies on a suitable distribution assumption and pre-defined class statistics, the fuzzy *c*-means clustering seeks to explore the coherence in the underlying data. In this sense, a fuzzy *c*-means clustering is similar to a conventional clustering process adopted in an unsupervised classification such as ISODATA as provided in the ERDAS software (see also Richards 1993). Moreover, as also seen in Section 4.2, the fuzzy *c*-means clustering is somehow versatile in the sense that a supervised method is “switched on” when class statistics are supplied, making it an attractive technique.

In this research, a fuzzy *c*-means clustering algorithm was programmed in FORTRAN 77 on VAX/VMS, whose code is provided in Appendix 4. This program was written in accordance with the algorithm described in Section 4.2. It is important to mention that this program requires large amount of space (384 k and 43 k bytes for the SPOT

HRV data and the Landsat TM data respectively) to store the ASCII data files loaded from LAN data files, since the existing remote sensing image processing software does not provide such a functionality.

The supervised fuzzy clustering was actually applied due to the unsatisfactory results from the unsupervised classification, as described in the previous chapter. In this case, the fuzzy *c*-means clustering program was used to calculate simply the membership values for each pixel in each of the five known classes (i.e., grassland, built-up land, wooded land, shrub land and water bodies).

An improvement on the algorithm as published in Bezdek, Ehrlich and Full (1984) was made in this fuzzy *c*-means clustering program when the supervised mode is switched on. This was actually to supply not only class centres as in the usual way, but also the class covariance matrix, thus tightening the control over the evaluation of FMVs and hence the result of classification. This original improvement was based on the reasoning that specifying a per-class covariance matrix rather than a single covariance matrix for the whole image data set would be more sensible in order to produce a fuzzy classification closer to expectations implied in the class statistics on a per-class basis.

To verify such a method, the data set shown in Table 4.2 and Figure 4.2 has been tested. Suppose there are three classes, into which the whole data set will be classified, as indicated by the three ellipses shown in Figure 6.1. In Figure 6.1, dots represent the individual pixels, while centres of classes are represented by small squares. Firstly, a global covariance matrix relevant to the whole data set is used, as in the unsupervised mode, where the global covariance matrix was supplied to calculate FMVs. As expected, the classified result is identical to that indicated in Table 4.2, though there are slight differences in the final U matrix, i.e., FMVs, because the final class centres formed in Figure 4.2 were slightly different in location from the three

class centres indicated by the small squares shown in Figure 6.1.

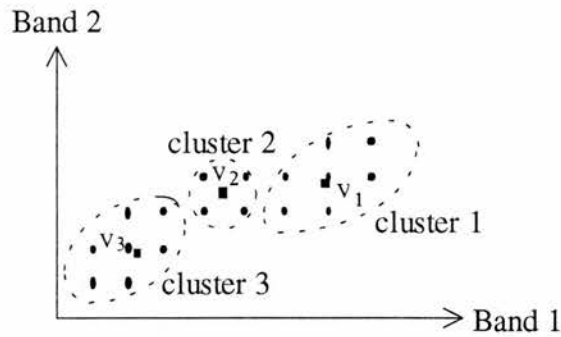


Figure 6.1 The improved algorithm on the supervised fuzzy  $c$ -means clustering, using the same data set as in Figure 4.2.

Then, the three classes are assumed to be defined by the three ellipses shown in Figure 6.1, where centres and covariance matrices can be calculated for each individual class. By such a per-class control over the calculation of FMVs, the classified result is exactly the same as defined by the three ellipses.

In addition to the fact that the improved method of fuzzy  $c$ -means clustering, in a supervised mode, helps to produce a classification with classes more tightly controlled by the class statistics, the sets of fuzzy surfaces generated will have lower degrees of fuzziness (thus higher accuracies) than those generated by usual implementations of fuzzy  $c$ -means clustering. This can be verified by calculating measures of entropy as described in Chapter 4 for the example in Figure 6.1. When a global covariance matrix is used, the resulting FMVs have an average measure of entropy of 0.86, while a lower measure of entropy of 0.69 is produced when using per-class statistics. Therefore, the improved method of the fuzzy  $c$ -means clustering will reduce the degree of fuzziness in the resulting set of fuzzy surfaces, thus making it more compatible with those fuzzy surfaces derived from photogrammetric data, which often possess lower degrees of fuzziness. This will become apparent later in this sub-section and in the next sub-section, where fuzzy surfaces built from remote sensing images



and those based on photogrammetric data are generated and subsequently analysed by calculating measures of entropy.

Based on this improved version of the supervised fuzzy *c*-means clustering algorithm, the class statistics (including class means and covariance matrices) were read out from signature data compiled from training samples, as described in Section 5.4.5. The training samples were then used as guides for the supervised fuzzy clustering. For both the SPOT HRV data and the Landsat TM data, the land cover type involving grass was divided into two subclasses, respectively relevant to worked allotment and grass/park land. The class statistics are tabulated in Appendix 4.

The fuzzy *c*-means clustering program was run in the supervised mode with class statistics supplied for both the SPOT HRV data and the Landsat TM data. The resulting U matrices consisted of the FMVs of each pixel belonging to the pre-defined classes, in which FMVs for the two subclasses of grassland cover needed to be merged. The FMVs for each class were transformed into standard ASCII file format suitable for generating an ERDAS image of single band. This image was transformed further to ARC/INFO GRID data, to allow three dimensional visualization of fuzzy surfaces. Fuzzy surfaces built from the SPOT HRV data and the Landsat TM data are illustrated in Figure 6.2 and 6.3 respectively.

The evaluation of how FMVs are partitioned among candidate classes is based, again, on the entropy measure described in Section 4.4. An entropy measure is minimised when a pixel has a full membership value (i.e., FMV = 100 %) in one class, and zero membership values (i.e, FMV = 0 %) in the remaining classes. In such a situation, the fuzzy classification is said to be of the highest accuracy (Foody 1995a). On the other hand, when FMVs are evenly distributed among all candidate classes, the entropy measure will be maximised at  $\log_2 c$ , where *c* is number of classes (*c* = 5 in this research).

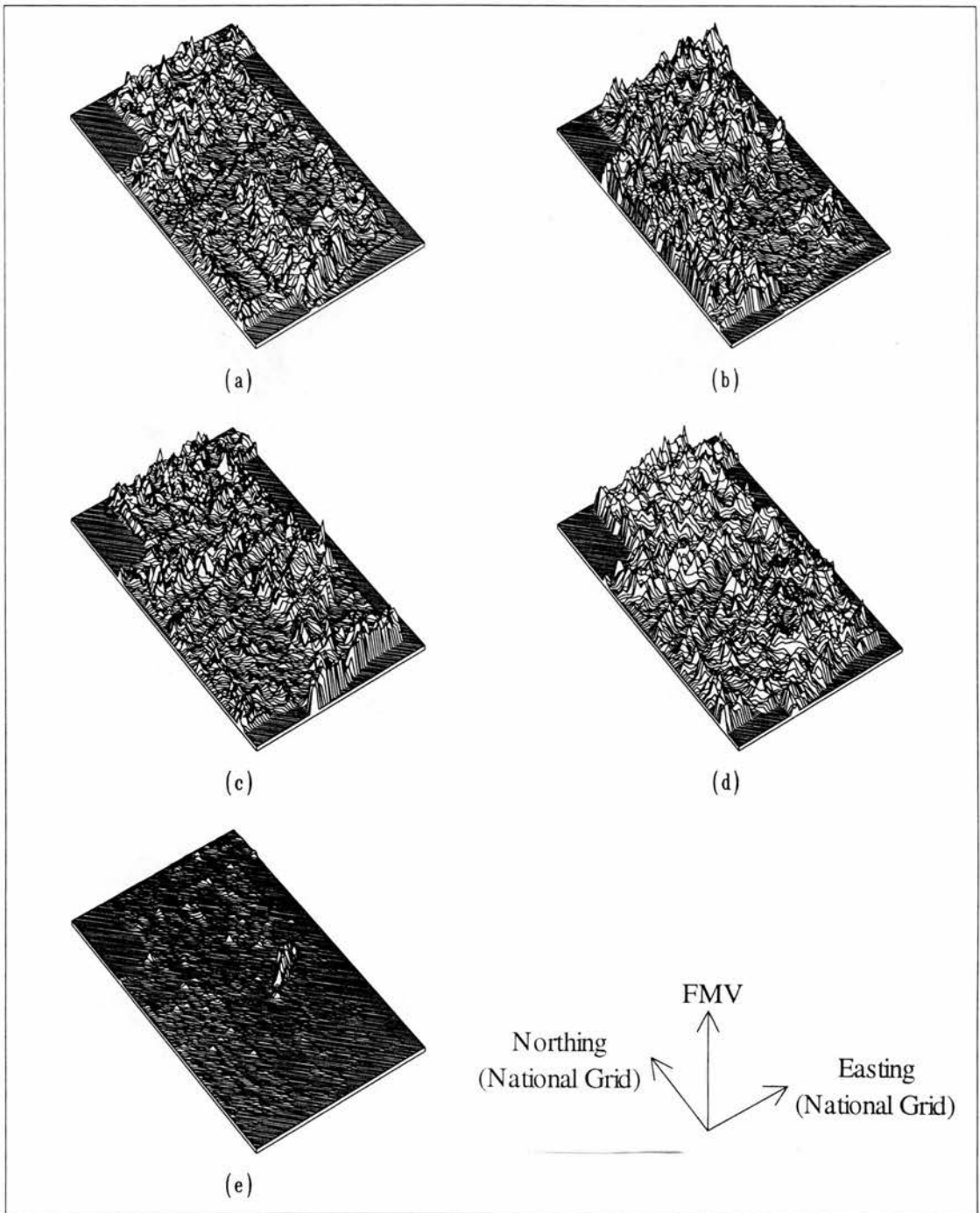


Figure 6.2 Perspective views of fuzzy surfaces of land cover derived from the SPOT HRV data: (a) grass, (b) built, (c) wood, (d) shrub, and (e) water. The peaks on diagrams indicate higher certainty, i.e., increasing likelihood of presence.

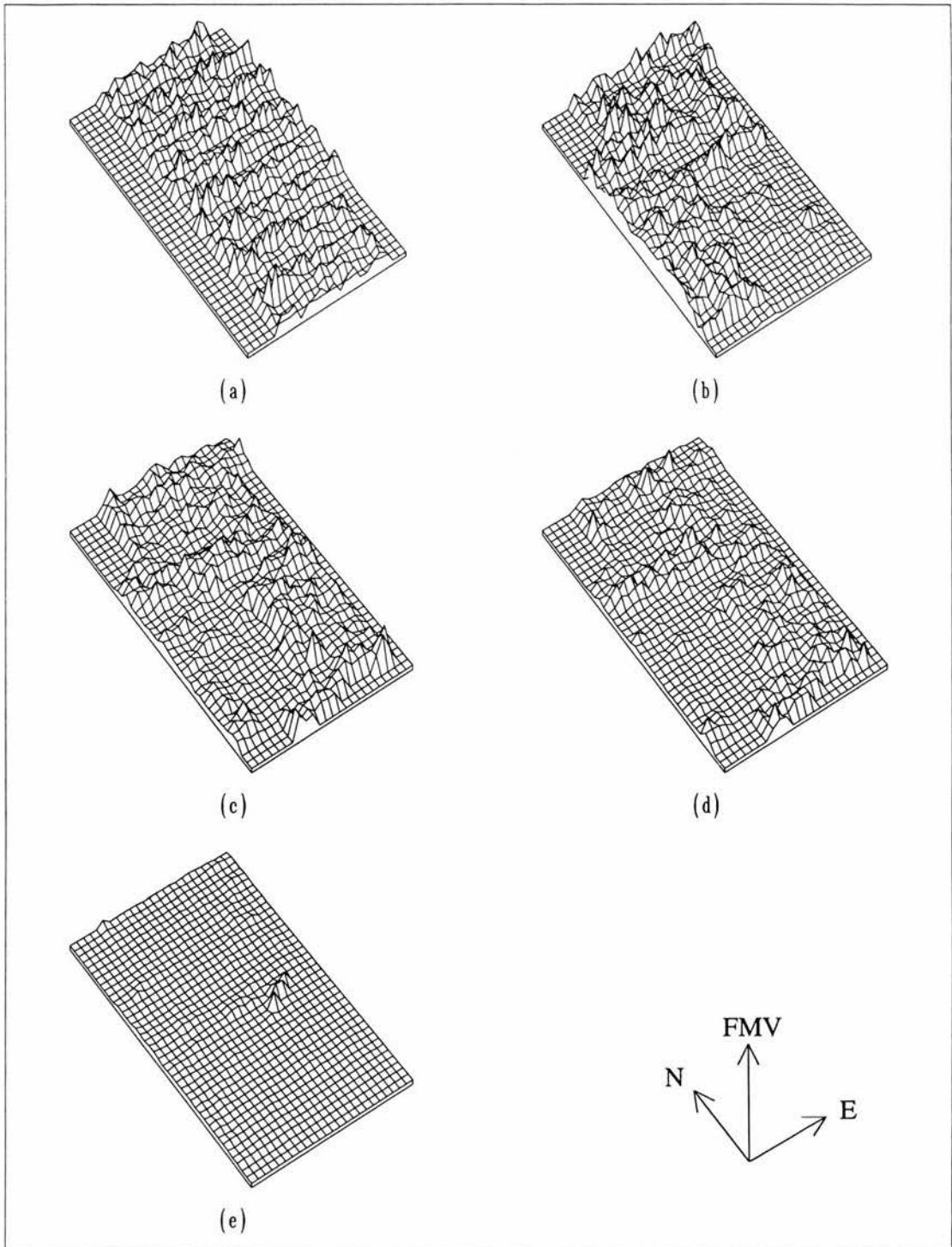
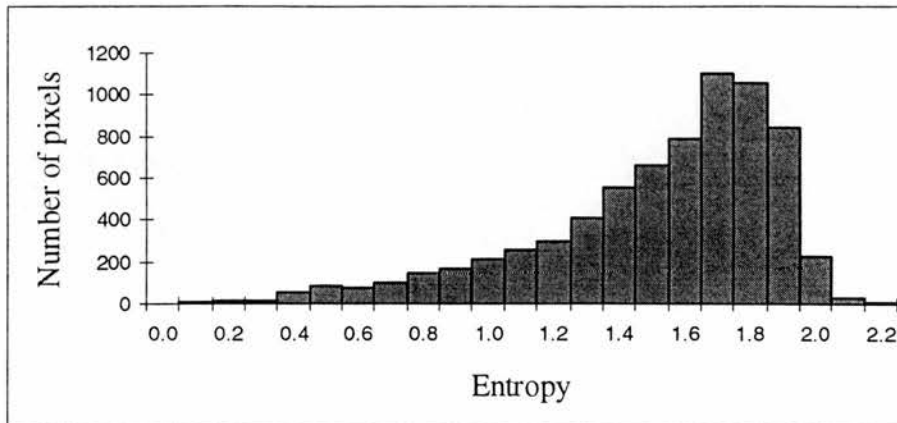
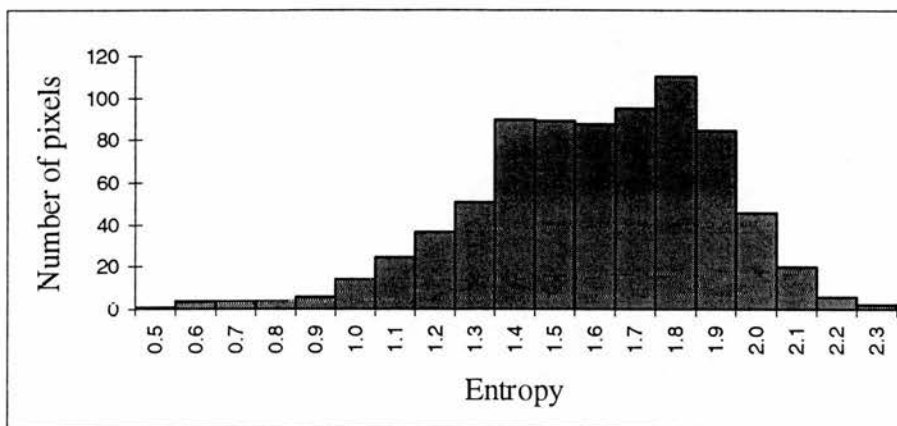


Figure 6.3 Perspective views of fuzzy surfaces of land cover derived from the Landsat TM data: (a) grass, (b) built, (c) wood, (d) shrub, and (e) water. The peaks on diagrams indicate higher certainty, i.e., increasing likelihood of presence.

The derived fuzzy surfaces were evaluated with respect to their entropy measures. This led to the histograms shown in Figure 6.4, where (a) and (b) represent distributions of entropy measures relevant to fuzzy surfaces derived from the SPOT HRV and the Landsat TM data respectively.



(a)



(b)

Figure 6.4 Histograms of entropy measures of fuzzy classification based on: (a) the SPOT HRV data; and (b) the Landsat TM data.

As shown in Figure 6.4, both histograms are negatively skewed towards higher entropy measures. This is particularly evident in Figure 6.4 (a). The means of entropy measures are 1.56 and 1.64 for fuzzy surfaces based on the SPOT HRV and the Landsat TM data respectively. Both values are quite high, indicating significant fuzziness in the fuzzy surfaces. This degree of fuzziness is clearly depicted on the

perspective views of fuzzy surfaces, where FMVs are shown as highly varied and rugged over short distances.

Comparatively, consistently larger proportions of pixels in the Landsat TM data register higher entropy measures, though there is a concentration of high entropy measures around the range between 1.70 and 1.90 on the SPOT HRV data, as shown in Figure 6.4. The result confirms that fuzzy classification based on the SPOT HRV data is slightly better than that based on the Landsat TM data. This is because remote sensing images of a higher spatial resolution should be able to capture finer detail on the ground than those with a lower spatial resolution, thus leading to less mixed pixels.

### **6.2.2 Fuzzy surfaces generated from photogrammetric data**

It was reviewed earlier that aerial photography has been used traditionally to derive useful information about land cover, natural resources and the environment. The information is usually presented on analogue maps or in digital forms, where the former can be easily derived from the latter. These kinds of maps typically portray the geographical reality as certain collections of discrete objects abstracted through the process of photo-interpretation and scribed onto maps as points, lines and polygons (see Figure 5.7).

As seen in Section 5.4, digitising from a reconstituted stereo model was of limited accuracy in terms of positions of objects, and suffered uncertainties due to assumptions of homogeneity which were usually unattainable in the case of urban and suburban land cover mapping. The uncertainties contained in a photogrammetrically digitised land cover map were compounded when such a map was used to check a classified remote sensing data product. This led to decreased agreements between the classified data and the photogrammetric data (the “alleged” ground truth data), and

hence low overall classification accuracies, as seen in Section 5.4.5.

Clearly, the information on the uncertainties involved in photo-interpretations and classifications was thrown away unwisely, and should have been retained if possible. As discussed in Section 4.3.1, sub-pixel component land cover proportions may be used to replace the heterogeneities typically suppressed during photo-interpretation, and thus form a simple type of fuzzy surfaces. Towards this end, photogrammetric data were rasterised at fine cell size (1 metre) via a built-in module in the ERDAS software. Data for the proportions of different sub-pixel components were then calculated by a small original FORTRAN 77 program (Appendix 5). This program performed the summation on a pixel by pixel basis (in accordance with the SPOT HRV and the Landsat TM pixel sizes respectively) with respect to the five land cover classes shown in Figure 5.7. The resulting sub-pixel component proportion data were stored as ARC/INFO GRID data files, and were later compared with the results derived from using indicator kriging, as described below.

As described in Section 4.3.2, indicator kriging is usually performed to update local information with neighbouring information to produce continuous surfaces of *a posteriori* probabilities of certain categories prevailing at a location. In terms of implementation, it has to be decided which geostatistical package to use. The author has accessed geostatistical packages including Geostatistical Software Library (GSLIB) (Deutsch and Journel 1992), GEOPACK (Yates 1995, pers. comm.) and simple kriging in the ARC/INFO package.

While the availability of kriging functionality in ARC/INFO is attractive for its convenience, several workers including Dungan (1995, pers. comm.) and Oliver and Webster (1990) found that simple kriging in ARC/INFO is limited to single spherical model for semivariograms, and its source code is poorly documented. Under such pre-imposed conditions, the users will not know exactly how the software is handling

complex data, what inner procedures the software is following, and with what criteria. Though GEOPACK is considered user-friendly with a menu-driven graphical interface, users are still kept away from more technical aspects such as semivariogram calculation and model fitting. GEOPACK was also found by the author to be incapable of performing kriging interpolation for large data sets. The remaining package GSLIB was therefore used in this empirical study.

GSLIB is a collection of geostatistical routines existing in source code that serves as a starting point for custom programs and application research. The source code provided adheres as closely as possible to the ANSI standard FORTRAN 77 programming language. However, GSLIB is by no means user-friendly, in particular due to the lack of a graphical, interactive semivariogram modelling program. GSLIB requires users to have a substantial understanding of the fundamental geostatistical principles and reasonable computer programming knowledge. Nevertheless, given the fact that flexibility rather than comfort is sometimes more important in the pursuit of research, GSLIB was used in combination with other general purpose statistical packages to perform interactive semivariogram model fitting.

After the selection of a suitable geostatistical package, attention is now given to collecting a set of classified samples or observations from aerial photographs as the first step of indicator kriging. As described in Section 4.3.2, these classified samples are usually taken from locations where the variable under study such as land cover can be considered “pure”, or representative of the named classes, similar to the process of locating training sample in remote sensing classification. Thus, it is assumed that a pure sample (that is, the class found at a point) should belong 100 percent to the named class, thus meaning 0 percent for all other classes. Often, the centres of land cover patches may be taken as pure samples of the corresponding land cover types. This has its supportive reasons in the epsilon band model, which assumes that the further a point is away from the boundary, and thus closer to the centre of a polygon,

the more likely it is to belong to the class of the polygon (Chrisman 1989).

Following this method of identifying pure samples, sets of classified pure samples were identified on the photogrammetric data as shown in Figure 5.7 (234 samples of grassland, 291 samples of built-up areas, 203 samples of wooded land, 76 samples of shrubland, and 11 samples of water bodies). These samples were re-confirmed by referring to the original aerial photographs and, if necessary, checking with the 1:5,000 scale aerial photographs. The resulting pure samples ( $2.5 \times 2.5$  square metres) are shown in Figure 6.5. They were then transformed to a grid coordinate system as required in GSLIB with grid cell size equal to  $2.5 \times 2.5$  square metres.

The data files to be used in GSLIB routines need to be in a format specified in the GSLIB user's guide. Actually, the files can be formatted by referring to the example below (a program is provided in Appendix 6):

- (1) 1st line explaining the data;
- (2) 2nd line specifying the number of fields or columns  $C$  in the data (including  $x, y$  coordinates);
- (3) 3rd line ~  $C+2$  th line specifying the names of individual fields (such as  $x, y$ , variable  $a, b, c$  and so on);
- (4)  $C+3$  th line onwards, each line written as a record of  $C$  fields with spaces separating them.

The experimental semivariograms were calculated by using the **gamv2** subroutine provided in GSLIB. As a rule, the parameter file needs to be properly specified. It may be a good practice to follow the example parameter files listed in the GSLIB user's guide. It is important to calculate the experimental semivariograms for at least three directions to detect any significant anisotropy: omnidirection, north-south, east-west.



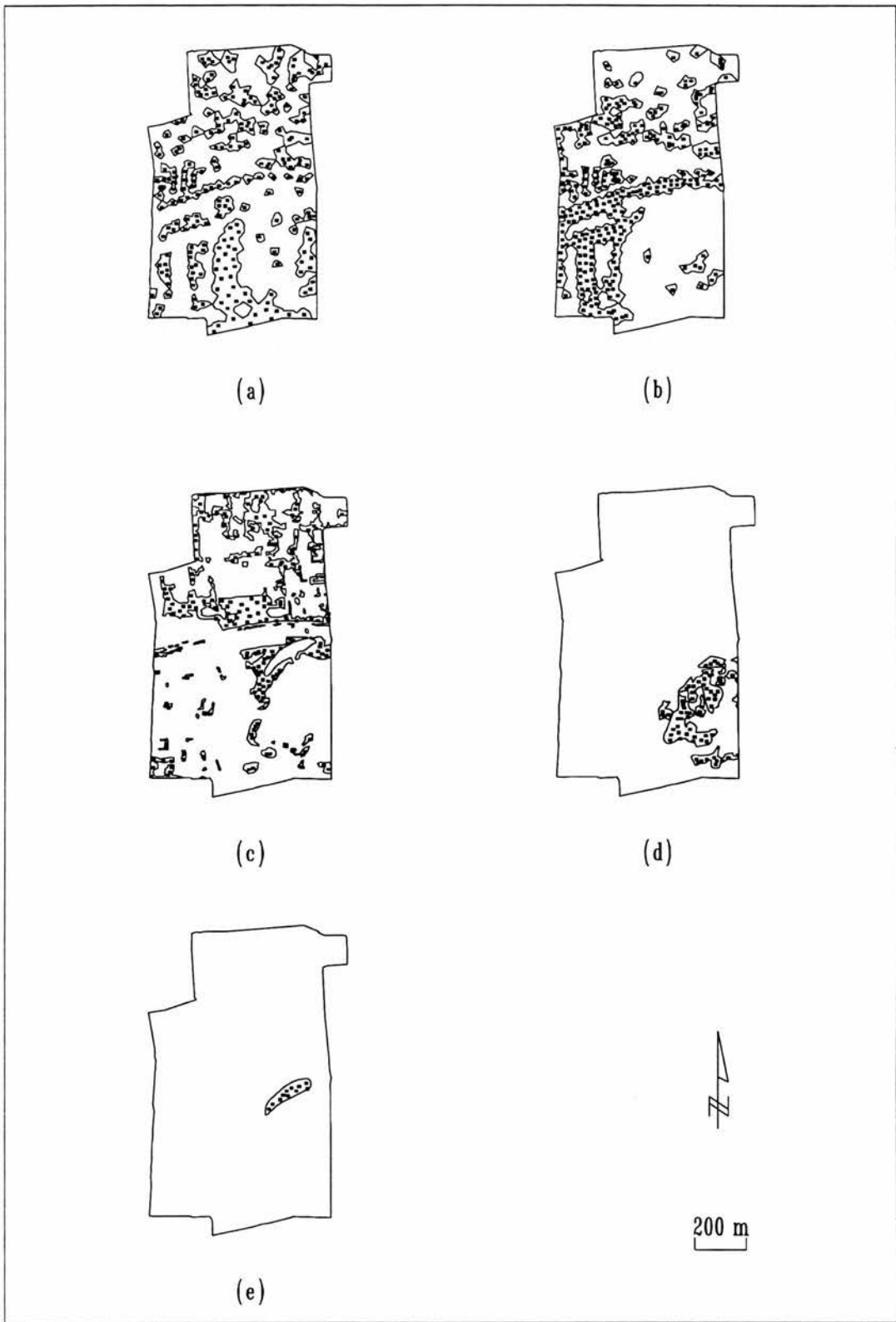


Figure 6.5 Classified pure point samples from the photogrammetric data:

(a) grass, (b) built, (c) wood, (d) shrub, and (e) water.

The outputs were then loaded into SPSS PC software available from the Computing Service at the University of Edinburgh to perform semivariogram model fittings. Prior to the model fitting, the scatter plots of the experimental semivariograms were examined visually in order to identify their upper bounds (known as sills), the ranges (as shown in Figure 4.5) and above all their general structures and forms. The SPSS PC software provides user-friendly, graphical and interactive interfaces to facilitate the linear or non-linear regression analysis in addition to a suite of general statistical functions and modules.

It is known that a spherical or exponential model semivariogram will reach an upper bound known as the sill. Thus, both spherical and exponential models were fitted to experimental semivariograms output from the **gamv2** subroutine. The results are summarised in Table 6.1 below.

Table 6.1 The correlation coefficients ( $r$ ) resulting from semivariogram model fitting (omnidirection semivariograms unless otherwise specified, sample size = 24)

Models	exponential	spherical
grass	0.93	0.95
built	0.92	0.92
wood north-south direction	0.69	0.70
east-west direction	0.72	0.74
shrub	0.89	0.90
water north-south direction	0.66	0.67
east-west direction	0.57	0.58

As indicated in Table 6.1, in general, the spherical model produced a better fit than the exponential model. Thus, the spherical model was used in this study. Figure 6.6 shows both the experimental semivariograms and fitted spherical model semivariograms. In Figure 6.6, scatter points represent the experimental

semivariograms, while solid lines are the fitted spherical model semivariograms:  $r(h)=c*(1.5*(h/a)-0.5*(h/a)^3)$  for  $h<a$ ;  $r(h)=c$  otherwise. The spherical model parameters are listed for different classes in Table 6.2.

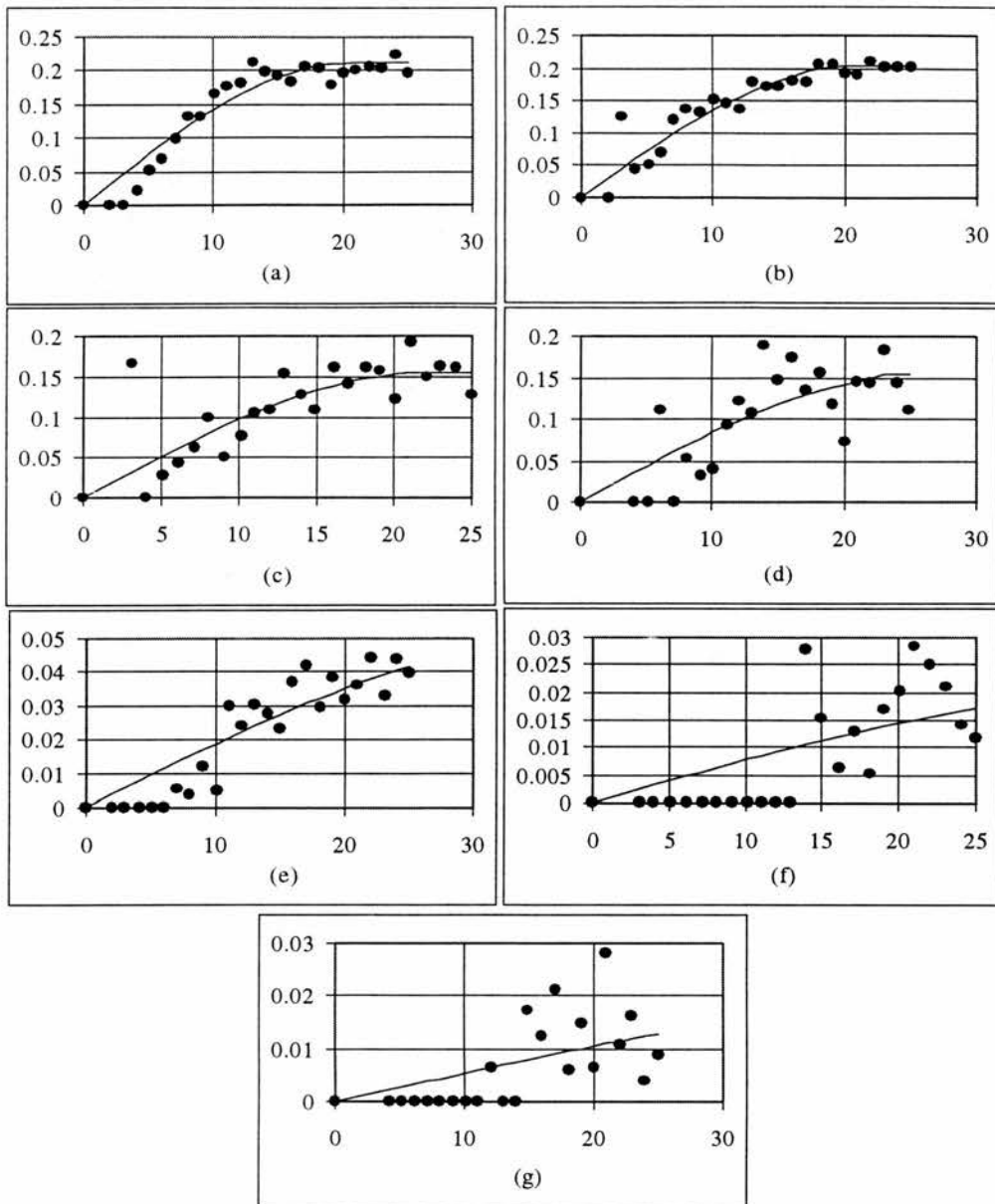


Figure 6.6 Semivariograms (Y-axis is semi-variance; X-axis is distance, in units of 2.5 metres): (a) grass, (b) built, (c) wood in north-south direction, (d) wood in east-west direction, (e) shrub, (f) water in north-south direction, and (g) water in east-west direction.

Table 6.2 The spherical model semivariogram parameters  
(omnidirection semivariograms unless otherwise specified, 1 unit = 2.5 metres)

Classes	Parameter a	Parameter c
grass	20.52	0.211
built	21.03	0.204
wood north-south direction	22.34	0.155
east-west direction	26.42	0.155
shrub	40.00	0.051
water north-south direction	40.00	0.021
east-west direction	57.70	0.021

It can be seen from Figure 6.6 that “grass” and “built” are best defined on their semivariogram, and thus have the highest correlation coefficients as indicated in Table 6.1. “Shrub” is also well defined semivariogram, while “wood” has rather dispersed semivariograms. “Water” is the least well defined by semivariograms, with the lowest correlation coefficients (0.67 and 0.58). While the results are quite reasonable (since all the correlation coefficients are significant at a confidence level of 95 %), there seems to be room for improving the model fitting in order to gain more advantages of using indicator kriging as opposed to other methods such as using sub-pixel component proportions.

Although refined model fitting is not the main concern of this thesis, a comparison of using indicator kriging against using sub-pixel component proportion data to construct fuzzy surfaces from photogrammetric data will be established in the next section.

Prior to running a kriging procedure, an evaluation of semivariogram models needs to be done by a cross-validation procedure. This helps to ensure the reliability of the

results from a kriging implementation and possibly optimise the implementation. The kriging procedure **okb2d** was then run with cell sizes of the output grids equal to the SPOT HRV and the Landsat TM pixel sizes, that is 10 m and 30 m respectively. The total numbers of grid cells were 7113 for the SPOT HRV data and 777 for the Landsat TM data. The kriging outputs were checked to ensure that FMVs sum to 1.0 (100 %) across all the classes for each cell, as described in Section 4.3. The standardised data were transformed to ASCII format files via a written FORTRAN program (in Appendix 7), which were further transferred to ARC/INFO GRID data files in order to make use of the three dimensional display and other surface modelling and analysing functions. The perspective views of fuzzy surfaces based on photogrammetric data for each of the five classes are shown in Figure 6.7, with grid cell size equal to that of the SPOT HRV data (10 m).

As in the case of fuzzy surfaces based on remote sensing data, the fuzzy surfaces as shown in Figure 6.7 can also be interpreted with respect to their fuzziness, or how FMVs are partitioned among all the candidate classes. This was done by using entropy, leading to the histogram shown in Figure 6.8.

As shown in Figure 6.8, the entropy measures turn out to be much smaller than those seen in Figure 6.4. The mean entropy value is 0.68, which is less than half the mean observed in Figure 6.4. The low entropy measures can be visually appreciated from Figure 6.7, where FMVs appear to be rather localised, hence less fuzzy than those observed in Figure 6.4. The results confirm that it is very sensible indeed to employ methods for fuzzy classifications of remotely sensed images such as that shown in the previous section, where an improved algorithm for the supervised fuzzy *c*-means clustering was described and tested with more control over the target classes and thus a reduced degree of fuzziness. Such a strategy helps to make fuzzy classified remotely sensed data more compatible with fuzzy classified photogrammetric data.

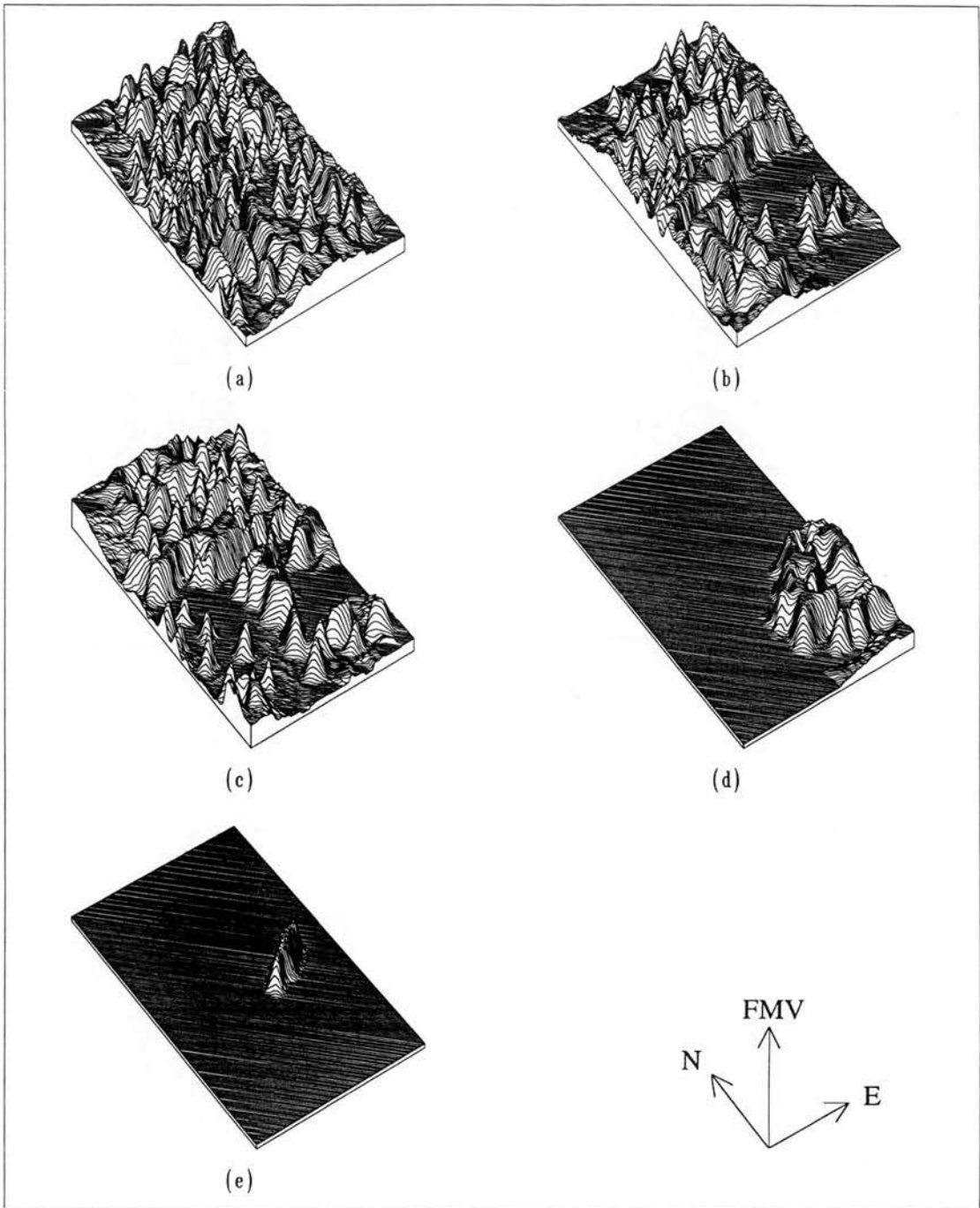


Figure 6.7 Perspective views of fuzzy surfaces of land cover derived from the photogrammetric data: (a) grass, (b) built, (c) wood, (d) shrub, and (e) water. The peaks on diagrams indicate higher certainty, i.e., increasing likelihood of presence.

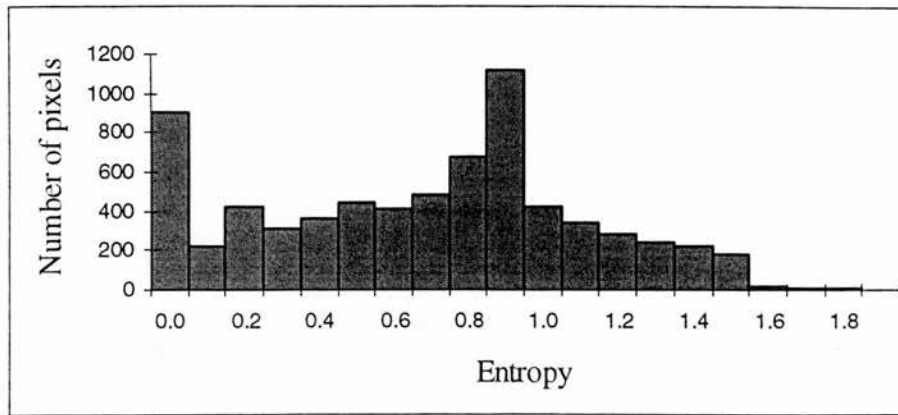


Figure 6.8 Histogram of entropy measures for fuzzy surfaces generated via indicator kriging

### 6.3 Analysing fuzzy surfaces

As established in the previous section, fuzzy surfaces have been derived from remote sensing by using fuzzy classification, and from photogrammetric data by using indicator kriging. For photogrammetric data, fuzzy surfaces were also derived from sub-pixel component land cover proportions. In some sense, these fuzzy surfaces stand as different classification products as opposed to conventional categorical maps containing a single class at any location. As in the case of comparing and analysing conventional categorical maps, fuzzy surfaces may also be compared and analysed, but using different methods. This procedure is developed below.

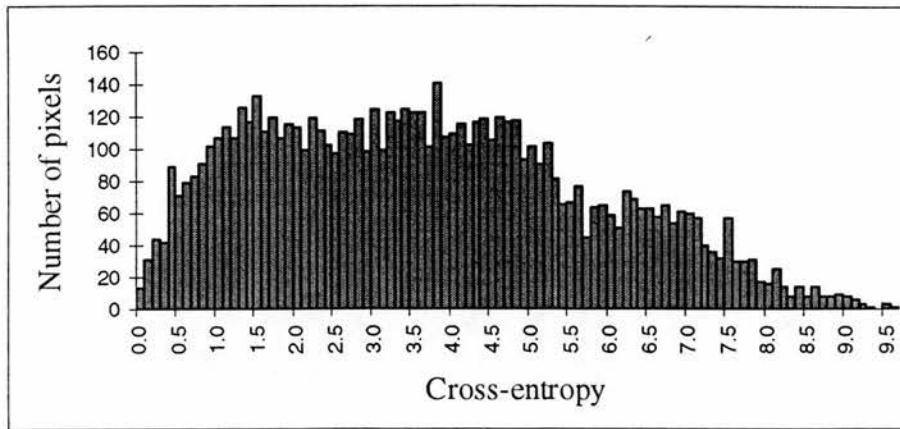
As shown in the previous section, the fuzziness versus discreteness of underlying fuzzy surfaces was evaluated by using entropy measures. The higher the entropy measure, the more “fuzzy” the specific set of fuzzy surfaces, and *vice versa*. According to entropy measures, fuzzy surfaces based on remote sensing data were seen to be highly varied, while those based on photogrammetric data were less varied.

As discussed in Section 4.4, entropy is a measure of the partition of fuzzy or probabilistic values among candidate classes. The high entropy measures for fuzzy surfaces based on remote sensing data imply that pixels are significantly mixed. Thus, it is difficult to interpret the accuracy of a fuzzy classification based on entropy measures. This is because pixels with a low or high entropy measure may still be accurately representing the ground situation.

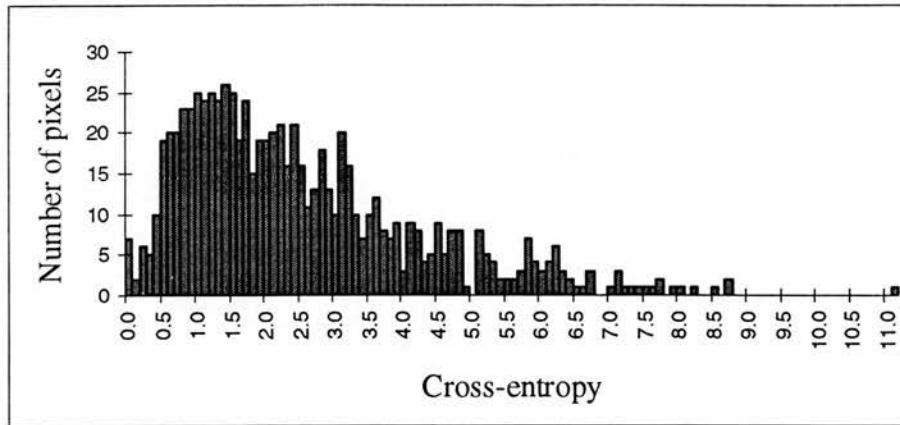
Cross-entropy, on the other hand, allows the closeness of a fuzzy classification to a fuzzy reference data set to be measured (Foody 1995a). For the purpose of this study, suppose fuzzy surfaces generated from photogrammetric data are to be used as fuzzy reference data for fuzzy classified remote sensing data. As mentioned previously, fuzzy reference data were derived using two different methods: (1) sub-pixel component land cover proportions, and (2) indicator kriging. Thus, the cross-entropy measures were calculated (1) between fuzzy classified remote sensing data and fuzzy reference data by using sub-pixel component land cover proportions, and (2) between fuzzy classified remote sensing data and fuzzy reference data derived using indicator kriging. The results are presented as Figure 6.9 and 6.10, respectively.

As shown in Figure 6.9, when using sub-pixel component proportions as fuzzy reference data, the means of cross-entropy measures for the SPOT HRV and the Landsat TM data are 3.67 and 2.49 respectively. Means of cross-entropy measures are reduced to 2.67 for the SPOT HRV data and 1.89 for the Landsat TM data when the fuzzy reference data are derived from indicator kriging, as shown in Figure 6.10.





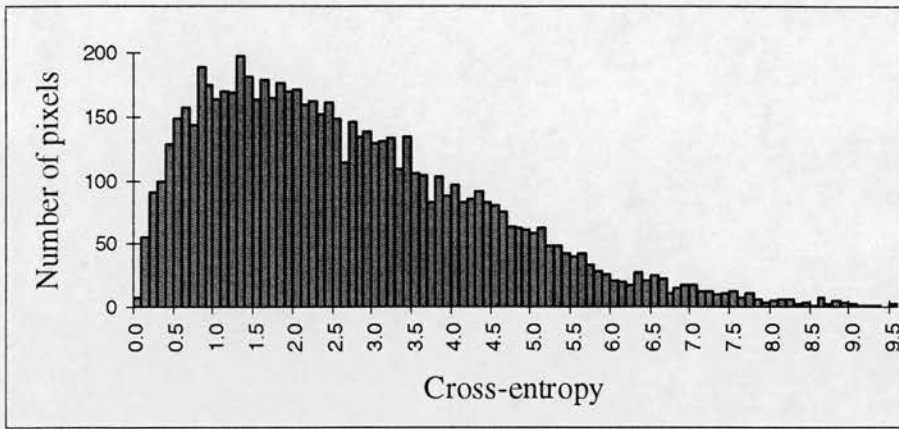
(a)



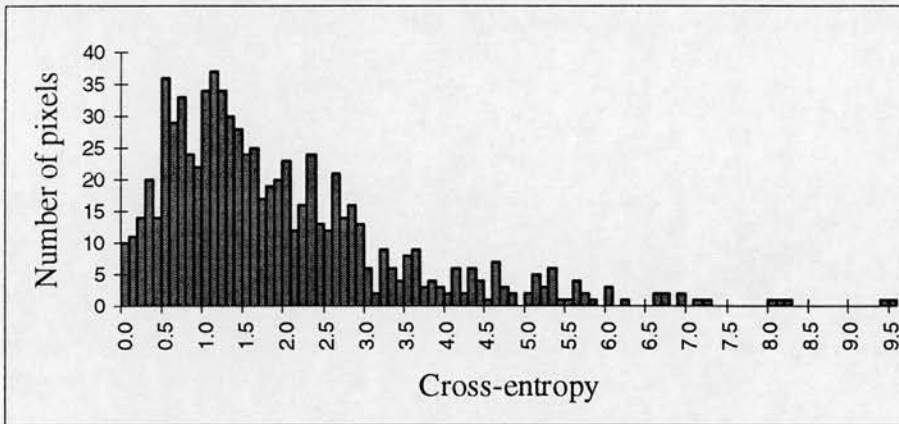
(b)

Figure 6.9 Histograms for cross-entropy measures between fuzzy classified remote sensing data and fuzzy reference data derived by using sub-pixel component proportions:

- (a) the SPOT HRV data, and
- (b) the Landsat TM data.



(a)



(b)

Figure 6.10 Histograms for cross-entropy measures between fuzzy classified remote sensing data and fuzzy reference data derived using indicator kriging: (a) the SPOT HRV data, and (b) the Landsat TM data.

Clearly, reduced cross-entropy measures for both the SPOT HRV and the Landsat TM data, as shown in Figure 6.10, suggest that fuzzy reference data derived by using indicator kriging lead to better closeness with fuzzy classified remote sensing data. In other words, results confirm that indicator kriging is more suitable as a means to generate fuzzy reference data from photogrammetric data than simply using sub-pixel component land cover proportion data as fuzzy reference data. One of the reasons why using sub-pixel component land cover proportions is less suitable is that the sub-pixel component proportions data are used on the assumption that the component

land covers can be accurately represented as mixtures of discrete polygons. This assumption is unlikely to be true of areas with significant fuzziness.

An additional line of evidence for checking the superiority of using indicator kriging over simply using sub-pixel component land cover proportions to construct fuzzy reference data is established below by calculating correlation coefficients between the fuzzy classified remote sensing data and the two different fuzzy reference data sets. A merit of calculating correlation coefficients is that correlation coefficients for individual classes can be calculated. Thus, a calculation of correlation coefficients may provide extra insights into the relative suitability of using sub-pixel component proportion data versus using indicator kriging to generate fuzzy reference data at the level of individual classes.

The correlation coefficients were calculated for each class using, firstly, fuzzy classified remote sensing data with fuzzy reference data derived by using sub-pixel component land cover proportions and, secondly, fuzzy classified remote sensing data with fuzzy reference data generated from indicator kriging. Results from correlation analyses are tabulated in Table 6.3, where fuzzy reference data are derived by: (I) sub-pixel component cover proportions, and (II) indicator kriging.

In terms of results, apart from “grass” for the Landsat TM data, all the correlation coefficients, as listed in Table 6.3, are significant at a confidence level of 5%. Also, as shown in Table 6.3, except for “water” in the case of the SPOT HRV data, consistently higher correlation coefficients were obtained when indicator kriging was used to derive fuzzy reference data for both the SPOT HRV and the Landsat TM data, although the magnitudes of gains are not quite the same.

Table 6.3 Correlation coefficient ( $r$ ) values between fuzzy classified data and their fuzzy reference data.

Classes	SPOT HRV data		Landsat TM data	
	7113 pixels		777 pixels	
	I	II	I	II
grass	0.43	0.49	-0.04	0.06
built	0.58	0.70	0.43	0.68
wood	0.52	0.52	0.28	0.48
shrub	0.25	0.30	0.28	0.34
water	0.72	0.68	0.10	0.45

The results reconfirm that, in general, using indicator kriging is more suitable than using sub-pixel component land cover proportions to derive fuzzy reference data, at least for similar purposes to those as in this study. In particular, the negative correlation coefficient obtained with the Landsat TM data when using sub-pixel component land cover proportions as fuzzy reference data highlights the potentially meaningless aspects of using this method. The only “abnormal” result for “water” had its root in the relatively poorly defined semivariogram as shown in Figure 6.6.

#### 6.4 Deriving categorical maps from fuzzy surfaces

As has been seen in Section 6.2, fuzzy surfaces were generated from remote sensing by a supervised fuzzy  $c$ -means clustering, and from photogrammetric data by using indicator kriging. These fuzzy surfaces are multidimensional and continuous, and so can be reduced down to simpler, discrete classifications to produce categorical maps as shown below.

As described in Section 4.4, categorical maps may be produced from fuzzy surfaces by using two interconnected operations: “maximising” and “slicing”. For the

classification of remote sensing images, several techniques are available, of which maximum likelihood classification is one of the most widely used. By using a maximum likelihood classification, pixels are assigned to classes, to which they have the maximum probability or likelihood of belonging, measured by specific class membership functions.

When the probabilities of a pixel or grid cell belonging to the set of candidate classes are known as in the case of fuzzy surface representations, the maximum likelihood classification becomes simply a maximising operation applied to the complete set of fuzzy surfaces built per category or class. The output will consist of the pixels or cells labelled as classes with maximum FMVs, which are accordingly assembled to form a corresponding map depicting the certainty levels for the named class at the level of individual pixels. This process is, to some extent, similar to the practice in a conventional classification of remote sensing images, where a classified data set is accompanied by a probability data layer, so that end users can place a specific confidence for each classified pixel.

The processing as outlined above was programmed into AML procedures within ARC/INFO GRID and ARCPLOT modules. The “hardened” categorical land cover maps were produced from the fuzzy surfaces generated in Section 6.2 (i.e. those shown in Figures 6.2, 6.3 and 6.7). Meanwhile, the maximum FMVs or probabilities constitute the by-products recording the underlying certainty levels for the output categorical maps. The resulting categorical maps are shown in Figure 6.11 ((a) for the SPOT HRV data, (b) for the Landsat TM data, and (c) for the photogrammetric data), where the categorical maps are draped onto the underlying certainty surfaces depicting the spatially varying certainty levels (FMVs or probabilities) with which a location belongs to its labelled class.

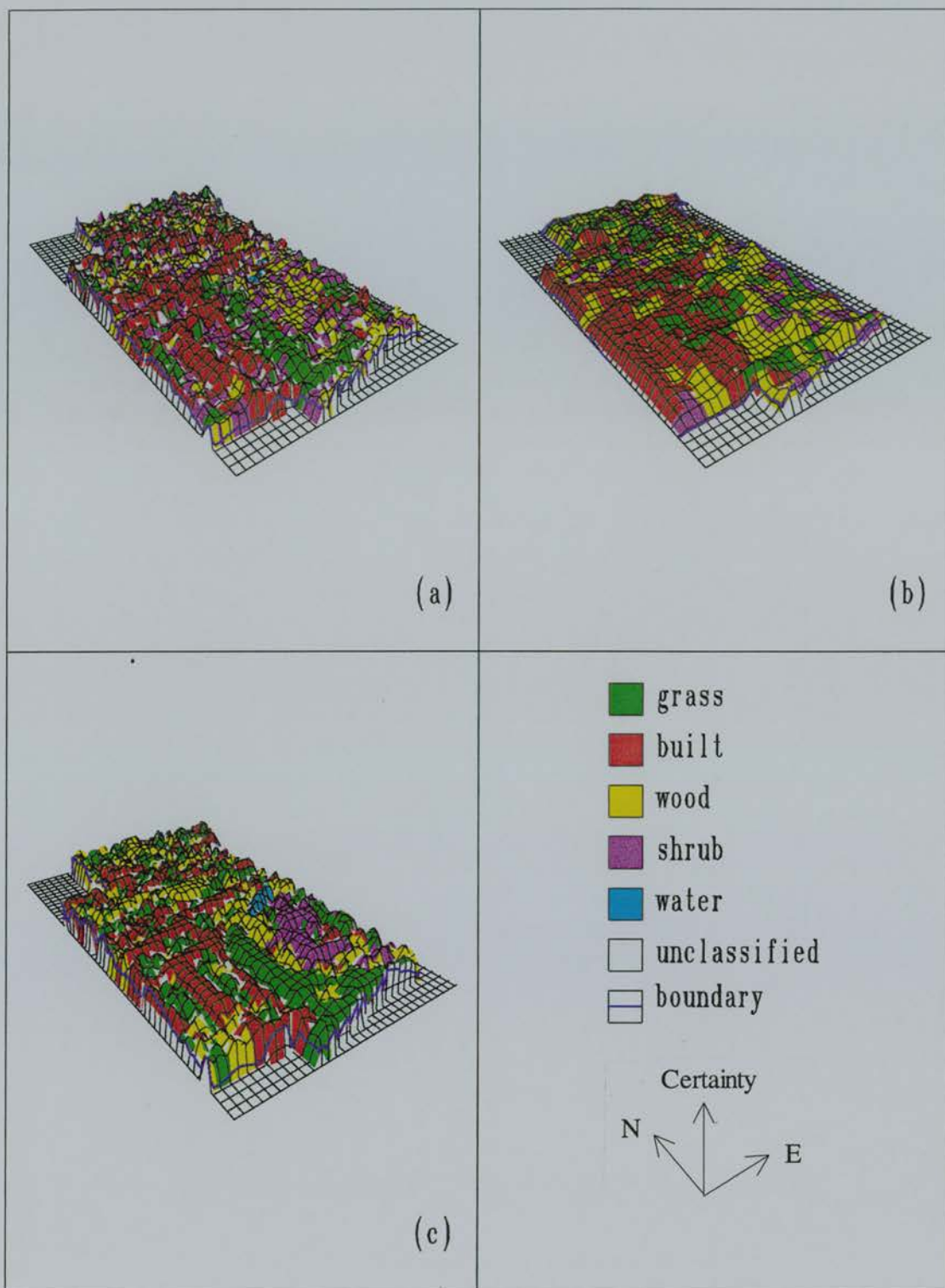


Figure 6.11 Categorical maps derived from the maximisation operation and draped onto their underlying certainty surfaces: (a) SPOT HRV data, (b) Landsat TM data, and (c) photogrammetric data.

The higher parts on these certainty surfaces indicate better accuracies, with which the categorical maps are associated, or *vice versa*. There are two main advantages held by such a surface-based approach to uncertainties. Firstly, the categorical maps presented in this way convey a strong sense of spatial variations of uncertainties inherent to spatial data. Secondly, spatially varying accuracies are readily available from consulting the underlying certainty surfaces (shown in Figure 6.11). This is seen to be lacking in traditional categorical maps, which are produced in the form of discrete polygons, each assigned a single class, with a spatially invariant level of accuracy applicable to, at best, individual classes.

Furthermore, based on the categorical maps presented in Figure 6.11, it is possible to perform slicing to derived another set of categorical maps, as described in Section 4.4. As an example, the certainty surface as shown in (c) of Figure 6.11 was sliced at a succession of thresholds: 40%, 55%, 70% and 85%, which led to categorical maps shown in (a), (b), (c) and (d) of Figure 6.12 below, respectively. The areas labelled as “unclassified” are those parts of the polygons where the FMVs for the classes with maximum FMVs (i.e., dominant classes) fall below the specified thresholds, while the shaded areas are where the FMVs for the classes with maximum FMVs are not less than the chosen thresholds.

As shown in Figure 6.12, when higher thresholds were applied, less pixels were classified, but these classified pixels were identified with greater certainty as belonging to their named classes. The change of spatial pattern observed when increasing or decreasing thresholds can be visualised in a sequence of categorical maps as those shown in Figure 6.12, serving to help users to chosen a suitable threshold. In practical terms, the set of categorical maps shown in Figure 6.12 can be considered as produced with: (a) relatively poor, (b) moderate, (c) good, and (d) outstanding levels of accuracy.

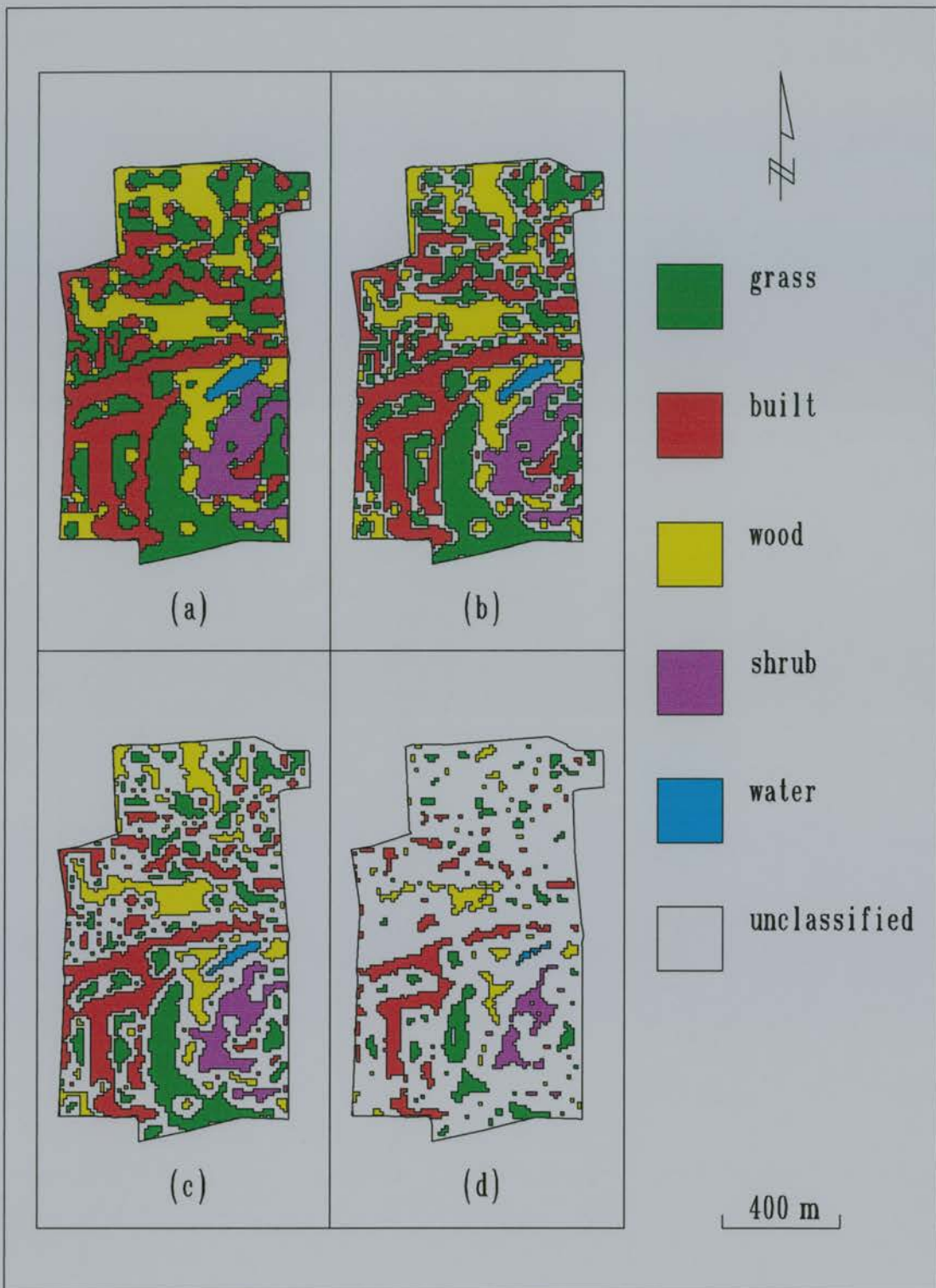


Figure 6.12 Categorical maps derived from the slicing operation based on the map shown in Figure 6.11 (c) at a succession of thresholds: (a) 40 %, (b) 55 %, (c) 70 %, and (d) 85 %.



The unclassified areas shown in Figure 6.12 are where relatively larger uncertainties exist, because their FMVs of belonging to the labelled classes fall below the prescribed threshold. Though irregular in shaped, these uncertainty zones can actually be viewed as some variants of the epsilon error band models, which may further be generalised to render the usual smooth-looking epsilon error band models. This confirms that uncertainties are better approached from fields rather than objects. An epsilon band model, as discussed in Chapter 3, is often used in an object-based model as a simplified representation of the otherwise complex pattern of uncertainty occurring in line objects (Pullar and Beard 1989). When the underlying phenomena are themselves complex, such as land cover and soil types, the derivations of epsilon band models would be difficult as no “ground truth” exists to be used as reference (Edwards and Lowell 1996). Now with a field-based method, it is possible to derive epsilon band models to represent the uncertainties relevant to polygon boundaries without major difficulties, as shown in Figure 6.12.

In summary, it has been seen so far that a variety of categorical maps might be produced by using maximising and slicing operators. A maximising operation generates a categorical map, in which each pixel is assigned a single and dominant class label, while its underlying certainty surface provides useful information on the spatially varying levels of accuracies for individual pixels. A slicing operation produces a categorical map, from which an epsilon band model might further be derived.

For the types of categorical maps derived from the maximising operation, effective visualisation of the uncertainties associated with the categorical maps was obtained by referring the categorical maps to the underlying certainty surface. On the other hand, for the types of categorical maps derived by the slicing operation, the changing pattern of the accuracies versus uncertainties was visualised by a sequence of categorical maps with increasing accuracies in the classified areas, but with widening

uncertainty zones represented by the unclassified areas. The results confirm that modelling uncertainties from fields rather than objects is theoretically sound and practically feasible.

## **6.5 Evaluating the accuracies of fuzzy classified remote sensing data**

As a variety of categorical maps was produced from fuzzy surfaces by using the maximising and slicing operations, it is possible to consider the ways by which the accuracies of these different categorical maps may be assessed. Chapter 4 has described both hard comparisons and soft comparisons that are useful for analysing fuzzy classified data. Thus, this section will show the comparisons between:

- (1) the categorical data derived by using the maximising operation (hard),
- (2) the categorical data derived by using the slicing operation (hard), and
- (3) the fuzzy classified data (soft).

For the purpose of this thesis, the categorical maps derived by using the maximising operator, as shown in Figure 6.11, are termed “hardened” versions of fuzzy classified data. Therefore, a comparison may be made between a hardened fuzzy classified remote sensing image and a hardened fuzzy reference data set (i.e., fuzzy classified photogrammetric data). As fuzzy reference data were derived by using: (1) sub-pixel component land cover proportion data, and (2) indicator kriging, it is possible to carry out the comparison mentioned above by checking the hardened fuzzy classified SPOT HRV and Landsat TM data against the two different sets of hardened fuzzy reference data. Because the hardened reference data used in the former kind of comparisons are, in fact, the original hard photogrammetric data shown in Figure 5.7, these two kinds of comparisons provide extra evidence that fuzzy reference data derived from indicator kriging are the more suitable for assessing the hardened fuzzy classified remote sensing images. Results are shown in Tables 6.4 and 6.5 for the former kind of comparisons, and in Tables 6.6 and 6.7 for the latter kind of comparisons.

Table 6.4 Error matrix of hardened fuzzy classified SPOT HRV data, using the original hard photogrammetric data as reference data

		Reference data					row total	row marginal
		grass	built	wood	shrub	water		
Classified data	grass	890	118	36	17	0	1061	0.15
	built	1357	1097	165	6	48	2673	0.38
	wood	253	26	631	139	14	1063	0.15
	shrub	1160	155	680	294	3	2292	0.32
	water	0	0	0	0	24	24	0.00
column total		3660	1396	1512	456	89	7113	
column marginal		0.51	0.20	0.21	0.06	0.01		

Overall agreement = 41%, chance agreement = 20%, Kappa = 26%

Table 6.5 Error matrix of hardened fuzzy classified Landsat TM data, using the original hard photogrammetric data as reference data

		Reference data					row total	row marginal
		grass	built	wood	shrub	water		
Classified data	grass	73	14	30	8	1	126	0.16
	built	195	87	42	4	0	328	0.42
	wood	117	20	82	29	10	258	0.33
	shrub	35	3	10	16	0	64	0.08
	water	1	0	0	0	0	1	0.00
column total		421	124	164	57	11	777	
column marginal		0.54	0.16	0.21	0.07	0.01		

Overall agreement = 33%, chance agreement = 23%, Kappa = 13%

Table 6.6 Error matrix of hardened fuzzy classified SPOT HRV data, using the hardened fuzzy reference data derived from indicator kriging as the reference data

		Reference data					row total	row marginal
		grass	built	wood	shrub	water		
Classified data	grass	760	219	58	24	0	1061	0.15
	built	663	1721	230	16	43	2673	0.38
	wood	170	62	662	146	23	1063	0.15
	shrub	885	224	794	384	5	2292	0.32
	water	0	0	0	0	24	24	0.00
column total		2478	2226	1744	570	95	7113	
column marginal		0.35	0.31	0.25	0.08	0.01		

Overall agreement = 50%, chance agreement = 23%, Kappa = 35%

Table 6.7 Error matrix of hardened fuzzy classified Landsat TM data, using the hardened fuzzy reference data derived from indicator kriging as the reference data

		Reference data					row total	row marginal
		grass	built	wood	shrub	water		
Classified data	grass	65	25	27	7	2	126	0.16
	built	96	195	34	3	0	328	0.42
	wood	85	17	110	36	10	258	0.33
	shrub	22	2	21	19	0	64	0.08
	water	0	0	0	0	1	1	0.00
column total		268	239	192	65	13	777	
column marginal		0.34	0.31	0.25	0.08	0.02		

Overall agreement = 50%, chance agreement = 18%, Kappa = 39%

For the SPOT HRV data, as shown in Table 6.4, there are serious confusions between “grass” and “built”, between “grass” and “shrub”, and between “wood” and “shrub”. For the Landsat TM data, as shown in Table 6.5, the confusions observed in the SPOT HRV data are reinforced. Besides, “grass” is seriously confused with “wood” for the Landsat TM data. Similar confusions were found in Section 5.4.5, where both the SPOT HRV and the Landsat TM images were classified by using conventional maximum likelihood classification.

Those confusions observed in Tables 6.4 and 6.5 are greatly reduced when the hardened fuzzy reference data derived from indicator kriging were used as the reference data, as shown in Tables 6.6 and 6.7. In this case, consistently better classification accuracies are obtained for both the SPOT HRV data and the Landsat TM data when using the original hard photogrammetric data as the reference data. This is particularly significant for the Landsat TM data, whose Kappa coefficient is increased two-fold. This is because the Landsat TM data, with a relatively coarse spatial resolution, will increase the number of mixed pixels, as verified by Figure 6.4, where the Landsat TM data are seen to be less accurately classified (with a higher measure of entropy) than the SPOT HRV data. This increased number of mixed pixels will, in turn, require the assumed reference data to be able to adapt to the significant fuzziness present. Therefore, it is further confirmed that indicator kriging is more suitable than using sub-pixel component land cover proportion data, especially in situations where fuzziness is a feature of not only remotely sensed images but also the adopted reference data.

As described in Section 6.3, categorical maps might also be derived from the slicing operation. This process works by selecting, on a quantitative basis, “pure” pixels according to a chosen threshold. As only those pixels considered pure are classified, increased agreements should be obtained by comparing pure pixels with their reference data.

Towards this end, fuzzy classified remote sensing images and fuzzy reference data derived by using indicator kriging are sliced, using the procedure described previously in Section 6.3. Specifically, the SPOT HRV and the Landsat TM data are sliced at thresholds of 54% and 41%, respectively, while their corresponding fuzzy reference data are sliced at thresholds of 77% and 72%. The “sliced” data are then used to construct error matrices, as shown as Tables 6.8 and 6.9 below.

As expected, in general, misclassifications observed in Tables 6.6 and 6.7 are further reduced, as shown in Tables 6.8 and 6.9. Overall classification accuracies and Kappa coefficients of agreement are increased by nearly half for both the SPOT HRV data and the Landsat TM data. In particular, only confusions between “grass” and “shrub”, and between “wood” and “shrub” remain based on comparisons between sliced fuzzy classified remote sensing images and sliced fuzzy reference data.

Table 6.8 Error matrix of sliced fuzzy classified SPOT HRV data, using sliced fuzzy reference data as the reference data

		Reference data					row total	row marginal
		grass	built	wood	shrub	water		
Classified data	grass	222	13	3	0	0	238	0.13
	built	129	886	8	0	0	1023	0.56
	wood	42	0	173	19	0	234	0.13
	shrub	107	7	136	81	0	331	0.18
	water	0	0	0	0	11	11	0.01
column total		500	906	320	100	11	1837	
column marginal		0.27	0.49	0.17	0.05	0.01		

Overall agreement = 75%, chance agreement = 34%, Kappa = 62%

Table 6.9 Error matrix of sliced fuzzy classified Landsat TM data, using sliced fuzzy reference data as the reference data

		Reference data					row total	row marginal
		grass	built	wood	shrub	water		
Classified data	grass	20	9	5	3	1	38	0.18
	built	9	89	1	0	0	99	0.46
	wood	7	0	43	12	1	63	0.30
	shrub	3	0	7	3	0	13	0.06
	water	0	0	0	0	0	0	0.00
column total		39	98	56	18	2	213	
column marginal		0.18	0.46	0.26	0.08	0.01		

Overall agreement = 73%, chance agreement = 33%, Kappa = 59%

Clearly, the classified accuracies will be increased further by setting higher thresholds for slicing. This is not going to be elaborated in this section. Attention is given to the “soft” comparison described in Section 4.4, in which not only the most likely classes but also the second likely classes are compared between a fuzzy classified remote sensing image and a fuzzy reference data set.

By a soft comparison, a pixel is considered to be correctly classified if the most likely class or the second most likely class agrees with the most likely class or the second most likely class labelled on its reference data. Soft comparisons accommodate the fuzziness by taking account of both the most likely and the second most likely classes. A soft comparison is deemed to be useful in situations where pixels are highly mixed. Therefore, both the most likely classes and second most likely classes are important. Results are next presented of the soft comparisons between fuzzy classified remote sensing images and fuzzy reference data derived by using indicator kriging, applying the condition described in Section 4.4, as shown in Tables 6.10 and 6.11.

Table 6.10 Error matrix based on the soft comparison between fuzzy classified SPOT HRV data and fuzzy reference data derived from indicator kriging

		Reference data					row total	row marginal
		grass	built	wood	shrub	water		
Classified data	grass	1635	10	14	0	0	1659	0.23
	built	53	2577	40	1	14	2685	0.38
	wood	66	21	1652	2	6	1747	0.24
	shrub	162	58	64	702	0	986	0.14
	water	0	0	0	0	36	36	0.01
column total		1916	2666	1770	705	56	7113	
column marginal		0.27	0.37	0.25	0.10	0.01		

Overall agreement = 93%, chance agreement = 28%, Kappa = 90%

Table 6.11 Error matrix based on the soft comparison between fuzzy classified Landsat TM data and fuzzy reference data derived from indicator kriging

		Reference data					row total	row marginal
		grass	built	wood	shrub	water		
Classified data	grass	131	3	1	2	0	137	0.18
	built	0	311	1	1	0	313	0.40
	wood	23	6	232	6	3	270	0.35
	shrub	5	2	0	48	0	55	0.07
	water	0	0	0	0	2	2	0.00
column total		159	322	234	57	5	777	
column marginal		0.20	0.41	0.30	0.07	0.01		

Overall agreement = 93%, chance agreement = 31%, Kappa = 90%



Significant increases in overall classification accuracies and kappa coefficients of agreement are observed in Tables 6.10 and 6.11. Classes are largely correctly classified. The great gains in the agreements explain, in another perspective, that a considerable amount of “misclassification” occurs due to the fuzziness and complexity existing in the real world. Thus, the second most likely classes count so much as not to be ignored, especially in situations where pixels are heavily mixed and fuzziness is intrinsic to both the remotely sensed data and the reference data, which is traditionally assumed to be absolutely discrete.

In summary, when fuzzy classified data are hardened, it is possible to assess their classification accuracies by using standard error matrix methods. While hardened fuzzy classified data are assessed in a way similar to that used for conventional categorical maps, fuzzy classified data can be compared with fuzzy reference data by either referring to those pixels considered pure in terms of fuzzy and probabilistic measures or applying soft methods, in which both the most likely classes and the second most likely classes are checked.

In order to aid the understanding of the results presented so far, a summary is given of the overall classification accuracies and the Kappa coefficients of agreement relevant to the different methods of assessing classification accuracies, as shown in Table 6.12. Increasing classification accuracies are obtained from using hardened fuzzy classified data and fuzzy reference data, through sliced fuzzy classified data and fuzzy reference data, to the soft comparisons of fuzzy classified data with fuzzy reference data. For the hardened fuzzy classified data, the hardened fuzzy reference data derived from indicator kriging were found more suitable to be used as the reference data than using hard photogrammetric data as the reference data.

Table 6.12 Summary results of the classification accuracies for the categorical maps derived in Section 6.4 using different methods of comparisons

Types of comparisons	Numbers of pixels	Overall classification accuracies	Kappa coefficients
1. hardened fuzzy classified remote sensing data			
a. using hard photogrammetric data as the reference data			
• SPOT HRV data	7113	41%	26%
• Landsat TM data	777	33%	13%
b. using hardened fuzzy reference data derived from indicator kriging as the reference data			
• SPOT HRV data	7113	50%	35%
• Landsat TM data	777	50%	39%
2. sliced fuzzy classified remote sensing data using sliced fuzzy reference data derived from indicator kriging as the reference data			
• SPOT HRV data	1837	75%	62%
• Landsat TM data	213	73%	59%
3. soft comparisons between fuzzy classified remote sensing data and fuzzy reference data derived from indicator kriging			
• SPOT HRV data	7113	93%	90%
• Landsat TM data	777	93%	90%

This fact confirms that a substantial amount of misclassification is due to much fuzziness (hence many mixed pixels) inherent in both the remotely sensed data and the reference data, which used to be assumed to have absolute accuracy. Therefore, methods which can adapt to this fuzziness will achieve better agreements between remotely sensed data and the reference data. In other words, fuzziness maintained by the fuzzy surface approach promoted in this thesis can be used to advantage in situations where the underlying phenomenon is essentially fuzzy.

## 6.6 Discussion

This chapter began with the generation of fuzzy surfaces from remotely sensed images by using an improved version of the fuzzy *c*-means clustering algorithm in a supervised mode, from photogrammetric data by using indicator kriging and, for comparisons, by using sub-pixel component land cover proportion data. While fuzzy surfaces were obtained from remotely sensed images by using a relatively straightforward program, fuzzy surfaces were derived from photogrammetric data via a somehow painstaking process. In this latter process, indicator kriging was eventually run but only after visually identifying a set of classified samples, transforming data formats and calculating semivariograms and fitting models. This very time-consuming and computing-intensive process was made more complicated by the need for off-line work from one system to another.

In addition to the relative degrees of difficulties, fuzzy surfaces derived from remotely sensed images were seen to be highly varied, while those derived from photogrammetric data were rather localised and less “fuzzy”. This difference in the degrees of fuzziness is probably because remotely sensed images existing in the form of multispectral reflectance data are intrinsically varied and complex, implying that large number of pixels will be mixed, in particularly for remotely sensed images with

coarse spatial resolutions, such as Landsat TM images. On the contrary, for photogrammetric data, fuzzy surfaces are derived by firstly identifying a set of pure samples, then interpolating the probabilities of finding a certain class at any location by using a suitable semivariogram model. It is important to note that these pure samples are, by nature, highly subjective and generalised, meaning that the complexity and fuzziness in the underlying phenomenon are hardly captured in full detail by the finite number of pure samples.

As described previously in this chapter, for the purpose of further discussion in the context of combined use of different types of fuzzy surfaces, the fuzzy surfaces derived from remotely sensed images were referred to as fuzzy classified data, while the fuzzy surfaces derived from photogrammetric data are termed fuzzy reference data. It was then seen that fuzzy reference data derived from indicator kriging are closer to the fuzzy classified data than those based on sub-pixel component land cover proportion data, because lower cross-entropy measures were obtained from the former than the latter. Extra evidence was obtained by calculating correlation coefficients, reinforcing the fact that indicator kriging is more suitable for generating fuzzy reference data than simply taking sub-pixel component land cover proportions as surrogate fuzzy reference data.

Furthermore, a variety of categorical maps was derived from fuzzy surfaces by using the maximising and slicing operators. The way that these categorical maps were visualised was very convincing and conveyed a strong impression of the spatial variabilities in the uncertainties of mapped data. Moreover, epsilon band models would be easily derived from the unclassified areas representing uncertainty zones of different patterns and widths. This fact confirms that uncertainties are better modelled from continuous fields than from discrete objects.

The effects of maintaining fuzziness in the mapped data on the accuracy assessment were then considered by using error matrices, which were derived from different methods of comparisons. Once again, it was confirmed that increased agreements can be obtained if the fuzziness is accommodated not only in the remotely sensed data, but also in the assumed reference data, if pure pixels are discriminated from mixed pixels on a quantitative basis, and if the full spectrum from the most likely classes to the least likely classes is explored.

In overview, the surface-based approach presented in this chapter has shown, firstly, that fuzzy surfaces, as opposed to conventional categorical maps consisting of discrete objects, can be derived from common types of data involved in land cover mapping without major difficulties. The derivation of fuzzy surfaces from remote sensing data is, in essence, a process of fuzzy classification by allowing partial and multiple memberships at each pixel, while the derivation of fuzzy surfaces from photogrammetric data is accomplished by sampling pure points representative of corresponding classes and then inferring fuzzy membership values at unsampled points through spatial interpolation.

Secondly, it has been shown that there is a variety of methods for analysing fuzzy surfaces and these methods are effective and flexible. This is because conventional categorical maps can be easily produced from the fuzzy surfaces, and the categorical maps produced also provide means to communicate uncertainties concerned in an informative way. One of the most impressive functions of fuzzy surfaces is that epsilon error band models can be generated from fuzzy surfaces, which would otherwise be difficult to map at the level of individual lines. Moreover, fuzziness retained by fuzzy surfaces can be usefully explored. This has been demonstrated through accommodating for fuzziness in processes such as selecting pure pixels and performing soft comparisons.

Comparing the object-based approach presented in Chapter 5 with the field-based approach presented in this chapter confirms that objects are not able to capture the fuzziness intrinsic to land cover mapping, and hence are not adequate by themselves for explicitly representing the spatially-varying nature of uncertainties, while fields permit an integral mechanism to maintain spatial variations and heterogeneities during spatial data acquisition, which are central to the construction and utilisation of fuzzy surfaces. In other words, the empirical study presented in Chapters 5 and 6 has established that land cover, as the phenomenon of specific interests to this thesis, is better mapped from fields than from objects.

## Chapter 7

### Conclusions

As a concluding chapter, this chapter gives a concise summary of the work accomplished in earlier chapters. Such a summary is followed by an outlook for further work along the direction established in this thesis.

#### 7.1 Handling uncertainties in spatial databases

This thesis has pursued the issue of uncertainties by following a systematic treatment of uncertainties in spatial databases. It was firstly discussed that uncertainties arise in GISs due to compounded effects of geographical abstraction, spatial data acquisition and geo-processing. Secondly, it was shown that uncertainties might be approached either as fields or as objects, the two fundamentally distinctive perspectives of geographical abstraction. However, the issue of uncertainties, in particularly those uncertainties involved in land cover mapping, is better addressed from fields rather than from objects. Thirdly, in pragmatic terms, it was described that uncertainties in categorical land cover data are best represented as sets of fuzzy surfaces. Under such a strategy, spatial variabilities and heterogeneities are maintained and may be usefully explored. Finally, results from an empirical study confirmed that field-based methods are better suited for effective handling of uncertainties in a variety of digital land cover data than their object-based counterparts. These points are reiterated below.

As detailed in Chapter 2, there are two fundamentally distinctive perspectives of spatial data modelling: discrete objects versus continuous fields, depending on whether the real world is viewed as populated by a collection of points, lines and areas or conceived as a set of single-valued functions defined everywhere. While object-based models are particularly effective for representing well-defined spatial

entities, fuzzy phenomena are better represented by field-based models which maintain the spatial variabilities and heterogeneities deemed important in geographical studies. The choice of spatial data models depends on the nature of the phenomenon under study, and is often related to the specific methods of data acquisition and database implementations. Such a choice will affect the extent to which uncertainties may be handled. Clearly, spatial data modelling is the most central issue in conducting research on uncertainties in spatial databases.

Land cover mapping typically uses land surveying, photogrammetric and remote sensing techniques as primary data acquisition methods, and uses map digitising as its main secondary data acquisition method. Errors have been found to occur in each of these methods: in land surveying, errors in coordinate positions and heights occur because of instrument design or miscalibration or in computational algorithm. Such errors may be systematic or random, but are generally small in amount. In photogrammetry, errors are associated with inadequate stereo-orientation, inadequate ground control or poor operator practice in line following or heighting. Errors in remote sensing can occur from incomplete radiometric and geometric correction, while the process of band selection and classification can produce results of greater or lesser completeness. Errors in map digitising result from the original map scale and hence degree of generalisation of the map detail, and from poor operator practice. Uncertainties also occur during spatial data acquisition due to the complex interaction of human and machine factors. Moreover, uncertainties inherent in diverse data layers will be propagated in geo-processing, such as in a map overlay. It is clear from the account above that research on GIS uncertainties must start from the fundamental issue of how the real world is perceived and measured in the first place, because this starting point underlies any systematic approach to the handling of GIS uncertainties.

In data acquisition methods, the attributes may be considered from the perspective of objects or of fields, and similarly uncertainties may be approached from the same two



perspectives, as has been seen comparatively in Chapter 3. In the object domain, objects are described by positions and attributes. Thus, there is a tradition in object-based models to address positional and attribute uncertainties separately, which may be incorporated in an extended geo-relational spatial database. Assuming that ground truth exists for discrete objects, it is possible to test the positional and attribute accuracies of objects by means of check surveys.

In the field domain, on the other hand, there is only one kind of uncertainty: the deviation of the value of a variable from its true value or the value assumed to be true. It is possible to use probability theory to describe uncertainties in fields, so that appropriate confidence levels can be placed for values of variables to be within certain intervals. It is, however, meaningless to talk about positional accuracies, unless discrete objects are referred to. This is because the objects employed to represent fields do not generally exist in the real world, but are results of the process of geographical abstraction. Also relevant is the fact that it is impossible to use the method of "testing against ground truth" for assessing the accuracies of fields, except for the case of the real, visible and stable topographical variables such as elevation. As the focus of this thesis, categorical variables such as land cover have, in many situations, to rely on fuzzy concepts and methods for proper representation and interpretation. It is thus necessary to use fuzzy surfaces to represent explicitly the uncertainties in categorical fields via the mechanism of partial and multiple memberships at any location.

While Chapter 3 has provided an outline for representing uncertainties in categorical fields by means of fuzzy surface, Chapter 4 has considered the possible ways by which two types of fuzzy surface are derived, from digital image processing (typically using remotely sensed data) and from visual interpretation of photogrammetric data. While there exist both supervised and unsupervised methods for fuzzy classifications based on remote sensing data, interpolating fuzzy surfaces from photogrammetric data uses

different types of functions, especially functions relating fuzzy membership values (FMVs) with distances, or geostatistical approaches. It has been argued that fuzzy *c*-means clustering and indicator kriging are the most suitable for deriving fuzzy surfaces from remote sensing data and photogrammetric data, respectively. A suite of methods for the analysis of fuzzy surfaces has been described, which serves to demonstrate how fuzzy surfaces, in contrast to traditional categorical maps, may be assessed and compared, and how the latter may be derived from the former. In a sense, Chapter 4 has confirmed that a sensible way for deriving measures of uncertainty lies in the data acquisition process itself, and uncertainties in fields are better considered as inherent components of the underlying phenomena, thus lending themselves to efficient exploration.

Chapters 5 and 6 have presented a case study in which both object-based and field-based approaches were implemented and compared in an urban-orientated spatial database where different types of digital land cover data with different formats and accuracies were combined. In Chapter 5, a co-registered hierarchy of data layers was collected via field surveying, photogrammetry (aerial photographs at 1:5,000 and 1:24,000 scales) and remote sensing (SPOT HRV and Landsat TM images) with a corresponding hierarchy of accuracies, so that each of the data layers has an independent data layer of higher accuracy to check against. It was then seen that object-based methods do not provide spatially varying information on the uncertainties in the underlying data and are thus inappropriate to model the complexity of uncertainties involved in digital land cover mapping, in particular when data with different levels of uncertainties are combined.

In order to overcome some of the shortcomings of object-based models for the handling of uncertainties, Chapter 6 has presented empirical results from using the surface-based approach. Fuzzy surfaces were derived from remote sensing images by using the fuzzy *c*-means clustering, and from photogrammetric data by using, firstly,

proportions of component land cover types in individual grid cells and, secondly, indicator kriging. Categorical maps were generated from fuzzy surfaces by using the “maximising” and “slicing” operations. The spatial variabilities of uncertainties in the underlying data were visualised by, firstly, draping the classified data onto the spatially varying surfaces of certainties, with which each location is correctly labelled, and, secondly, producing a series of categorical maps with increasing thresholds, above which each location is classified, so that locations classified have increasing accuracies while locations unclassified might be generalised to form suitable epsilon band models. Comparisons were made by using hardened fuzzy classified data and sliced fuzzy classified data, in order to assess the different levels of accuracies obtainable when using fuzzy classified data. These comparisons were followed by so-called “soft” comparisons, in which both the most likely and the second most likely classes were scrutinised.

It has been found that fuzzy surfaces based on remote sensing images are more varied than those based on photogrammetric data, as explained by the measures of entropy. For the latter types of fuzzy surfaces, indicator kriging has been found to be more suitable as a constructor of fuzzy surfaces than resolving proportions of component land cover types in individual grid cells, because closer associations between fuzzy classified remote sensing data and photogrammetric data were obtained from using the former technique, indicator kriging. It was then found that a variety of categorical maps can be derived from fuzzy surfaces produced in a straightforward manner. The categorical maps presented have greatly enhanced the visualisation of uncertainties inherent in common types of digital land cover data. Moreover, the categorical maps produced by using the slicing operation provided basic and valuable information, from which an object-based presentation of uncertainties in categorical maps, i.e., epsilon band models, could be easily derived. Comparisons of fuzzy classified data confirm that fuzzy methods are superior to conventional hard methods, especially in situations

where fuzziness is intrinsic to both remote sensing images and photogrammetric data, the supposed reference data.

At this stage, it is useful to restate the main points covered in earlier chapters and reviewed previously in this section. A summary is shown in Table 7.1, conforming to the work presented in this thesis in following the fundamental distinction between discrete objects and continuous fields. Table 7.1 may be seen as an update and extension of Table 1.1 where past work on GIS uncertainties was briefly documented.

Table 7.1 Object-based versus field-based approaches to GIS uncertainties

	Object-based	Field-based
Suitability	for phenomena better viewed as well-defined entities having positions and attributes, such as roads and land parcels	for phenomena, mainly fuzzy, better conceived as a set of single-valued functions defined everywhere, such as land cover and soil type
Emphasis	discretisation, so highly abstract	spatial variability and heterogeneity, so very close to the reality
Description	definable and separable positional and attribute uncertainties	only in the values of variables concerned such as land cover types, both probabilistic and fuzzy measures are useful
Assessment	through independent sources of higher accuracies: a hierarchy of digital land cover data with a corresponding hierarchy of accuracies, experimental in nature	as an integral component of data acquisition: classification is fuzzy but objective, visual interpretation is assumed hard but implies fuzziness
Representation	extended geo-relational databases: attached to groups of objects and attributes	surfaces, in the case of categorical variables, using fuzzy surfaces
Analysis	complex	effective and efficient

As shown in Table 7.1, while both object-based and field-based approaches are useful in their respective contexts, the latter is seen to be better than the former, in particular in the case of land cover mapping. More importantly, field-based approaches are worth further investigation because analysis via continuous fields is seen to be more effective and efficient than that via discrete objects, which, in turn, enhances the opportunity for field-based approaches to be more adaptable to modelling uncertainties than object-based approaches, as developed in Section 7.2.

Fuzzy concepts and methods, as the main foundation of this thesis, were synthesised and incorporated as one of the few novel uses in the context of urban and suburban land cover mapping, where data of varying accuracies are frequently merged. The claim for novelty is that spatial data involved in such an environment as urban and suburban land cover mapping lie in a position somewhere between discrete objects and continuous fields. Results have verified that fuzzy surface-based methods are useful to represent and handle, in particular, the uncertainties present in common types of digital land cover data in urban and suburban areas. By using fuzzy surface-based methods, complex spatial variations of spatial data and their uncertainties are visualised, interpreted and analysed on a quantitative basis. Such a strategy holds considerable potentials for routine GIS applications. It has then been argued that the fuzzy surface-based approach presented in this thesis provides a good starting point for comprehensive research into the spatial variability of diverse spatial data and their uncertainties.

In addition to the novel aspects mentioned above, there are other characteristics that make this thesis a valuable asset to the continuing and growing debate on GIS uncertainty issues. Firstly, fuzzy and probabilistic concepts have been clarified with special reference to common types of digital land cover data involved in urban and suburban areas. This clarification helps to organize the arguments presented in later chapters on the case study and may be used to guide further work on GIS

uncertainties. Secondly, an improvement has been made to the fuzzy *c*-means clustering algorithm, which uses per-class statistics, as opposed to the global statistics used in conventional algorithms, to strengthen class statistics in the resulting fuzzy classified data (thus reducing their degree of fuzziness) and make fuzzy classified remote sensing data more compatible with fuzzy classified photogrammetric data. Thirdly, geostatistics has been successfully applied in the construction of fuzzy surfaces from photogrammetric land cover data, which, as a collection of discrete polygons, are traditionally assumed to be of absolute accuracy, both positional and attribute. This may lead to improved recognition of photogrammetry as the means of acquiring spatial data for relevant applications, and thus may greatly help to expand its scope as a traditional discipline. GIS communities will benefit more from aerial photography and photogrammetric techniques if they share the same insights and understanding.

In overview, the fuzzy surface-based approach pursued in this thesis reinforces the value of a “fuzzy” view in the handling of uncertainties. Such a fuzzy view will help to enhance human perception about the nature of the spatial data and the underlying geographical phenomena. It is further considered that the research presented in this thesis stands as an interesting and viable topic for the future. This is especially important when it is recognised that natural links exist between fuzzy surfaces and geostatistics, as will be elaborated in the next section.

## **7. 2 Prospects for the surface-based approach to uncertainties in GISs**

As summarised in the previous section, the fuzzy surface-based approach is more suitable for representing and handling uncertainties in spatial data than the object-based alternative. This is especially true in the context of urban and suburban land cover mapping where real geographical phenomena cannot be represented perfectly by either discrete objects associated with exactly valued attributes or continuous

fields. Land cover in such areas, as a specific phenomenon, displays much complexity, which must be properly mapped. Fuzzy surfaces provide a mechanism suitable for depicting such complexity.

It has to be realised that using fuzzy surfaces to represent spatial data, irrespective of their underlying data formats, may need expanded storage and involve complicated computational implementations. This should be well justified by the increasing need for information on spatial data quality and decreasing costs of hardware and software overheads. Nevertheless, it is highly desirable to incorporate error-handling functions into widely used GIS packages, so that more users can easily handle the spatial data and their uncertainties in a straightforward manner.

The question arises, however, as to how uncertainties are predicted for data products derived from certain GIS operations such as map overlay, when the fuzzy surfaces recording spatially varying uncertainties are generated by using methods similar to those presented in earlier chapters of this thesis. It may not be possible to predict the uncertainties in overlaid maps straightforwardly, even if all individual maps contain information on uncertainty at each location. This is because the uncertainty obtainable in an overlaid map relies on both the uncertainties existing in individual map layers and the spatial correlation of different uncertainties, as described in Chapter 3. The spatial correlation between uncertainties in different map layers needs to be quantified in order to produce an uncertainty map for an overlaid map. Similar problems exist in the outputs of derivative data products. For example, slope data are often calculated from DEMs. One may wish to find out the uncertainty in slope data. Suppose uncertainty for elevation data is available. The uncertainty in slope data will depend on the uncertainty in elevation data, and the spatial correlation of the uncertainty in neighbouring points, which is not available from the uncertainty of elevation. In other words, uncertainty in elevation alone is not sufficient to estimate the uncertainty in slope data, unless the amount of spatial correlation in uncertainty is known

(Goodchild 1994, pers. comm.).

As developed in earlier chapters, geostatistical approaches are designed to work on spatially correlated and inter-correlated random variables representing many geographical phenomena. They use semivariograms to quantify the amount of spatial dependence present, which is deemed crucial to the studies on error propagation. It was seen previously in this thesis that kriging, as a branch of the geostatistical approach, has been developed to deal with categorical variables as well as continuous variables. For continuous variables, kriging is preferably used as a special kind of spatial interpolation method, through its capability of accounting for spatial dependence intrinsic to many mapped phenomena. For categorical variables, indicator kriging may be used. Indicator kriging estimates the conditional (posterior) probability distribution without making assumptions about the form of the prior distribution functions. This feature makes geostatistics particularly attractive for its application in deriving fuzzy maps directly from a set of classified samples, as shown in Chapters 4 and 6.

There is another branch of geostatistical approaches: stochastic simulation. Stochastic simulation is the process of building alternative, equally probable, high resolution models of the spatial distribution of a variable in a field. The simulation is said to be conditional if the resulting realisations honour the hard data values at their locations. While most interpolation methods based on geostatistics seek to provide a best local estimated value for each unsampled point without specific regard to the resulting spatial statistics of the estimates, a stochastic simulation is more concerned with retaining spatial variability than spatial smoothing. In a recent paper, Journel (1996) predicted that stochastic simulation with its underlying principle of spatial dependence will play an important role for research on GIS uncertainties and the spatial data themselves.



There are two related objectives for stochastic simulation. The first is to create realisations that honour, subject to statistical fluctuations, the histogram, semivariogram and the conditioning data. These realisations avoid the smoothing effect of kriging if used as estimated maps. The second objective is to model uncertainty through multiple realisations, which would otherwise be difficult to derive for GIS operations incorporating several map layers. This is of particular value to those seeking to understand how uncertainties are propagated during geo-processing, thus allowing for their prediction. A sequential Gaussian method can be used to simulate additional realisations to reproduce the spatial variability, which is useful in determining the spatial uncertainties of maps, such as those produced by intersection of map layers of interest. In terms of stochastic simulation, of particular interest to research in GIS uncertainties is error simulation method based on geostatistics.

It becomes apparent that geostatistics, in particular stochastic simulation, offers theoretically sound and practically effective functions to research on modelling uncertainties in GISs. When spatial data are represented as fuzzy surfaces consisting of fuzzy or probabilistic measures, the links between surface data and geostatistics appear straightforward, and thus geostatistics for modelling uncertainties in overlaid fuzzy maps becomes a natural choice.

GISs have been developed to facilitate the complex spatial decisions that occur in the environment of spatial databases. GISs have also increased the need for better and more comprehensive error-handling functionality in the computerised spatial data processing environment, because the reliability of decision makings based on GISs becomes extremely important and complicated when the amount of spatial data is growing rapidly and steadily, and when uncertainties are propagated pervasively. Therefore, it is anticipated that geostatistics will have an increasingly important role to play in fostering new generations of GISs that are equipped with functionalities for handling uncertainties.

## References

- Abler R.F., 1987, The National Science Foundation Centre for Geographical Information and Analysis. *International Journal of Geographical Information Systems*, Vol. 1, No. 4, pp. 303-26.
- Altman D., 1994, Fuzzy set theoretic approaches for handling imprecision in spatial analysis. *International Journal of Geographical Information Systems*, Vol. 8, No. 3, pp. 271-89.
- American Society for Photogrammetry and Remote Sensing (ASPRS), 1989, ASPRS interim accuracy standards for large scale line maps. *Photogrammetric Engineering and Remote Sensing*, Vol. 55, pp. 1038-40.
- Aronoff S., 1989, *Geographical information systems: a management perspective*. Ottawa: WDL Publications.
- Aspinall R.J. and Pearson D.M., 1994, Describing and managing data quality for categorical maps in GIS. In: *Proceedings GISRUK,94*, Leicester, pp.161-70.
- Bailey R.G., 1988, Problems with using overlay mapping for planning and their implications for geographic information systems. *Environmental management*, Vol. 12, No. 1, pp. 11-17.
- Baker R., DeSteiguer J., Grant D. and Newton M., 1979, Land-use/land-cover mapping from aerial photographs. *Photogrammetric Engineering and Remote Sensing*, Vol. 45, No. 5, pp. 661-8.
- Beard M.K., 1994, Accommodating uncertainty in query response. In: *Proceedings 6th International Symposium on Spatial Data Handling*, Edinburgh, pp. 240-53.
- Bedard Y., 1987, Uncertainties in land information systems databases. In: *Proceedings AutoCarto 8*, pp. 175-84.
- Berry J.K., 1987, computer-assisted map analysis: potentials and pitfalls. *Photogrammetric Engineering and Remote Sensing*, Vol. 53, No. 10, pp. 1405-10.
- Berry J.K., 1993, *Beyond mapping: concepts, algorithms, and issues in GIS*. Colorado: GIS world, Inc..
- Bezdek J.C., Ehrlich R. and Full W., 1984, FCM: the fuzzy *c*-means clustering algorithm. *Computers and Geosciences*, Vol. 10, No. 2 and 3, pp: 191-203.
- Bierkens M.F.P. and Burrough P.A., 1993a, The indicator approach to categorical soil data. I: theory. *Journal of Soil Science*, Vol. 44., pp. 361-8

- Bierkens M.F.P. and Burrough P.A., 1993b, The indicator approach to categorical soil data. II: applications to mapping and land use suitability analysis. *Journal of Soil Science*, Vol. 44., pp. 369-81.
- Blakemore M., 1984, Generalisation and error in spatial databases. *Cartographica*, Vol. 21, No. 2 and 3, pp. 131-9.
- Bolstad P., Gessler P. and Lillesand T.M., 1990, Positional uncertainty in manually digitized map data. *International Journal of Geographical Information Systems*, 1990, Vol. 4, No. 4, pp. 399-412.
- Bregt A.K. 1991, Mapping uncertainty in spatial data. In: *Proceedings EGIS'91*, pp. 149-54.
- Burrough P.A., 1986, *Principles of Geographical Information Systems for Land Resources Assessment*. Oxford: Clarendon.
- Burrough P.A., 1987, Multiple sources of spatial variation and how to deal with them. In: *Proceedings AutoCarto 8*, pp. 145-54.
- Burrough P.A., 1992, Are GIS data structures too simple minded? *Computers and geosciences*. Vol. 18, No. 4, pp. 395-400.
- Burrough P.A., 1994, Accuracy and error in GIS. In: Green D.R. and Rix D. (eds.), *The AGI source book for geographic information systems 1995*, pp. 87-91.
- Burrough P.A. and Frank A.U. (eds.), 1996, *Geographic objects with indeterminate boundaries*. Basingstoke: Taylor and Francis.
- Burrough P.A., van Rijn R. and Rikken M., 1993, Spatial data quality and error analysis issues: GIS functions and environmental modeling. In: *Proceedings of the 2nd International Workshop on Integrating Geographic Information Systems and Environmental Modeling*, Colorado.
- Campbell J.B., 1987, *Introduction to Remote Sensing*. New York: The Guilford Press.
- Carter J.R., 1988, Digital representations of topographic surfaces: an overview. In: *Technical Papers 1988 ASPRS Annual Conventions*, Vol. 5, pp. 54-60.
- Carter P. and Gardner W.E., 1977, An urban management information service using Landsat imagery. *Photogrammetric Record*, Vol. 9, pp. 157-71.
- Chorley R.J., 1966, The application of statistical methods to geomorphology. In: Dury G.H. (ed.), *Essays in Geomorphology*. London: Heinemann.

- Chrisman N.R. 1989, Modelling error in overlaid categorical maps. In: Goodchild M. F. and Gopal S. (eds.), *Accuracy of Spatial Databases*, London: Taylor and Francis, pp. 21-34.
- Chrisman N.R., 1991, The error component in spatial data. In: Maguire D.J., Goodchild M. F. and Rhind D.W.(eds.), *Geographical Information Systems: principles and applications*, London: Longman, pp. 165-74.
- Congalton R.G., 1991, A review of assessing the accuracy of classification of remotely sensed data. *Remote Sensing of Environment*, Vol. 37, No. 1, pp. 35-46.
- Cooper M.A.R. and Cross P.A., 1988, Statistical concepts and their application in photogrammetry and surveying. *Photogrammetric Record*, Vol. 12, No. 71, pp. 637-63.
- Cowen D.J., 1988, GIS versus CAD versus DBMS: what are the differences? *Photogrammetric Engineering and Remote Sensing*, Vol. 54, No. 11, pp. 1551-5.
- Cressie N., 1993, Geostatistics: a tool for environmental modellers. In: Goodchild M.F., Parks B.O. and Steyaert L.T. (eds.), *Environmental modelling with GIS*, New York: Oxford University Press, pp. 414-21.
- Dahlberg R.E., 1986, Combining data from different sources. *Surveying and Mapping*, Vol. 46, No.2, pp. 141-9.
- Davis F.W. and Simonett D.S., 1991, GIS and remote sensing, In: Maguire D.J., Goodchild M. F. and Rhind D.W.(eds.), *Geographical Information Systems: principles and applications*, London: Longman, Vol. 1, pp. 191-213.
- Deutsch C. and Journel A.G., 1992, *GSLIB: Geostatistical Software Library and User's Guide*. New York and Oxford: Oxford University Press.
- Dougenik J., 1979, WHIRLPOOL: a generic processing for polygon coverage data. In: *Proceedings AutoCarto 4*, Vol. 2, pp. 304-11.
- Drummond J.E., 1987, A framework for handling error in geographic data manipulation. *ITC Journal*, Vol. 1, pp. 73-82.
- Drummond J.E. and Ramlal B., 1992, A prototype uncertainty sub-system implemented in ITC's ILWIS PC-based GIS, and tested in a Dutch land reallocation project. In: *Proceedings EGIS'92*, Vol. 1, pp. 234-43.
- Dunn R., Harrison A.R. and White J.C., 1990, Positional accuracy and measurement error in digital databases of land use: an empirical study. *International Journal of Geographical Information Systems*, Vol. 4, No. 4, pp. 385-98.

- Dutton G., 1992, Handling positional uncertainty in spatial databases. In Proceedings of the 5th International Symposium on Spatial Data Handling, Charleston, pp. 460-9.
- Edwards G. and Lowell K.E., 1996, Modelling uncertainty in photointerpreted boundaries. *Photogrammetric Engineering and Remote Sensing*, Vol. 62, No. 4, pp. 373-91.
- Englund E.J., 1993, Spatial simulation: environmental applications, In: Goodchild M.F., Parks B.O. and Steyaert L.T. (eds.), *Environmental Modeling with GIS*, New York: Oxford University Press, pp. 432-7.
- Estes J.E., 1992, Remote sensing and GIS integration: research needs, status and trends. *ITC Journal*, pp. 2-9.
- Fisher P.F., 1991, Spatial data sources and data problems. In: Maguire D.J., Goodchild M. F. and Rhind D.W. (eds.), *Geographical Information Systems: principles and applications*, London: Longman, pp. 175-86.
- Fisher P.F., 1994, Probable and fuzzy models of the viewshed operation. In: Worboys M.F. (ed.), *Innovations in GIS*, London: Taylor and Francis, pp. 161-75.
- Fisher P.F. and Pathirana S., 1989, Urban boundary detection from Landsat imagery: a GIS and knowledge-based approach to MSS data. In: *Technical Papers 1989 ASPRS/ACSM Annual Convention*, Vol. 4, pp. 93-101.
- Flowerdew R., 1991, Spatial data integration. In: Maguire D.J., Goodchild M. F. and Rhind D.W. (eds.), *Geographical Information Systems: principles and applications*, London: Longman, pp. 375-87.
- Foody G.M., 1992, On the compensation for chance agreement in image classification accuracy assessment. *Photogrammetric Engineering and Remote Sensing*, Vol. 58, No. 10, pp. 1459-60.
- Foody G.M., 1995a, Cross-entropy for the evaluation of the accuracy of a fuzzy land cover classification with fuzzy ground data. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 50, No. 5, pp. 2-12.
- Foody G.M., 1995b, Fully fuzzy supervised image classification. In: *RSS'95 Remote Sensing in Action*, Southampton, pp. 1187-94.
- Foody G.M., 1996, Relating the land-cover composition of mixed pixels to artificial neural network classification outputs. *Photogrammetric Engineering and Remote Sensing*, Vol. 62, No. 5, pp. 491-9.
- Foody G.M. and Cox D.P., 1994, Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions. *International Journal of Remote Sensing*, Vol. 15, No. 3, pp. 619-31.

- Foody G.M., Campbell N.A., Trodd N.M. and Wood T.F., 1992, derivation and applications of probabilistic measures of class membership from the maximum likelihood classification. *Photogrammetric Engineering and Remote Sensing*, Vol. 58, No. 12, pp. 1335-41.
- Forster B., 1983, Some urban measurements from Landsat data. *Photogrammetric Engineering and Remote Sensing*, Vol. 49, No. 12, pp. 1693-707.
- Gautam N.C., 1976, Aerial photo interpretation techniques for classifying urban land use. *Photogrammetric Engineering and Remote Sensing*, Vol 42, No. 6, pp. 815-22.
- Goodchild M.F., 1989a, Modelling error in objects and fields. In: Goodchild M.F. and Gopal S. (eds.), *Accuracy of Spatial Databases*, London: Taylor and Francis, pp. 107-13.
- Goodchild M.F., 1989b, Modelling error in spatial databases. In: *Proceedings GIS/LIS'89*, pp. 154-62.
- Goodchild M.F., 1991, Issues of quality and uncertainty. In: Muller J.C. (ed.), *Advances in Cartography*, London and New York: Elsevier Science Publication Ltd., pp. 113-39.
- Goodchild M.F., 1992, Geographical data modelling. *Computers and Geosciences*, Vol 18., No. 4, pp. 401-8.
- Goodchild M.F., 1993, The state of GIS for environmental problem solving. In: Goodchild M.F., Parks B.O. and Steyaert L.T. (eds.), *Environmental Modelling with GIS*, New York and Oxford: Oxford University Press, pp. 8-15.
- Goodchild M.F. and Lam N., 1980, Areal interpolation: a variant of the traditional spatial problem. *Geo-Processing*, Vol. 1, No. 3, pp. 297-312.
- Goodchild M.F., Lin C.C. and Leung Y., 1994, Visualizing fuzzy maps. In: Hearnshaw H.M. and Unwin D.J. (eds.), *Visualization in Geographical Information Systems*, New York: John Wiley and Sons, pp. 158-76.
- Goodchild M.F., Sun G.Q. and Yang S.R., 1992, Development and test of an error model for categorical data. *International Journal of Geographical Information Systems*, Vol. 6, No. 2, pp.87-104.
- Gopal S. and Woodcock C., 1994, Theory and methods for accuracy assessment of thematic maps using fuzzy sets. *Photogrammetric Engineering and Remote Sensing*, Vol. 60, No. 2, pp. 181-8.

- Guptill S.C., 1989, Inclusion of accuracy data in a feature based, object-oriented data model. In: Goodchild M. F. and Gopal S. (eds.), *Accuracy of Spatial Databases*, London: Taylor and Francis, pp. 91-8.
- Guptill S. and Morrison J. (eds.), 1995, *Elements of spatial data quality*. Oxford: Elsevier Scientific / International Cartographic Association.
- Harley J.B., 1975, *Ordnance Survey Maps: a descriptive manual*. Southampton: Ordnance Survey.
- Harris R., 1987, *Satellite Remote Sensing: an introduction*. London and New York: Routledge and Kegan Paul.
- Heuvelink G.B.M. and Burrough P.A., 1993, Error propagation in cartographic modelling using Boolean logic and continuous classification. *International Journal of Geographical Information Systems*, Vol. 7, No. 3, pp. 231-46.
- Heuvelink G.B.M., Burrough P.A. and Stein A., 1989, Propagation of errors in spatial modelling with GIS. *International Journal of Geographical Information Systems*, Vol. 3, No. 4, pp. 303-22.
- Hunter G.J. and Goodchild M.F., 1995, Dealing with error in spatial databases: a simple case study. *Photogrammetric Engineering and Remote Sensing*, Vol. 61, No. 5, pp. 529-37.
- 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. 123-30.
- Janssen L. and van der Wel F., 1994, Accuracy assessment of satellite derived land cover data: a review. *Photogrammetric Engineering and Remote Sensing*, Vol. 60, No.4, pp. 419-26.
- Jensen J.R. (ed.), 1983, *Urban/suburban land use analysis*. In: Colwell R.N. (ed.), *Manual of Remote Sensing*, Falls Church, Virginia: The Sheridan Press.
- Jensen J.R., Cowen D.J., Halls J. et al., 1994, Improved urban infrastructure mapping and forecasting for BellSouth using remote sensing and GIS technology. *Photogrammetric Engineering and Remote Sensing*, Vol. 60, No. 3, pp. 339-46.
- Joao E., Herbert G., Openshaw S. and Rhind D., 1992, Magnitude and significance of generalization and its effects. In: *Proceedings EGIS'92*, Vol. 1, pp. 711-21.
- Journel A.G., 1996, Modelling uncertainty and spatial dependence: stochastic imaging. *International Journal of Geographical Information Systems*, Vol. 10, No. 5, pp. 517-22.

Kaufmann A., 1975, *Introduction to the Theory of Fuzzy Subsets*. New York: Academic Press.

Kennedy M., 1996, *The Global Positioning System and GIS*. Middlesex: Ann Arbor Press.

Key J.R., Maslanik J.A. and Barry R.G., 1989, Cloud classification from satellite data using a fuzzy sets algorithm: a polar example. *International Journal of Remote Sensing*, Vol. 10, No. 12, pp. 1823-42.

Kirby R.P., 1988, Measuring the areas of rural land use parcels. In: Bunce R.G.H. and Barr C.J. (eds.), *Rural Information for Forward Planning*, Grange-over-Sands: Institute of Terrestrial Ecology, pp. 99-104.

Kirby R.P., 1992, The 1987-1989 Scottish national aerial photographic initiative. *Photogrammetric Record*, Vol. 14, No. 80, pp. 187-200.

Kirby R.P. and Zhang J.X., 1993, A built-in paradigm for error modelling in an urban orientated GIS. In: *Proceedings of the International Colloquium on Advances in Urban Spatial Information and Analysis*, Wuhan (China), pp. 257-64.

Langford M. and Unwin D.J., 1994, Generating and mapping population density surface within a geographical information system. *The Cartographic Journal*, Vol. 31, pp. 21-6.

Lanter D.P. and Veregin H., 1990, A lineage meta-database program for propagating error in geographic information systems. In: *Proceedings GIS/LIS'90*, Vol. 1, pp. 144-53.

Lillesand T.M. and Kiefer R.W. 1994, *Remote Sensing and Image Interpretation*, 3rd edition. New York: John Wiley and Sons.

Lo C.P., 1971, A typological classification of buildings in the city centre of Glasgow from aerial photographs. *Photogrammetria*, Vol. 27, pp. 135-57.

Lowell K.E., 1992, On the incorporation of uncertainty into spatial data systems. In: *Proceedings GIS/LIS'92*, Vol. 2, pp. 484-93.

Lowell K.E., 1994, An uncertainty-based spatial representation for natural resources phenomena. In: *Proceedings 6th International Symposium on Spatial Data Handling*, Edinburgh, pp. 933-44.

Lunetta R.S., et al., 1991, Remote sensing and geographic information system data integration: error sources and research issues. *Photogrammetric Engineering and Remote Sensing*, Vol. 57, No. 6, pp. 677-87.



- Maguire D.J., 1991, An overview and definition of GIS. In: Maguire D.J., Goodchild M. F. and Rhind D.W. (eds.), *Geographical Information Systems: principles and applications*, London: Longman, pp. 9-20.
- Mark D.M. and Csillag F., 1989, The nature of boundaries on area-class maps. *Cartographica*, Vol. 21, pp. 65-78.
- Mark D.M., Lauzon J.P. and Cebrian J.A., 1989, A review of quadtree-based strategies for interfacing coverage data with digital elevation models in grid form. *International Journal of Geographical Information Systems*, Vol. 3, No. 1, pp. 3-14.
- Mather P.M. 1987, *Computer Processing of Remotely Sensed Images: an introduction*. Chichester: Wiley.
- Middelkoop H., 1990, Uncertainty in a GIS: a test for quantifying interpretation output. *ITC Journal*, Vol. 3, pp. 225-32.
- NCDCDS (National Committee for Digital Cartographic Data Standards), 1988, The proposed standard for digital cartographic data. *The American Cartographer*, Vol. 15, No.1, pp. 9-140.
- NCGIA (National Centre for Geographic Information and Analysis), 1989, The research plan of the National Centre for Geographic Information and Analysis. *International Journal of Geographical Information Systems*, Vol. 3, No.2, pp. 117-36.
- Neto F., 1996, Airborne and spaceborne imagery for mapping. In: *Surveying World*, Vol. 4, No. 2, pp. 27-9.
- Newcomer J.A. and Szajgin J., 1984, Accumulation of thematic map errors in digital overlay analysis. *The American Cartographer*. Vol. 11, No.1, pp. 58-62.
- Oliver M.A. and Webster R., 1990, "Kriging": a method of interpolation for geographical information systems. *International Journal of Geographical Information Systems*, Vol. 4, No.3, pp. 313-32.
- Peuquet D.J., 1984, A conceptual framework and comparison of spatial data models. *Cartographica*, Vol. 21, No. 4, pp. 66-113.
- Pullar D. and Beard K., 1989, Specifying and tracking errors from map overlay. In: *Proceedings GIS/LIS'89*, pp. 79-87.
- Richards, J.A., 1993, *Remote Sensing Digital Image Processing: an introduction* (second, revised and enlarged edition). Berlin: Springer-Verlag.
- Robinson V.B., 1988, Some implications of fuzzy set theory applied to geographic databases. *Computers, Environment, and Urban Systems*. Vol. 12, No. 2, pp. 89-98.

- Rhind D., 1987, Recent developments in geographical information systems in the UK. *International Journal of Geographical Information Systems*, Vol. 1, No. 3, pp. 229-41.
- Rhind D., 1988, A GIS research agenda. *International Journal of Geographical Information Systems*, Vol. 2, No. 1, pp. 23-8.
- Rosenfield G.H. and Fitzpatrick-Lins K., 1986, A coefficient of agreement as a measure of thematic classification accuracy. *Photogrammetric Engineering and Remote Sensing*, Vol. 52, No. 2, pp. 223-7.
- Skidmore A.K. and Turner B.J., 1988, Forest mapping accuracies are improved using a supervised nonparametric classifier with SPOT data. *Photogrammetric Engineering and Remote Sensing*. Vol. 54, No. 10, pp. 1415-21.
- Stevens S.S., 1946, On the theory of scales of measurement. *Science*, Vol. 103, pp. 677-80.
- Swain P.H. and Davis S.M., 1978, *Remote Sensing - The Quantitative Approach*. New York: McGraw-Hill.
- Thapa K. and Bossler J., 1992, Accuracy of spatial data used in geographic information systems. *Photogrammetric Engineering and Remote Sensing*, Vol. 58, No. 6, pp. 835-41.
- Thapa K. and Burtch R.C., 1991, Primary and secondary methods of data collection in GIS/LIS. *Surveying and Land Information Systems*, Vol. 51, No. 3, 1991, pp. 162-70.
- Theobald D.M., 1989, Accuracy and bias issues in surface representation. In: Goodchild M.F. and Gopal S.(eds.), *Accuracy of Spatial Databases*, London: Taylor and Francis, pp. 99-106.
- Tomlinson R.F., 1987, Current and potential uses of geographical information systems: The North American experience. *International Journal of Geographical Information Systems*, Vol. 1, No. 3, pp. 203-18.
- Torlegard A.K.I., 1986, Some photogrammetric experiments with digital image processing. *Photogrammetric Record*, Vol. 12, No. 68., pp. 175-96.
- Townshend J.R.G., 1981, *Terrain Analysis and Remote Sensing*. London: George Allen and Unwin.
- Townshend J.R.G., 1992, Land cover. *International Journal of Remote Sensing*, Vol. 13, No. 6and7, pp. 1319-28.

- van der Wel F.J.M., Hootsmans R.M. and Ormeling F., 1994, Visualisation of data quality. In: MacEachren A.M. and Taylor F. (eds.), *Visualisation in Modern Cartography*, Oxford: Pergamon, pp. 313-31.
- Veregin H., 1989, Error modelling for the map overlay operation. In: Goodchild M.F. and Gopal S. (eds.), *Accuracy of Spatial Databases*, London: Taylor and Francis, pp. 3-18.
- Veregin H., 1995, Developing and testing of an error propagation model for GIS overlay operations. *International Journal of Geographical Information Systems*, Vol. 9, No. 6, pp. 595-619.
- Vonderohe A.P. and Chrisman N.R., 1985, Tests to establish the quality of digital cartographic data: some examples from the Dane county land records project. In: *Proceedings Auto-Carto 7*, pp. 552-9.
- Walsh S.J., Lightfoot D.R. and Butler D.R., 1987, Recognition and assessment of error in GIS. *Photogrammetric Engineering and Remote Sensing*, Vol. 53, No.10, pp. 1423-30.
- Wang F., 1990, Improving remote sensing image analysis through fuzzy information representation. *Photogrammetric Engineering and Remote Sensing*, Vol. 56, No. 8, pp. 1163-9.
- Wang, F. and Hall G.B., 1996, Fuzzy representation of geographical boundaries in GIS. *International Journal of Geographical Information Systems*, Vol. 10, No. 5, pp. 573-90.
- Weber C., 1994, Per-zone classification of urban land cover for urban population estimation. In: Foody and Curran (eds.), *Environmental Remote Sensing from Regional to Global Scales*, Chichester: John Wiley and Sons, pp. 142-9.
- Webster R. and Oliver M.A., 1990, *Statistical Methods in Soil and Land Resources Survey*. New York: Oxford University Press.
- Welch R.A., 1987, Integration of photogrammetric, remote sensing and database technologies for mapping applications. *Photogrammetric Record*, Vol. 12, No. 70, pp. 409-28.
- Wolf P.R., 1987, *Adjustment Computations: practical least squares for surveyors*. Landmark Enterprises.
- Wolf P.R., 1988, *Elements of Photogrammetry*, 2nd edition. McGraw-Hill.
- Wood T.F. and Foody G.M., 1989, Analysis and representation of vegetation continua from Landsat Thematic Mapper data for lowland heaths. *International Journal of Remote Sensing*, Vol. 10, No. 1, pp. 181-9.

Wood T.F. and Foody G.M., 1993, Using cover-type likelihoods and typicalities in a geographic information system data structure to map gradually changing environments. In: Haines-Young R., Green D.R. and Cousins S.H. (eds.), *Landscape ecology and GIS*, London: Taylor and Francis, pp. 141-6.

Zadeh L.A., 1965, Fuzzy sets. *Information and Control*, Vol. 8, pp. 338-353.

## Appendix 1. The list of ground control points by land surveying

ID	X (National Grid coordinates)	Y (National Grid coordinates)	Z (above sea level O.D.)	Type
1			63.03	Z
2			67.96	Z
3			69.02	Z
4			70.11	Z
5			81.95	Z
6			82.03	Z
7			78.50	Z
8			72.70	Z
9	326774.66	670954.20	73.52	XYZ
10			64.27	Z
11			60.90	Z
12	326650.02	671369.59	59.27	XYZ
13			65.65	Z
14			57.71	Z
15			57.76	Z
16			58.87	Z
17			61.58	Z
18			62.79	Z
19			72.81	Z
20			77.39	Z
21			70.70	Z
22			69.67	Z
23			75.55	Z
24			76.73	Z
25			69.78	Z
26			67.30	Z
27			75.35	Z
28			78.17	Z
29			87.19	Z
30			106.48	Z
31			96.58	Z
32			88.40	Z
33			82.15	Z
34			75.77	Z
35			80.30	Z
36			79.99	Z
37	324820.02	671308.53	81.92	XYZ
38			98.28	Z
39			94.45	Z
40			68.15	Z

/cont.

ID	X	Y	Z	Type
	(National Grid coordinates)		(above sea level O.D.)	
41			58.19	Z
42			57.53	Z
43			58.95	Z
44			73.06	Z
45			70.31	Z
46	327359.87	669707.36	94.81	XYZ
47			91.64	Z
48			98.33	Z
49			105.77	Z
50			106.95	Z
51			107.91	Z
52	326449.05	669576.67	121.73	XYZ
53			115.61	Z
54			127.10	Z
55			127.35	Z
56	325280.12	670053.00	129.69	XYZ
57			136.54	Z
58			135.47	Z
59	326911.42	670444.41		XY
60	325906.33	671107.08		XY
61	327321.52	671113.93		XY
62	327356.87	669694.63		XY
63	326541.28	669555.84		XY
64	326031.50	669860.13		XY
65	327249.78	671084.66		XY
66	326413.63	669604.50		XY
67	326019.57	669849.50		XY

**Appendix 2. The list of densified control points by photogrammetric block adjustment**

ID	X (m)	Y(m)	Z (m)
	(National Grid coordinates)		(above sea level O.D.)
Model 1.			
101	324795.16	670343.50	104.85
109	324981.72	670358.19	105.66
30	325019.53	670417.76	106.48
105	325355.50	670485.91	104.40
27	325275.99	670975.11	75.35
28	325098.76	670923.05	78.17
29	325135.50	670774.06	87.19
32	324870.17	670724.18	88.40
31	324848.84	670605.78	96.58
33	324858.08	670840.87	82.15
34	324865.24	671015.56	75.77
36	324993.58	671253.66	79.99
37	324820.02	671308.53	81.92
108	325164.05	671349.56	79.92
91	325395.95	671121.36	77.47
35	324838.91	671226.84	80.30
24	325444.66	671261.37	76.73
Model 2.			
23	325527.07	671269.49	75.55
171	325512.65	670590.13	148.37
69	325439.30	670643.93	164.76
110	325809.88	670584.44	130.45
120	325804.71	670917.95	90.02
26	325638.35	670988.98	67.30
25	325444.86	670984.11	69.78
21	325837.67	671298.82	70.70
114	325767.37	671413.69	73.69
132	326266.91	671350.27	61.62
Model 3.			
2	326243.44	671136.03	67.96
166	325876.35	670531.88	121.86
111	325848.39	670397.00	119.31
61	325906.34	671107.080	67.92
22	325883.19	671182.97	69.67
38	326071.70	670771.54	98.28
39	325972.18	670843.01	94.45
5	326291.46	670851.02	81.95

/cont.

ID	X (m)	Y(m)	Z (m)
	(National Grid coordinates) (above sea level O.D.)		

123	326198.39	670600.05	102.24
4	326036.49	671001.92	70.11
3	326028.76	671033.89	69.02
1	326419.40	671250.43	63.03

Model 4.

11	326661.65	671303.20	60.90
144	326408.23	670479.22	82.55
143	326674.09	670409.04	74.39
12	326650.02	671369.59	59.27
6	326474.45	670803.82	82.03
7	326665.54	670787.33	78.50
10	326706.19	671146.14	64.27
66	327249.78	671084.66	55.83

Model 5.

42	327004.49	671345.35	57.53
60	326911.42	670444.41	115.21
13	326994.34	670530.46	65.65
14	327129.65	670317.53	57.71
152	327247.86	670566.99	52.20
43	327180.74	671024.70	58.95
41	327116.63	671141.28	58.19
40	326904.04	671023.62	68.15
9	326774.66	670954.20	73.52
8	326789.47	670879.27	72.70
78	327089.07	671391.57	56.49

Model 6.

182	325138.32	669754.34	160.45
190	325004.75	670095.59	124.70
56	325280.13	670053.00	129.69
57	325495.93	669981.07	136.54
174	325615.57	669686.52	145.52
169	326016.39	669609.44	148.36

Model 7.

170	325655.44	670281.73	71.96
58	325747.96	669923.78	135.47
172	325874.41	669634.35	144.56
168	325901.73	670548.26	120.25
167	325945.11	670278.06	109.17
81	326089.28	670192.22	63.58

/cont.



ID	X (m)	Y(m)	Z (m)
	(National Grid coordinates) (above sea level O.D.)		

54	326089.37	669841.99	127.10
55	326018.39	669845.83	127.35
68	326019.57	669849.50	128.22
18	326186.21	670168.89	62.79

Model 8.

17	326272.34	670163.15	61.58
52	326449.05	669576.67	121.73
16	326565.88	670178.67	58.87
67	326413.63	669604.50	125.25
50	326979.30	669628.81	106.95

Model 9.

15	326809.32	670194.17	57.76
159	326981.06	669928.17	80.23
53	326690.38	669499.28	115.61
51	326949.28	669566.78	107.91
20	327122.26	669934.54	77.39

Model 10.

45	327283.72	670080.50	70.31
154	327533.56	670444.61	56.36
49	327074.03	669536.03	105.77
48	327092.18	669704.49	98.33
47	327213.68	669763.18	91.64
46	327359.87	669707.36	94.81
44	327336.93	669970.18	73.06

**Appendix 3. The class statistics used in the supervised fuzzy classification described in section 6.2.1**

1. SPOT HRV image

Classes		band 1	band 2	band 3	
grass-I	mean vector		154.19	178.64	72.12
	variance and	band 1	109.88	71.68	105.34
	covariance	band 2	71.68	53.32	66.50
	matrix	band 3	105.34	66.50	154.06
grass-II	mean vector		145.82	166.86	141.54
	variance and	band 1	18.39	11.12	-1.63
	covariance	band 2	11.12	16.34	3.06
	matrix	band 3	-1.63	3.06	30.92
built	mean vector		188.98	196.51	227.89
	variance and	band 1	1424.34	1364.28	581.80
	covariance	band 2	1364.28	1398.07	651.28
	matrix	band 3	581.80	651.28	457.57
wood	mean vector		106.63	113.58	69.27
	variance and	band 1	95.10	73.25	-101.26
	covariance	band 2	73.25	76.81	-104.90
	matrix	band 3	-101.26	-104.90	262.53
shrub	mean vector		122.04	124.41	96.32
	variance and	band 1	124.27	101.57	15.67
	covariance	band 2	101.57	117.06	16.36
	matrix	band 3	15.67	16.36	92.94
water	mean vector		40.50	38.40	146.60
	variance and	band 1	12.90	11.08	28.14
	covariance	band 2	11.08	29.39	24.73
	matrix	band 3	28.14	24.73	80.63

2. Landsat TM image

Classes		band 3	band 4	band 5	
grass-I	mean vector		26.55	105.07	26.15
	variance and	band 3	4.89	-7.76	7.46
	covariance	band 4	-7.76	67.30	-4.79
	matrix	band 5	7.46	-4.79	15.21
grass-II	mean vector		30.33	68.44	28.33
	variance and	band 3	2.18	-2.91	0.74
	covariance	band 4	-2.91	7.16	-0.67
	matrix	band 5	0.74	-0.67	0.67
built	mean vector		42.38	49.75	37.72
	variance and	band 3	26.86	-2.16	18.59
	covariance	band 4	-2.16	23.29	-3.71
	matrix	band 5	18.59	-3.71	16.53
wood	mean vector		23.60	65.54	18.11
	variance and	band 3	3.48	4.38	1.41
	covariance	band 4	4.38	79.72	14.00
	matrix	band 5	1.41	14.00	5.19
shrub	mean vector		26.15	71.44	18.51
	variance and	band 3	4.08	-3.94	-0.65
	covariance	band 4	-3.94	54.44	13.78
	matrix	band 5	-0.65	13.78	4.51
water	mean vector		22.77	22.92	11.38
	variance and	band 3	2.52	1.94	0.59
	covariance	band 4	1.94	5.37	0.32
	matrix	band 5	0.59	0.32	2.66

## Appendix 4. The FORTRAN 77 program for the fuzzy *c*-means clustering

```
c-----
c This program was written and tested by J Zhang in November 1994. It is based
c on a program by J Bezdek et al. (1984), designed for fuzzy c-means clustering.
c-----
c description of variables and arrays:
c   nn - total number of pixels
c   n0 - actual number of pixels (within bounds)
c   nb - number of spectral bands (attributes)
c   nc - number of clusters desired
c   mc - the maximum number of clusters (classes)
c   mb - the maximum number of spectral dimensions
c   iu - integer variable for choice of initial u-matrix
c   d - distance array (squared)
c   y - image array
c   temp - for character string processing flexibility
c   tmp - temporary control character
c   u1,u2,v - arrays named in accordance with the fcm algorithm
c   tf - logical variable for choosing Zhang/Bezdek's programme
c-----
parameter (nn=25*41,mb=7,mc=15,nb=2)
character fcmfn*12,tmp,temp*47
integer y(nn,mb)
real d(nn,mc),u1(nn,mc),u2(nn,mc),v(mc,mb),
+   cov(mb,mb,mc),incov(mb,mb,mc),
+   mat1(nb,nb),mat2(nb,nb),mat3(mb)
logical tf
1  write(*,'(1x,a30)')
+ ' 1 - switch on supervised mode'
  write(*,'(1x,a34)')
+ ' 2 - set initial u matrix randomly'
  write(*,'(1x,a34)')
+ ' 3 - set random non-fuzzy u matrix'
  write(*,'(1x,a46)')
+ ' 4 - set an almost uniform u matrix as a start'
  read(*,'(i1)') iu
  if(iu.lt.1 .or. iu.gt.4) goto 1
  tf=.false.
  if (iu.eq.1) then
    write(*,'(1x,a39)')
+ 'use per-category v,u calculation (y/n)?'
    read(*,'(a)') tmp
    if (tmp.eq.'y' .or. tmp.eq.'y') tf=.true.
  end if
  write(*,'(1x,a34)')'please enter the class number - nc:'
```

```

read(*,'(i2)') nc
rm=2.5
pp=1.0/(rm-1.0)
thrshd=0.001
lmax=1000
c -----
c  read image data
c-----
      i=1
      open(1, file='clstr.dat', status='old')
5     read(1, '(i2,1x,i2)', end= 10) y(i,1),y(i,2)
      i=i+1
      goto 5
10    n0=i-1
      close(1)
      if (n0.eq. 18) then
          write (*,'(1x,a22)') 'data points confirmed!'
      endif
      do 65 i=1, n0
          if (iu.eq. 3) then
              x=6.
c-----because there are 6 classes-----
              call random(ix,x,i)
              write(*,'(1x,i5)') ix
              endif
              do 60 j=1,nc
                  if (iu.eq.3) then
                      u1(i,j)=0.0
                      u2(i,j)=0.0
                      if (j.eq.ix) then
                          u1(i,j)=1.0
                          u2(i,j)=1.0
                      end if
                  else if (iu.eq.2) then
                      x=16384.
                      itp=i*3+j
                      call random(ix,x,itp)
                      e=real(ix)/16384.0
                      write(*,'(1x,f9.2)') e
                      u1(i,j)=e
                      u2(i,j)=e
                  else if (iu.eq.4) then
                      x=16384.
                      itp=i*3+j
                      call random(ix,x,itp)
                      e=real(ix)/1638400.0
                      u1(i,j)=1.0/real(nc)-e

```

```

        u2(i,j)=1.0/real(nc)-e
        write(*,'(1x,f9.2)') e
    end if
60    continue
65    continue
c-----
c  read covariance matrix per category
c-----
    if (tf .eq. .true.) then
        open(1, file='sign.dat', status='old')
        do 125 i=1,nc
            read(1,'(1x,2f9.2)') (v(i,i1),i1=1,2)
            write(*,'(1x,2f9.2)') (v(i,i1),i1=1,2)
            read(1,'(1x,2f9.2)') ((cov(i1,i2,i),i2=1,2),i1=1,2)
            write(*,'(1x,2f9.2)') ((cov(i1,i2,i),i2=1,2),i1=1,2)
            write(*,'(1x,a35)') '-----'
c        read(*,'(a)') tmp
125    continue
        close(1)
c -----
c  calculation of cov matrix
c -----
        else
            write(*,'(1x,a21)')'for the whole sample:'
            write(*,'(1x,a38)')'calculating variance-covariance matrix'
            do 13 j=1,nb
                mat3(j)=0.0
                do 12 i=1,n0
                    mat3(j)=mat3(j)+y(i,j)
12            continue
13            mat3(j)=mat3(j)/n0
                do 15 i=1,nb
                    do 15 j=1,nb
                        cov(i,j,1)=0.0
                        do 16 k=1,n0
16                cov(i,j,1)=cov(i,j,1)+(y(k,i)-mat3(i))*(y(k,j)-mat3(j))
15            cov(i,j,1)=cov(i,j,1)/n0
                write(*,'(1x,2f9.2)') (mat3(j),j=1,2)
                write(*,'(1x,2f9.2)') ((cov(i,j,1),j=1,2),i=1,2)
            end if
c-----
c  calculating the inverse of covariance
c-----
            write(*,'(1x,a33)')'calculating inverse of cov matrix'
            if (tf.eq..true.) then
                itp=nc
            else

```

```

        itp=1
    end if
    do 126 l=1,itp
        do 127 i=1,nb
            do 127 j=1,nb
                mat1(i,j)=cov(i,j,l)
127    continue
        call inverse(nb,mat1,mat2)
        do 128 i=1,nb
            do 128 j=1,nb
                incov(i,j,l)=mat2(i,j)
128    continue
        write(*,'(1x,2f9.2)') ((incov(i,j,l),j=1,2),i=1,2)
        do 129 i=1,nb
            do 529 j=1,nb
                tl=0.0
                do 121 k=1,nb
121    tl=tl+cov(i,k,l)*incov(k,j,l)
                mat1(i,j)=tl
529    continue
129    continue
        write(*,'(1x,a19)')'check for identity:'
        do 530 i=1,nb
            do 530 j=1,nb
                mat2(i,j)=0.0
                if (j.eq.i) mat2(i,j)=1.0
                if (abs(mat1(i,j)-mat2(i,j)) .gt. 0.001) then
                    write(*,'(1x,a37)')
+                'matrix inversion calculation failure!'
                end if
530    continue
        write(*,'(1x,2f9.2)') ((mat1(i,j),j=1,2),i=1,2)
        write(*,'(1x,a10)')'-----'
c    read(*,'(a)') tmp
126 continue
c-----
c fuzzy c-means clustering algorithm beginning
c-----
        it=0
105 continue
        if (iu.eq.1) goto 399
        it=it+1
        if (it.gt.lmax) goto 1000
        write(*,'(//1x,a20,i5/)') '----- iteration -----',it
        write(*,'(1x,a16)')'--- v matrix ---'
        do 120 i1=1,nc
            do 115 i0=1,nb

```

```

    tp0=0.0
    tp1=0.0
    tp2=0.0
    do 110 i2=1,n0
        tp0=u1(i2,i1)**(rm)
        tp1=tp1+tp0
        tp2=tp2+y(i2,i0)*tp0
110    continue
        if (tp1.eq. 0.) then
            write(*,'(1x,a34)') 'divered by zero, v - matrix error!'
            stop
        else
            v(i1,i0)=tp2/tp1
        end if
115    continue
        write(*,'(1x,i5,3f9.2)') i1,(v(i1,j),j=1,nb)
c      write(*,'(1x,a10)') '-----'
c      read(*,'(a)') tmp
120    continue
c-----
c   calculating squared mahalanobis distance d-matrix
c-----
399 do 131 j=1,nc
    itp=j
    if(tf .eq. .false.) itp=1
    do 136 i=1,n0
        do 132 k=1,nb
            tl=0.
            do 133 l=1,nb
                tap=y(i,l)-v(j,l)
                tl=tl+tap*incov(l,k,itp)
133        continue
                mat3(k)=tl
132    continue
            tt=0.
            do 134 k=1,nb
                tap=y(i,k)-v(j,k)
                tt=tt+tap*mat3(k)
134        continue
                d(i,j)=tt+0.000001
136    continue
c      write(*,'(1x,i5,6f9.2)')i,(d(i,j),j=1,nc)
131    continue
        write(*,'(1x,a28)')'end of d-matrix calculation.'
c-----
c   modified u matrix
c-----

```



```

do 130 i1=1, n0
  do 130 i2=1,nc
    tp0=0
    tp1=(1.0/d(i1,i2))**(pp)
    do 135 i3=1,nc
      tp0=tp0+(1.0/d(i1,i3))**(pp)
135  continue
      u2(i1,i2)=tp1/tp0
130  continue
140  continue
      exp=0.0
      do 340 i=1,n0
        do 340 j=1,nc
          tap=abs(u2(i,j)-u1(i,j))
          if(tap.gt.exp) exp=tap
340  continue
          h=0.0
          f=0.0
          do 500 i=1,nc
            do 500 k=1,n0
              h=h+u2(k,i)*alog(u2(k,i)+0.000001)
500  f=f+u2(k,i)*u2(k,i)
          f=f/real(n0)
          h=-h/real(n0)
          rjm=0.0
          do 510 i=1,nc
            do 510 k=1,n0
              rjm=rjm+d(k,i)*(u2(k,i)**(rm))
510  continue
          write(*,'(1x,a15,f9.3)')
          + 'minimising jm :',rjm
          write(*,'(1x,a22,2f9.3/)')
          + 'validity fuctions f,h:',f,h
          if (iu.eq.1) goto 1000
          if (exp .gt.thrshd .or. it.eq.1) then
            do 150 i1=1,n0
              do 150 i2=1,nc
                u1(i1,i2)=u2(i1,i2)
150  continue
              goto 105
            end if
            goto 1000
999  write(*,'(1x,a19)')'error in file i/o !'
1000 continue
c-----
c  write out the clustering output
c-----

```

```

write(*,'(1x,a42)') 'please enter the file name to store output'
read(*,'(a12)') fcmfn
open(1,file=fcmfn,status='new')
write(1,'(1x,a19,1x,i2)') 'spectral dimension:',nb
write(1,'(1x,a19,1x,i2)') 'number of clusters:',nc
write(1,'(1x,a29,i1)') 'u matrix initialised by mode:',iu
write(1,'(1x,a7,1x,3f10.2)') 'jm,f,h:',rjm,f,h
do 210 i=1,nc
  write(1,'(1x,2f9.2)') (v(i,j),j=1,nb)
  write(*,'(1x,2f9.2)') (v(i,j),j=1,nb)
210 continue
do 200 i1=1,n0
  write(1,'(1x,3f7.1)',err=999)
+ (int(1000*u2(i1,i2))/10.0,i2=1,nc)
  write(*,'(1x,3f7.1)') (int(1000*u2(i1,i2))/10.0,i2=1,nc)
200 continue
close(1)
end

c-----
subroutine inverse(m,a,b)
c-----
dimension a(m,m),b(m,m)
do 5 i=1,m
  do 5 j=1,m
    b(i,j)=0
    if(j.eq.i) b(i,j)=1.
5 continue
do 100 loop=1,m
  ab=abs(a(loop,loop))
  id=loop
  do 10 i=loop,m
    if(abs(a(i,loop)).gt. ab) then
      ab=abs(a(i,loop))
      id=i
    end if
10 continue
if (ab.lt. 0.0000001) then
  write(*,'(1x,a28)') 'singular matrix encountered!'
  stop
end if
if (id.eq.loop) goto 35
do 30 j=1,m
  tb=b(loop,j)
  ta=a(loop,j)
  b(loop,j)=b(id,j)
  a(loop,j)=a(id,j)
  b(id,j)=tb

```

```

    a(id,j)=ta
30  continue
35  continue
    tap=a(loop,loop)
    do 40 j=1,m
        b(loop,j)=b(loop,j)/tap
        a(loop,j)=a(loop,j)/tap
40  continue
    do 50 i=loop+1,m
        tap=a(i,loop)
        do 50 j=1,m
            b(i,j)=b(i,j)-tap*b(loop,j)
            a(i,j)=a(i,j)-tap*a(loop,j)
50  continue
100 continue
    do 300 loop=m-1,1,-1
        do 200 i=loop,1,-1
            tap=a(i,loop+1)
            do 200 j=1,m
                b(i,j)=b(i,j)-tap*b(loop+1,j)
                a(i,j)=a(i,j)-tap*a(loop+1,j)
200 continue
300 continue
    return
end

```

```

c-----
subroutine matmpl(m,n,l,a,b,c)

```

```

c-----
    dimension a(m,n),b(n,l),c(m,l)
    do 10 i=1,m
        do 10 j=1,l
            tl=0.
            do 20 k=1,n
20      tl=tl+a(i,k)*b(k,j)
            c(i,j)=tl
10  continue
    return
end

```

```

c-----
subroutine random(ix,y,n)

```

```

c-----
    ibig=16384
    iseed=15625
    new=iseed
    do 10 i=1,n
        new=new*iseed
        new=mod(new,ibig)

```

```
    if (new.lt.0) new=ibig+new
10  continue
    ix=new*y/real(ibig)+1
    if (ix.lt.1 .or. ix .gt.y) then
        print *, 'random number generator fault!'
        stop
    else
c   print *, 'next number',irnd
    end if
    end
```

## Appendix 5. The FORTRAN 77 program for deriving proportions of sub-pixel component land cover types

```

c-----
c   This program was designed for deriving ground truth data from fine
c   resolution grid file PHGREF.GIS
c   written and tested by J Zhang in November 1994.
c   GIS - GIS array
c   TEMP - For character string processing flexibility
c   TMP - Temporary contrl character
c-----

      parameter (nc=5)
      character gisnam*12,output*12,tmp,temp*47,yn1*12,yn2*12
      integer gis(0:nc),gis0(0:nc)
      n0=25*41
      write(*,'(1x,a36)')'please enter file name for gis data:'
      read(*,'(a12)') gisnam
      write(*,'(1x,a30)')'please enter output file name:'
      read(*,'(a12)') output
      write(*,'(1x,a31)')'please specify the window size:'
      read(*,'(i5)') ips

c -----
c   read image/gis data out
c-----
c   inquire(file=gisnam,sequential=yn1,formatted=yn2,recl=iyn3)
c   write(*,'(1x,a12,a12,i5)') yn1,yn2,iyn3
c   stop
      open(1,file=output,access='direct',recl=12,status='new')
      open(2,file=gisnam,status='old')
31  format(a17)
41  format(2i5,i6)
      read(2,fmt=31,err=999,end=100) temp(1:17)
      read(2,'(a)',err=999,end=100) tmp
      read(2,fmt=31,err=999,end=100) temp(1:17)
      read(2,fmt=31,err=999,end=100) temp(1:17)
      write(*,'(1x,a32)')'gis file headers read out.'
      nr0=0
      line=1
21  do 55 l=1,ips
      np=1
20  do 30 i=0,nc
      gis(i)=0
30  continue
      do 45 i=1,ips
      read(2,fmt=41,end=100) ix,iy,ic
      if (i.eq.1) then
      ix0=ix

```

```

        iy0=iy
    else
        if (iy.ne.iy0) then
            backspace(2)
            icon=1
            goto 50
        end if
    end if
    gis(ic)=gis(ic)+1
45  continue
    icon=0
50  read(1,rec=nr0+np,err=56) (gis0(j),j=0,5)
c   if (nr0+np.eq.1) then
c     write(*,'(7i5)') nr0+np,(gis0(j),j=0,5)
c   end if
    do 57 j=0,5
        gis(j)=gis(j)+gis0(j)
57  continue
56  write(1,rec=nr0+np) (gis(j),j=0,5)
c   if (nr0+np.eq.1) then
c     if (l.eq.ips) then
c       write(*,'(1x,7i5)') nr0+np,(gis(j),j=0,5)
c     end if
        if (l.eq.ips) then
            dd=0.0
            do 70 j=0,5
70         dd=dd+gis(j)
            if (dd.eq.ips*ips) write(*,'(1x,a11,2i5)')
+           'full pixel:',np,line
            end if
            if (icon.eq.0) then
                np=np+1
                goto 20
            end if
55  continue
        nr0=nr0+np
        write(*,'(1x,a28,2i5)')
+         'one window line (no.,recno.)',line,nr0
        line=line+1
        goto 21
999 write(*,'(1x,a9)') 'i/o error'
        goto 101
100 write(*,'(1x,a15)') 'end of gis file'
101 close(1)
    close(2)
end

```

## Appendix 6. The FORTRAN 77 program for transforming point sample data to GSLIB data format

```
c-----
c This program was designed as a transformation procedure to generate
c GSLIB data file from ERDAS DIGSCRN point sampling data files.
c The input data are binary data (corresponding to classes 1,2,3,4 and 5,
c stored in files point1/2/3/4/5.dat), the output data file will be in the
c format illustrated below:
c Title.....
c 7
c xlocation
c ylocation
c grass
c built
c wood
c shrub
c water
c x y var1 var2 var3 var4 var5
c .....
c Written and tested by J Zhang in June 1995.
-----
* Version 23rd June 1995
  character*12 fromdf, todf
  character tmp*3
  character PAUSE,STR*25,CH1*40
  dimension node(5)
  write(*,'(1x,a35)') 'Please enter the to_data file name:'
  read(*,'(a12)') todf
  open(2,file=toadf,status='new')
  write(*,'(1x,a34)') 'Please specify the No. of classes:'
  read(*,'(i2)') iclass
  write(*,'(1x,a40)') 'Please enter the grid cell size (dx,dy):'
  read(*,*) dx
  read(*,*) dy
  write(2,'(1x,a43)')
+ 'GSLIB data file relevant to land cover data'
  write(2,'(1x,i2)') iclass+2
  write(2,'(1x,a9)') 'xlocation'
  write(2,'(1x,a9)') 'ylocation'
  do 5 i=1,iclass
    write(2,'(1x,a6,i2)') 'class_',i
5  continue
  write(*,'(1x,a37)') 'Please enter the from_data file name:'
  read(*,'(a12)') fromdf
  open(1,file=fromdf,status='old')
```

```

13  format(5I5)
14  format(F12.2,F12.2,A1)
15  format(4I5,4X,A1)
21  read(1,fmt=13,err=999)npt,igis,itmp,mod,npts
    read(1,fmt=14,err=999) xmin,ymin,ch
    write(*,'(1x,2f14.6)') xmin,ymin
    read(1,fmt=14,err=999) xmin,ymax,ch
    write(*,'(1x,2f14.6)') xmin,ymax
    read(1,fmt=14,err=999) xmax,ymax,ch
    write(*,'(1x,2f14.6)') xmax,ymax
    read(1,fmt=14,err=999) xmax,ymin,ch
    write(*,'(1x,2f14.6)') xmax,ymin
    write(*,'(a14)') 'end of header!'
    write (*,'(1x,38a)') 'please specify xmin, ymin as required:'
    read(*,*) xmin
    read(*,*) ymin
30  read(1,fmt=15,err=999,end=35) npt,igis,itmp,mod,ch
    do 31 j=1,5
        node(j)=0
        if (j.eq.igis) node(j)=1
31  continue
    do 40 i=1,npt
        read(1,fmt=14,err=999) x,y,ch
        nc=1+ifix((x-xmin)/dx)
        nr=1+ifix((y-ymin)/dy)
        write(2,'(2x,i3,1x,i3,5i2)') nc,nr,(node(j),j=1,5)
40  continue
    goto 30
35  write(*,'(1x,a22)') 'end of from_file read!'
    close(1)
    close(2)
    write(*,'(1x,a26)') ' end of data file transfer!'
    goto 1000
999 write(*,'(1x,a15)') 'file i/o error!'
1000 continue
    end

```



## Appendix 7. The FORTRAN 77 program for transferring kriging outputs to ASCII format suitable for ERDAS LAN data files

```
-----  
c This program works to transfer the outputs from indicator kriging procedures  
c in GSLIB to ASCII format suitable for ERDAS LAN data files, which can then  
c be read into ARC/INFO GRID and ARCPLOT modules to facilitate surface data  
c manipulation and visualisation.  
c Witten and tested by J Zhang July 1995  
-----  
      parameter (nx1=75,ny1=122,nx2=25,ny2=41,nc=5)  
      character ksurlan*12,krigfn*12,tmp,temp*66  
      dimension ksurdat(nx1,ny1,nc)  
-----  
      write(*,'(1x,a52)')  
+ 'please remember resultant lan-ascii data file name:'  
      read(*,'(a12)')ksurlan  
      write(*,'(1x,a)')  
+ 'is the kriging output for tm or spot (t/s) ?'  
      read(*,'(a1)') tmp  
      nx=nx1  
      ny=ny1  
      if (tmp.eq. 't' .or. tmp.eq.'t') then  
          nx=nx2  
          ny=ny2  
      end if  
      do 100 i=1,nc  
          write(*,'(1x,a40)')  
+ 'please enter file name for kriging data:'  
          write(*,'(1x,a7,i3)') 'class -',i  
          read(*,'(a12)') krigfn  
          open(1,file=krigfn,status='old')  
          read(1,'(a66)') temp  
          read(1,'(a1)') tmp  
          read(1,'(a8)') temp(1:8)  
          read(1,'(a20)') temp(1:20)  
          ipt=0  
          ix=1  
          iy=ny  
110 ksurdat(ix,iy,i)=-9  
115 read(1,'(f8.3,1x,f8.3)',err=999,end=120) est,var  
          write(*,'(1x,2f9.3)') est,var  
          if (est.eq. -999.0) est=-9  
          if (est.gt.1.0) est=1.0  
          if (est.lt.0.0 .and. est.gt. -9) est=0.0  
          if (est.ne. -9) then  
              ksurdat(ix,iy,i)=ifix(100.0*est+0.5)
```

```

end if
write*,'(1x,3i5)' ix,iy,ksurdat(ix,iy,i)
ipt=ipt+1
ix=ix+1
if (ipt.eq.nx) then
  iy=iy-1
  ix=1
  ipt=0
end if
if (iy.lt.1) goto 115
goto 110
120 continue
write*,'(1x,a32,i5)'
+ 'end of kriging datafile reading ',iy
close(1)
100 continue
c-----
30 format(1x,a41)
31 format(a17)
41 format(2i5,5i6)
open(1,file=ksurlan,status='new')
write(1,'(1x,a5,a12)',err=999) 'f1 : ',ksurlan
write(1,'(1x,a)',err=999) ' '
write(1,fmt=31,err=999) '          f1'
write(1,fmt=30,err=999)
+ ' x y b1 b2 b3 b4 b5'
do 121 i=1,ny
do 121 j=1,nx
  irec=(i-1)*nx+j
  write(1,fmt=41,err=999) j,i,(ksurdat(j,i,k),k=1,nc)
121 continue
write*,'(1x,a36)' 'end of lan-ascii data file writing!'
c-----
goto 1000
999 write*,'(1x,a15)' 'file i/o error!'
goto 1000
1000 continue
close(1)
end

```