Toward Verification of a Natural Resource Uncertainty Model

by

Trevor John Davis
B.Sc., University of Victoria, 1992
M.Sc., University of Victoria, 1994

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES
(Department of Geography)

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

June, 1999

Department of Geography

The University of British Columbia
Vancouver, Canada

Date July 4/99

DE-6 (2/88)

Abstract

Natural resource management models simplify reality for the purpose of planning or management. In much the same way, an uncertainty model simplifies the many uncertainties that pervade the natural resource management model. However, though a number of uncertainty models have been developed, there has been little work on verifying such models against the uncertainty they purport to represent. The central research question addressed by this work is 'can a natural resource management uncertainty model be verified in order to evaluate its utility in real-world management?' Methods to verify uncertainty models are developed in two areas: uncertainty *data* models, and uncertainty *propagation* through process models. General methods are developed, and then applied to a specific case study: slope stability uncertainty in the southern Queen Charlotte Islands. Verification of two typical uncertainty data models (of classified soils and continuous slope) demonstrates that (in this case) both expert opinion inputs and published error statistics underestimate the level of uncertainty that exists in reality. Methods are developed to recalibrate the data models, and the recalibrated data are used as input to an uncertainty propagation model. Exploratory analysis methods are then used to verify the output of this model, comparing it with a high-resolution mass wastage database--itself developed using a new set of tools incorporating uncertainty visualisation. Exploratory data analysis and statistical analysis of the verification shows that, given the nature of slope stability modelling, it is not possible to directly verify variability in the model outputs due to the existing distribution of slope variability (based on the nature of slope modelling). However, the verification work indicates that the information retained in uncertainty-based process models allows increased predictive accuracy--in this case of slope failure. It is noted that these verified models and their data increase real-world management and planning options at all levels of resource management. Operational utility is demonstrated throughout this work. Increased strategic planning utility is discussed, and a call is made for integrative studies of uncertainty model verification at this level.

Table of Contents

Abstract ii
List of Tables vii
List of Figures viii
Acknowledgements x

Chapter One: Introduction 1
1.1. The Problem 1
1.2. Major Questions 3
1.3. Research Organisation 5
1.4. Contribution to Knowledge 6

Chapter Two: Background 8
2.1. Introduction 8
2.1.1. GIS and Uncertainty 9
2.1.2. Chapter Layout 10
2.2. Error and Uncertainty 10
2.2.1. Uncertainty Defined 11
2.2.2. Quality 12
2.2.3. Subdivisions of Uncertainty 13
2.2.3.1. Positional Uncertainty 14
2.2.3.1.1. Registration 14
2.2.3.1.2. Other Sources 15
2.2.3.1.3. Lines and Areas 15
2.2.3.2. Attribute Uncertainty 16
2.2.3.3. Temporal Uncertainty 19
2.2.4. Subdivision by Source 20
2.2.4.1. Inherent Uncertainty 20
2.2.4.2. Data Collection and Input Uncertainty 21
2.2.4.3. Data Interpretation 21
2.2.4.4. Data Entry 22
2.2.4.5. Data Manipulation Uncertainty 23
2.2.4.6. Propagation 23
2.2.4.7. Generalisation Issues 24

2.2.5. Measures of Uncertainty 25

2.3. Uncertainty Modelling 26
2.3.1. Modelling 27
2.3.1.1. Modelling with GIS 27
2.3.1.2. Methods of Modelling 28
2.3.1.2.1. Bayesian Probability 29
2.3.1.2.2. Dempster-Shafer's Theory of Evidence 30
2.3.1.2.1. Non-Monotonic Logic 31
2.3.1.2.2. Fuzzy Sets 32
2.3.1.2.3. Linking Fuzzy Sets With Attribute Data 35
2.3.1.2.4. Combining Fuzzy Classifications 36
2.3.1.2.5. Cardinal Values 37
2.3.2. Propagation of Uncertainty 37
2.3.2.1. Arithmetic Propagation 38
2.3.2.1. Monte Carlo 40
2.3.3. Uncertainty in Continuous Data 41
2.3.4. Summary 43

2.4. Communication of Uncertainty 43

2.5. Uncertainty in Forestry Data and Models 45
2.5.1. Uncertainty in Soil and Terrain Modelling 47
2.5.1.1. Soil 47
2.5.1.2. Elevation 48

2.6. Research Gaps 50
2.6.1. Uncertainty Modelling 50
2.6.2. Validation 50
2.6.3. Linking Uncertainty Management with Decision Making 51

2.7. Summary 52

Chapter Three: Modelling and Storing Measures of Uncertainty in Inventory . 54

3.1. Introduction 54

3.2. Methodology 56
3.2.1. The Corridor of Transition Model 57
3.2.1.1. The Semantic Import Model 57
3.2.1.1.1. Polygon Centres 59
3.2.1.1.2. Polygon Boundaries 61
3.2.2. DEM Randomisation 65
3.2.3. Combining Error and Uncertainty 67

3.3. Slope Stability Modelling 67
3.3.1. Combining and Summarising 68

3.4. Case Study 69

3.5. The Boundary Model and Attribute Uncertainty 70

3.6. The Monte Carlo Procedure 71

3.7. Results 73
3.7.1. Problems and Work Required 75
3.7.1.1. Verifying Parameters 75
3.7.1.2. Prediction 76
3.7.1.3. Reporting and Communicating Uncertainty 77

3.8. Summary 78

Chapter Four: Verification of Model Inputs 79

4.1. Introduction 79

4.2. Background 81
4.2.1. Fuzzy Classification 82
4.2.2. Maximum Likelihood 85
4.2.3. Continuous Classes - Fuzzy Clustering 86
4.2.3.1. Background - Fuzzy Clustering 86
4.2.3.2. Applying Fuzzy Clustering to Confirmatory Sampling 89
4.2.3.3. Structure of the Classes in Attribute Space 90
4.2.3.4. Nature of the Sample 92
4.2.3.5. Metrics and Measures for Membership Values 93
4.2.4. Summary of Theoretical Work 97

4.3. Application to Parameter Verification and Tuning 97
4.3.1. Samples Required 98
4.3.2. Methodology 99
4.3.2.1. Cross-Correlation 103
4.3.2.2. Data Summary 106
4.3.3. Results 107
4.3.4. Applying Changes 110

4.4. Discussion 113

4.5. Calibration of Continuous Data 114

4.6. Conclusions 117

Chapter Five: Evaluation of Uncertainty Model Output 120

5.1. Introduction 120

5.2. Methodology 123

5.3. Results 124
5.3.1. Comparison of Means 125
5.3.2. Alternative Realisations 128
5.3.3. Variance 129
5.3.4. Expected vs. Actual 131
5.3.5. Zonal Spatial Limits 133
5.3.6. Spatial Constraints 136

5.3.7. Comparison: Old vs. New ... 136

5.4. Discussion ... 138

5.5. Conclusions ... 140

Chapter Six: Discussion ... 141

6.1. Introduction ... 141

6.2. Resource Management ... 142

6.2.1. Strategic Level ... 143

6.2.2. Tactical Level ... 144

6.2.3. Operational Level ... 146

6.3. Uncertainty Model Validation ... 147

6.4. Further Research ... 148

6.5. Summary ... 154

Chapter Seven: Conclusions ... 155

7.1. Research Timeline ... 155

7.2. Research Questions ... 156

Bibliography ... 160

Appendix A: The Hard k-Means and Fuzzy c-Means Algorithms ... 174

Appendix B: GPS Accuracy Statistics ... 177

Appendix C: Cross Correlograms and Significance Tests for Sample Transects ... 179

Appendix D: Uncertainty Visualisation in the Development of a New Data Update Tool for GIS ... 182

Appendix E: Development of the Mass Wastage Database ... 220

List of Tables

3.1. Misclassification matrices derived from the SI model. 59
3.2. Soil characteristics and estimated standard deviations. 72
4.1. Calculation of the Mahalanobis distance for one sample. 106
4.2. Maximum values of correlation coefficients. 112
4.3. Original and updated misclassification matrices. 113
5.1. Summary statistics for slope stability predictions, based on mean values. 126

List of Figures

1.1. A view of data flow and feedback (verification) loops in resource management. 3
2.1. Location probability of a survey coordinate in 2-D space. 13
2.2. Epsilon boundary model of a digitised line. 15
2.3. Examples of probability density functions for line or digitising error. 16
2.4. Boolean and fuzzy classification models. 34
2.5. Four probabilistic functions of spatial boundary uncertainty. 42
3.1. Alternative centroid models for variable polygon shapes 60
3.2. The variable ridge model of a polygon's centre. 61
3.3. The 'corridor of transition' model for spatial boundary uncertainty. 63
3.4. Perspective view of the fuzzy surface representing soil type 1. 65
3.5. Perspective view of the fuzzy surface representing soil type 1. 65
3.6. Location of the Louise Island study site 70
3.7. Detail of transition corridors. 71
3.8. The three types of surfaces resulting from the uncertainty modelling routine. 73
3.9. Maximum likelihood summary of slope stability factor-of-safety. 75
3.10. The spatial distribution of standard deviation of slope stability factor-of-safety 75
3.11. An example of an application-specific data summary. 75
3.12. The worst-case-scenario summary. 75
4.1. Simplified view of p-dimensional attribute space, fuzzy classes and samples. 84
4.2. Notional distribution of individuals in attribute space. 87
4.3. Hard classes and continuous classes. 87
4.4. Mahalanobis distance in a 3-D attribute space. 88
4.5. Classes viewed as structures in (A,B,C) attribute space. 88
4.6. A new individual at an intergrade position between class A and B. 91
4.7. Centroids. 92
4.8. A sample hyper-polygon. 93
4.9. Sample to class distance defined using fuzzy sets in attribute space. 96
4.10. An overview of transect locations on the Louise Island test site. 100
4.11. Idealised transects and the effects of shifting them within uncertainty bounds. 102
4.12. Simplified examples of cross-correlograms. 105
4.13. The sequence of polygons 'encountered' on each transect. 108
4.14. Differences between measured slope and TRIM modelled slope. 116
5.1. The Lyell Island study area. 123
5.2. Relative frequency using an ML realisation. 127
5.3. Relative frequency using a worst-case realisation. 128
5.4. Factor-of-safety values for slide zones relative to number of cells. 129

5.5. Factor-of-safety values for non-slide cells relative to number of cells. 130
5.6. The previous two figures graphed using cumulative values. 130
5.7. A scatter graph of variance vs. factor-of-safety for slide zones. 131
5.8. Population (random subset) vs. standard deviation. 132
5.9. Worst case realisation factor-of-safety vs. standard deviation. 133
5.10. Worst case realisation using the upper 50% of slide zones. 133
5.11. Relative frequency of the relative position in each slide. 134
5.12. Position of low FS predicted areas relative to slides. 135
5.13. A comparison of predictive accuracy between the original and updated models. 137
5.14. The location of differences between the two models relative to slide zones. 138
6.1. Typical graph of normally distributed uncertainty 148

Acknowledgements

This project was funded in part by Forest Renewal British Columbia (project #HQ96078-RE). The author wishes to thank the following (in chronological order): Dr. Olaf Niemann for providing the original data and contacts; Dr. Peter Keller for assistance with contacts, grant management and considerable help with brainstorming; my supervisor, Dr. Brian Klinkenberg, who started the ball rolling, brainstormed, edited, and performed all the other thankless supervisory duties; Todd Golumbia and his crew at Gwaii Haanas AMB/Parks Canada for logistical support (and some delicious fish)—specifically, Warden Debbie Gardiner for support on Lyell (and for climbing all those slides) and Warden Dennie Chretien for assisting with the aerial video work; Pat Bartier, Gwaii Haanas GIS analyst, for making data available and also helping with logistics; Macmillan Bloedel and their employees and contractors who assisted with data and logistics on Louise Island; Dr. Rosaline Canessa for assisting with field work and deterring bears; John Barker and others at Western Forest Products for making data available and supporting the research extensions; the staff at the Malcolm Knapp Research Forest for logistical and field assistance; and, throughout it all, my wife Wanda.

Chapter One

Introduction

1.1. The Problem

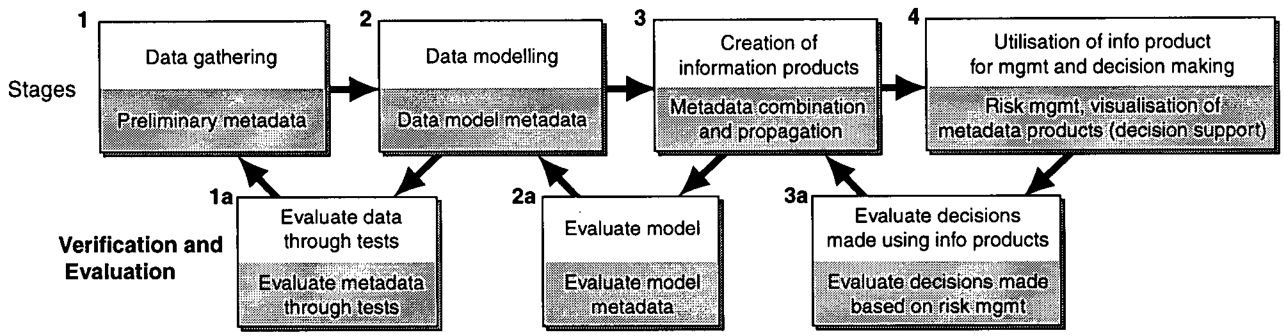
Effective natural resource management requires significant amounts of information. To be useful, this information must be up to date and accurate, must cover the entire management area, and must be in a useable form. However, resource management is rarely practised in an optimal information environment. If data are simply missing, out of date, or in the wrong form, the solutions to these problems are often straightforward; however, data *accuracy* issues are a far less tractable problem. A key ingredient in increasing the effectiveness of natural resource management decision-making is the development of efficient and realistic models of data accuracy. A model of data accuracy differs from a measure of accuracy. The latter allows only evaluation, while the former allows both evaluation and further manipulation.

A number of models have been developed that purport to realistically represent one or more aspects of data accuracy. However, a crucial deficiency exists in this research area. When a standard model is developed, an important phase in the development process is the testing and verification of the model. The question must be asked: does the new representation of the resource or resource-based process (the model) *accurately* represent the resource or process in the real world? Without this type of verification process the utility of the model is in question. The entire field of data accuracy modelling suffers from a notable lack of verification work.

The research presented in this document focuses on the issue of accuracy (or, to be more precise, 'uncertainty') model verification. The central research question is: **can a natural resource management uncertainty model be verified in order to evaluate its utility in real-world management?** For example, a model that claims to represent uncertainty in slope stability might be used to reduce the amount of road construction in an area where uncertainty is high. The question becomes: is this justified by the *actual* uncertainty? Answering the apparently simple research question involves addressing a wide range of issues. For one, an uncertainty model can operate at any one of a number of levels, from the representation of an attribute of a particular resource (such as soil cohesion) through to the representation of aspects of the human decision making process (such as risk management in resource allocation). The uncertainties at each level are quite different in nature. Verification at each level will require unique methods. Therefore, it will be necessary to focus on particular aspects.

A second issue in answering the principal research question is that there is no straightforward *yes* or *no* solution. There exists no single simple statistic to determine if uncertainty as modelled matches the actual level of uncertainty. An uncertainty model can represent 'soft' information such as ' too what degree can I trust this value', or 'what is the level of risk associated with this decision' (as compared to easily verified 'hard' data such as 'percent slope' or 'soil class'.) Therefore, it will be necessary to make use of surrogate measures and exploratory analysis to approach the answer to the principal question. Nevertheless, the research question is a crucial one.

The research presented herein represents a major step in an overall research program intended to integrate uncertainty management into natural resource decision making. This research focuses on modelling and verifying uncertainty at both the data gathering and information product modelling stages. In the process of developing the verification methods and results, some initial work is also performed on the integration of uncertainty models into real world management decision making. Presentation of this latter work serves to highlight the need for integrative research; however, such research can only be effective once all the pieces of the puzzle are in place.



Contributions to Metadata Modelling

1. Measure and record uncertainty in data (Ch. 3)
2. Fuzzy model/mathematical model of data uncertainty (Ch. 3)
3. Propagation through process model (Ch. 3)
4. (Touch on) visualisation and risk management (Ch. 3,6)

1a. Evaluate fuzzy and mathematical model parameters (Ch. 4)
2a. Evaluate process model (slope stability) uncertainty (Ch. 5)
3a. Discuss . . . future research (Ch. 6)

Figure 1.1. A view of data flow and feedback (verification) loops in resource management. The shaded boxes represent metadata flows.

1.2. Major Questions

The organisation of this research, and how it fits in to the flow of information in resource management, is presented in Figure 1.1. This figure represents a simplified typical information flow from initial data gathering through to the decision making stage. Work on uncertainty management superimposes another level onto each box in which the focus shifts to metadata (shaded sections). Metadata is data about data -- information that explains the sources, fitness for use, and other similar factors. The feedback loops, representing data verification exercises, are well established for standard data collection. For example, the output of a forest growth model would be calibrated through field checks. However, the feedback loops and verification of *metadata* do not typically occur.

The first box (data gathering) represents the initial stages at which raw data gathering takes place. The 'data modelling' stage (box two) represents the reduction of these data to a useable form, such as classification or downscaling (although these first two stages are often combined in tasks such as remote sensing). The principal purpose of metadata at the data modelling stage is to quantify how well the data model represents reality. The feedback loop from the data modelling to the gathering stage is typically performed through procedures such as classification accuracy checks or other types of spot sampling. However, the focus of the *metadata* feedback loop is to determine if the predicted variability matches the variability on the ground. Such variability may be a function of the data gathering (e.g., sensor precision or resolution) or of the classification stage (e.g., reduction from cardinal to interval data). These metadata can be based on a wide variety of items, from the inherent variability in the classification scheme through to spatial variability inaccuracies caused by polygonal structures. The inherent complexity of quantified metadata creates an equally complex problem in verification work.

The third box in Figure 1.1 represents information products. These can include anything from simple data overlays (e.g., a soil and ownership layer combination) through to complex simulation models that use many inputs. Again, the feedback loops are well established for standard modelling—one simply compares model output with reality through a sampling scheme. However, metadata propagation procedures for both simple and complex models are a relatively new area of inquiry. This refers to the process of combining the metadata associated with the major inputs to a model in such a manner that the resultant is representative of the model's output uncertainty. There are no established methods for comparing these complex metadata with the variability that exists in reality. While a number of metadata propagation models have been attempted (e.g., Dunn *et al.* 1990), there have been little or no attempts to verify their utility.

The act of decision-making based on information products is represented in the fourth box. The feedback loop here is the evaluation of the decisions—which would typically lead to better information products. The metadata loop here focuses on evaluating the quantification, summary and presentation of 'risk' as it affects decisions.

The research presented in this document includes contributions to each of the stages presented in Figure 1.1, as summarised at the bottom of the figure. The key contributions focus on the first two feedback loops (1a and 2a). While the third loop (decision evaluation) will be discussed and demonstrated, a proper evaluation would require a far larger project scope.

The specific questions asked in this research are:

1. What are appropriate methods for modelling data uncertainty in natural resource management, making use of information typically available?
2. How appropriate are these methods (1), and how can this 'appropriateness' be determined? Specific questions include:
   - 2a. How effective is gathering metadata from expert opinion?
   - 2b. How effective is gathering metadata from published variability statistics?
3. What are appropriate methods for propagating these metadata through to information products (i.e., using a typical type of natural resource model)?
4. How appropriate are these methods (3), and how can this 'appropriateness' be determined?
5. What are some of the implications of the methods outlined above for resource management decision making?

The principal focus of this work is answering questions #2 and 4—those concentrating on verification of metadata. The principal focus of the discussions evolving from these answers is the implications for management—question #5. Questions #1 and #3 will be addressed; however, the discussion will partially draw upon research conducted previously by the author.

This research makes use of specific models and a specific resource sector. The models (slope stability) and the resource sector (forestry management) have been chosen for their broad applicability. The theory and procedures developed herein can be applied to a wide range of models and management regimes. Links to the broader field of 'resource management' are noted throughout the document.

1.3. Research Organisation

This dissertation is organised as follows: after this introduction, the second chapter contains the background and specific research justification for all that follows. It is demonstrated that the traditional methods of modelling natural resources are inadequate. Uncertainty, uncertainty modelling, uncertainty propagation, and links with management and decision making in the resource sector are each examined in turn. The various research fields are described with an eye to highlighting the factors that link them. The third chapter presents the methodology and results of earlier uncertainty modelling work undertaken by the author—work that developed the basic forms of the model used in the remainder of this document. In this chapter, one specific method of modelling and propagating natural resource uncertainty is used to highlight the potential of the field.

The following chapter presents new theoretical and applied work on verifying uncertainty models and sampling for uncertainty values. The model presented in the preceding chapter is ground-truthed, and techniques are developed to tune the model so that its results match these ground data

Chapter Five moves to the next stage of verification—with a focus on the metadata *generated* by the modelling procedure. The output of the model developed in Chapter Three is verified using field data. Methods are developed to address the comparison of variability-focused metadata with yes-or-no confirmation data.

The following chapter takes a further step up in metadata complexity by examining some possible methods of integrating uncertainty management into real-world decision making. This discussion chapter presents examples developed during the verification work, which also serve to demonstrate some possible applications of this research in decision support. Implications for management in both forestry and resource management in general are discussed. Recommendations for future research as also presented here. The final chapter summarises the overall conclusions.

1.4. Contribution to Knowledge

This research contributes to knowledge in the field of uncertainty analysis on both theoretical and applied fronts. Two major theoretical areas are explored: 1) the integration of confirmatory sampling into a spatially variable uncertainty model, with a focus on the representation of fuzzy classes in attribute space and the development of several measures of comparing samples and classes; and 2) the development of methods for verification of uncertainty model output, with a focus on comparing complex multivariate variability data with ground samples. As a secondary contribution, procedures are also developed for integrating oblique data with planimetric databases in GIS, with a focus on capturing and using registration and digitising uncertainty as spatially-variable metadata. Chapter Two will detail how these theoretical contributions fit into the overall discipline of uncertainty research.

Applied contributions include: 1) verification of 'expert knowledge' on soil spatial structure in a slope stability model; 2) development and evaluation of a set of tools for measuring spatial parameters from oblique image data; 3) testing of a slope stability uncertainty model using a temporal landslide database developed with these tools; and 4) evaluation of the modelling techniques and tools developed in 1-3 through a case study.

Uncertainty is a crucial issue in resource inventory data, as well as almost all other types of spatial data. Yet it is unlikely that a general purpose 'error button' will ever be developed—the problems faced are too diverse and also too application-specific. There are so many ways of approaching the problems of uncertainty models, visualisation, sampling, etc., that developments in one of these areas is rarely applicable to others. Therefore, by focusing on a crucial set of problems (verification) and integrating them into a 'cradle-to-grave' uncertainty management task, it is hoped that this research will both demonstrate and increase the utility of uncertainty modelling for natural resource management in general.

Chapter Two

Background

2.1. Introduction

'Gulliver's Travels' contains the story of a cartographer from a small kingdom. In his quest for greater and greater mapping accuracy, he created maps at larger and larger scales. Eventually, he found himself working at a 1:1 scale; unfortunately, there was no kingdom left to describe—it was filled with his creation. Maps and spatial databases are an abstraction of reality. Details are filtered out in order to clarify information. The 1:1 scale map rather defeats this purpose (as well as being somewhat difficult to fold). Sampling, filtering—in fact any abstraction—leaves a gap between representation and reality. This gap, in essence, is data uncertainty.

In the days when cartographers risked their lives in leaky boats trying to chart the unknown reaches of the world, uncertainty was, to say the least, very high. Yet no ship captain expected to use these charts in any precise way. It was sufficient to know that there was a big piece of land somewhere in a general westerly direction. Blank areas or straight lines meant unknowns. If a map of a particular piece of coastline was drawn in some detail, you would expect that it roughly corresponded to reality, yet you would be foolish to bet your life on the location of a particular shoal. Uncertainty was built into the structure and conventions of map making. In any case—any map was better than no map at all.

As cartography matured and information on maps became more precise, the issue of uncertainty still remained largely part of the map-making process. If the source data did not support it, then a competent cartographer would simply not draw a 1:5,000 map. Within a map, the thickness and style of lines could be used to indicate local spatial uncertainty, or uncertainty in a particular class of objects. Other visual techniques could be utilised to draw attention to degraded or missing information, or source data that had a limited life-span. While control of map making remained within the hands of cartographers, uncertainties were largely understood.

Then came the computer 'revolution'; the control of spatial data began to slip out of the cartographers' hands. In a short span of years the production of maps—formerly solely the realm of specialists—became possible using simply a home computer; a six-year old (or a newspaper columnist) became able to turn out professional looking maps with the touch of a button. Yet such software is only capable of imitating what was the simplest part of the cartographer's job: the mechanical drawing of the map. Communication skills are not so easily emulated. Computers also took the analysis of spatial data out of the cartographer's hands. Automated area and perimeter calculations soon gave way to data overlays, topographic analysis and spatial statistics—all of which are available to anyone who knows which button to press. In the areas of both cartography and spatial analysis the control of data uncertainty was wrested from the arms of cartographers, and soon began to be a problem; or, to be more precise, a non-problem—it was virtually ignored.

2.1.1. GIS and Uncertainty

Some cartographic and most analytical tasks formerly performed by cartographers have been passed on to geographic information systems (GIS). These often massive programs have revolutionised the manipulation of spatial data. However, some of the basic assumptions and structures built into GIS foster this 'non-problem' of data uncertainty. A GIS enables spatial data to be viewed and manipulated at virtually any scale. It is, in fact, a scaleless working environment. If addressed at all, scale information will normally only accompany data for display purposes (e.g., what size to print a label). A typical GIS stores data at a resolution that is capable of locating a point down to a tolerance of less than the width of a hydrogen atom (using double precision with a local co-ordinate system). It also typically reports all information: co-ordinates, areas, etc., with the same excessive precision. These two characteristics—lack of scale and the reporting of extreme precision—virtually eliminated the implicit recognition of uncertainty found in manual cartography.

After a number of years and a considerable amount of frustration on the part of users, GIS and spatial analysis researchers began to examine how this implicit uncertainty could be made explicit. It took many more years for basic conceptual work to appear, in which the nature of spatial data uncertainty was examined, terms were defined, and eventually standards put in place. This field of inquiry is still in its infancy, due in part to the complexity of the problem, and also in part to the reluctance of users to accept the fact that an answer of lesser precision can be more 'correct'.

2.1.2. Chapter Layout

This chapter examines the current state of research into error and uncertainty in spatial data as it relates to the various elements of resource inventory, with a specific focus on forestry aspects. It does not attempt to trace the evolution of thought in this field, as this 'evolution' does not represent a co-ordinated effort towards a clear goal. Instead, what appears in the literature is a haphazard series of incremental steps on many diverse fronts towards numerous disparate goals. Therefore, the chapter is organised in a manner reminiscent of the overall document. Terms and basic concepts are defined and presented first, followed by a discussion of some of the major areas of application. Relevant research into the modelling of uncertainty is then presented, with later sections focusing on output issues relevant to this added dimension of data. The final section details research into uncertainty in natural resource management—in particular forestry data—and its implications for this current research.

2.2. Error and Uncertainty

Geographic information systems can be found on the desks of public utility planners, natural resource scientists, and almost anyone else concerned with spatially referenced data. Utility management focuses primarily on straight lines and unambiguous locations. Detailed GIS co-ordinates and precise analytical routines cater to a utility manager's desire to see the world in black-and-white. In contrast, when the same data structures and models are utilised in the resource sector, they potentially bear little resemblance to the spatial characteristics of the information being captured or modelled. Here, entities in question might be better represented by shades of grey.

This dichotomy between imprecise reality and its precise digital representation has given rise to a rapidly growing research area in which spatial data experts grapple with the implications of uncertainty and error analysis, while cartographers focus on the special problems of communicating uncertainty. As this field has developed, researchers have fanned out across a broad front: advancing error analysis (Chrisman 1989; Chrisman 1991), locational and feature uncertainty analysis (Burrough *et al.* 1992), error propagation methods (Heuvelink and Burrough 1993), the visualisation of uncertainty (MacEachren 1992; Goodchild *et al.* 1994), and numerous related topics.

The self-referential problem of uncertainty about uncertainty terminology has been a notable stumbling block in this avenue of inquiry. The primary terms, namely 'error' and 'accuracy', are commonly used interchangeably—compounding the problem. Most writers would agree that 'error' refers to deviations from a 'true' value. Almost all resource data contain some degree of error; however, as the 'true' value is generally unknown, the error cannot be easily quantified and stated in the same manner as errors in a numerical model might be. The term 'accuracy', as defined by Buttenfield and Beard (1994), is a more easily quantifiable alternative. In their definition it refers to measures of discrepancy from a modelled or assumed value.

2.2.1. Uncertainty Defined

Here, the term 'uncertainty' is utilised to include both of the above, as well as to extend these concepts. In its broadest sense, uncertainty refers to knowledge of possible deviation from a 'true' value, but without precise knowledge of the magnitude. It is not as broad a term as data 'quality' which, in its commonly accepted definition (Guptill and Morrison 1995) includes items that cannot be subjected to test or verification, such as lineage or completeness. Uncertainty may exist for many reasons: inability to measure precisely, alterations in values during processing (e.g., manipulation or classification) or, at a more fundamental level, natural variability in the phenomena being measured. Uncertainty is not necessarily an absolute, since the resolution of the dataset or analysis may be a factor.

2.2.2. Quality

Uncertainty is the focus of this work, yet the other elements of data *quality* play important supporting roles. Before delving into the subdivisions of uncertainty, it is important to step back and put it in context with other metadata elements. As introduced earlier, a U.S. committee on data standards (NCDCDS 1988) produced an influential document that attempts to categorise the major elements of data quality. It includes the following:

1. **lineage**: the history of the data and the operations performed on it;
2. **completeness**: the extent of data coverage (spatial or attribute) relative to the complete real-world object (e.g., the subset of soil attributes in the database relative to *all* possible attributes) ;
3. **positional accuracy**: the closeness of spatial co-ordinates to the 'true' values (or values accepted as true);
4. **attribute accuracy**: as above, but with reference to the attributes of the spatial location;
5. **logical consistency**: for example, the appropriateness of a chosen classification scheme; and
6. **temporal information**: includes references to periodicity, the temporal range (shelf life) of the data, and other relevant descriptive temporal elements.

These categories function well in a descriptive sense, but are not oriented towards implementation of quality tracking or analysis. In fact, the descriptive nature of these categories has prompted many agencies to implement metadata through descriptive add-ons to their spatial and attribute GIS layers or data products. The US Geological Survey (USGS) and many Canadian government agencies have taken this approach. Given the metadata files and some interpretative information, a knowledgeable user can make general decisions about the utility of a data product for their purposes—sometimes. Lack of standards makes it difficult to compare products from different agencies, or at times even within a single agency. Positional accuracy can usually be described with a small set of numbers; however, an item such as 'logical consistency' can be interpreted in many ways.

Three items from this list *do* lend themselves to a more quantitative implementation: positional accuracy, attribute accuracy, and temporal information. Only when metadata such as these are stored quantitatively does it become possible to mathematically manipulate these values, to follow them through overlays or models, and to present them visually. A qualitative understanding of the implications of these metadata is still an important ingredient, for only then can a user determine data's fitness for use; however, quantitative metadata expands the utility of this information considerably. A focus on numerical aspects of quality brings the discussion back to the realm of uncertainty.

2.2.3. Subdivisions of Uncertainty

By limiting this discussion to the measurement of natural resource data, uncertainty can be subdivided into these three broad areas: positional, attribute, and temporal. Positional uncertainty is a well defined topic area in geomatics. Those making positional measurements, notably surveyors, are accustomed to imagining a bell curve of uncertainty that exists along both the horizontal axes (Figure 2.1) and on the vertical axis as well. Much of surveying science is designed to minimise these uncertainty envelopes. However, when the study of uncertainty is expanded to more complex objects than 'points', numerous other issues appear. For example, how does uncertainty vary along the line drawn between two known points? How does uncertainty behave in an

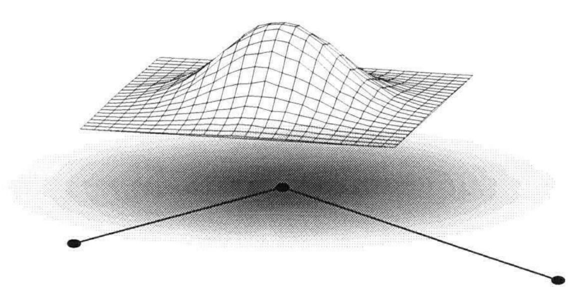


Figure 2.1. Location probability of a survey coordinate in 2-D space.

overlay of two data layers? The following sections present the primary sources of uncertainty in each of these areas, and provide an overview of research into these topics.

2.2.3.1. POSITIONAL UNCERTAINTY

Uncertainty in position is certainly the most tractable of the issues discussed here. It is the basic problem that cartographers dealt with through line size and scale, and one that continues to occupy much of the efforts of geomaticians and surveyors. The three co-ordinates used to define a point in space may be mathematically compared with 'true' values, providing simple measures of positional uncertainty. Of course, the elusiveness of 'true' values compounds this problem, as does measurement accuracy, data entry error, etc.

2.2.3.1.1. Registration

Positional uncertainty is an issue at several points in the processing of spatial data. The first of these is the registration of the data source to a reference value. This might involve registering an air photo using survey markers and ground control points, or the registration of a satellite image to a reference dataset. Survey markers are undoubtedly the best spatial reference point available. First and second order control points are established with extremely high accuracy, using mathematics that correct for the earth's curvature, as well as triangulation within the survey grid. The co-ordinates of a control point are subject to errors in the reference datum and ellipsoid; however, the magnitude of these errors is very small in a local context.

High order survey control points are rarely used in the registration of satellite images or air photos. Lower order control points, GPS derived control points, or existing planimetric dataset points are more commonly used. Each of these sources is subject to various inaccuracies. Lower order control points are not subject to strict controls over placement, and may have positional errors significantly higher than their 'parents'. GPS points are subject to the many types of error associated with GPS data (see Owens and McConville 1996), and existing datasets have already been subject to registration, and so act to compound uncertainty.

The process of image registration might involve simply shifting the image until the control points line up with minimal error, skewing the image (independent x and y stretching), or 'rubbersheeting', which allows every co-ordinate in the image to shift. Each of these methods generate non-uniform registration error for every point in the image. However, these values are normally summarised in a single value such as root mean squared (RMS) error (and then typically ignored). Recent research has begun to address this loss of information during registration. New methods of performing and summarising registration accuracy have been developed (Mather 1995; Buiten and van Putten 1997), and methods of employing multiple representations are proving useful (Djamdji 1993; Fonseca and Manjunath 1996). A key element is not to simply perform the best possible registration, but to also maintain the uncertainty information for later processing (e.g., Delavar 1997).

2.2.3.1.2. Other Sources

Positional uncertainty also occurs at later stages of spatial data processing; however, typically the entities being manipulated are of a higher order than points. Both line and area entities (vectors and polygons) are built out of point data, yet involve uncertainties of a different nature (see below). Raster datasets are somewhat simpler; however, uncertainties generated during the manipulation of raster datasets can be complex in nature.

2.2.3.1.3. Lines and Areas

During the period of transition from paper to digital datasets (still continuing in some sectors), manual digitising was the primary method of vector data input. Studies of digitising uncertainty constitute a major part of the field of uncertainty analysis. One of the first important works in this field (Perkal 1966) introduced the concept of an 'epsilon band', which later led to the epsilon distance model of cartographic lines (Peucker 1975; Chrisman 1982). In this model the assumption is that a cartographic line (i.e., the proper location of a feature) is surrounded on each side by an area of constant width epsilon, and that the digitised representation of the line will lie somewhere within that area (Figure 2.2). The distribution of the location of the lines may then be described by some type of distribution function, such as a probability density

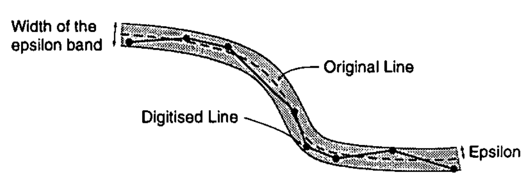


Figure 2.2. Epsilon boundary model of a digitised line, where the true location of the line is assumed to lie within the epsilon band.

function (PDF). A number of functions are possible, depending on the assumptions used (Figure 2.3). The epsilon *band* can use various measures, such as maximum deviation (range), interquartile range, or some other. There is no single accepted definition of this model. It is explored in some depth by several authors (e.g., Dunn *et al.* 1990; Chen and Finn 1994). Others have gone on to examine the structure of uncertainty on the line segment *between* the digitised points, allowing the epsilon band some flexibility (Chrisman 1982; Dutton 1992).

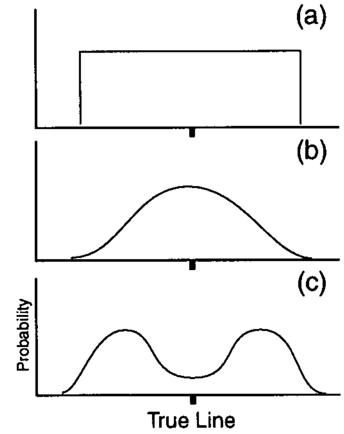


Figure 2.3. Examples of probability density functions for line or digitising error: (a) rectangular, (b) bell-shaped, (c) bimodal (Adapted from Dunn *et al.* 1990)

Most recently, stochastic methods of addressing uncertainty in vector objects have begun to appear. For example, Hunter and others (Hunter 1995; Hunter *et al.* 1996) have developed a method of imposing controlled stochastic changes in the spatial location of all vector objects, allowing stochastic simulation in a vector environment. Others (Youcai and Wenbao 1997) utilise similar methods to address digitising error specifically. Other methods, such as moving bands, spatially autoregressive and Markov processes have been described (see Haining *et al.* 1983).

Most of the above research applies specifically to database objects representing linear features in the real world. However, database vectors are also used to represent more abstract structures such as soil polygons. Here, uncertainties in spatial locations still apply, but are generally rendered insignificant due to the magnitude of uncertainty generated through the abstraction of reality: namely attribute uncertainty.

2.2.3.2. ATTRIBUTE UNCERTAINTY

Positional and attribute information are stored separately in most spatial data models. The uncertainties in each are often determined by quite different processes, and in this case are termed *separable* (Goodchild 1991). For example, the boundary of a clearcut might be digitised from an orthophoto, with the accompanying registration and digitising error; yet the attributes are determined through interpretation and field sampling. Uncertainties in these values have no bearing on spatial uncertainty.

However, initial forest inventories do not necessarily contain obvious spatial discontinuities between stands. In a typical procedure the stand polygons might be outlined on imagery, followed by an iterative process of ground surveys leading to changes in stand boundaries. Here, positional and attribute uncertainty are *not* separable, as the process of determining attributes is linked to the process of determining boundaries. A forest cover map would typically contain a mix of separable and inseparable uncertainties. Most research treats both types as completely separable, although a few studies (notably Mark and Csillag 1989; Brimicombe 1993) attempt a synthesis.

It should be noted that many of these problems of separability are partially a result of, or compounded by, the data model utilised. Spatial data may be modelled in two basic ways: as discrete objects (e.g., vector model or object-oriented model) or as continuous fields (e.g., raster model). When fields are used, the attributes modelled are usually not sharply bounded, and so separability is less of an issue (Goodchild 1989).

In contrast to positional uncertainty, attribute uncertainty has received considerably less attention in research and analysis. This is hardly surprising, as the dimensions of the problem are vast, and many of the uncertainties resist easy quantification. Attribute uncertainty can arise in several general ways. Error or imprecision in field measurements is perhaps the simplest to deal with. More complex are uncertainties generated due to the way the data object represents complex reality. For example, a point object might be used to represent a city at a particular scale. Polygons are often used to represent transitions between soil types. The transition, which extends over some distance, is represented by a sharp discontinuity. The attributes assigned to such a polygon will have varying degrees of validity throughout its spatial extent (Mark and Csillag 1989). The simplification of reality necessary for data storage and modelling imposes these types of attribute uncertainty.

Simplification of attributes themselves also generates uncertainties. The process of classification splits a continuous reality into discontinuous parcels for ease of sampling, storage and analysis.

However, the taxonomy involved may produce an incomplete or even misleading description of the attribute. Problems include:

Internal purity: the degree to which a random sample matches the class descriptor can be abysmally low for some data types—notably soils.

Class boundaries: some classes may be functionally similar, visibly similar, or both. Other classes may be extremely distinct. Rigid class boundaries do not allow this distinction. A sample is either class A or B, even if its' properties are similar to both.

Sampling error: field or laboratory error introduces a random element into classification or study comparison. Although multivariate statistical techniques can often ameliorate such problems, resulting indices or principal component scores are not necessarily easy to interpret (Burrough 1989).

Once attributes are combined in a modelling scenario these uncertainties can play havoc with the results. Within a particular speciality the user often has some implicit understanding of the attribute uncertainties. However, when complex environmental models draw from numerous disciplines for their source data, a lack of attention to uncertainty at the source leaves the model's user with no choice but to trust the datasets implicitly. It then becomes difficult to estimate even the general variability in the results.

Some researchers have proposed alternative data structures that better describe attributes without abandoning the concept of a categorical coverage. Spline values might be used to describe environmental gradients within a polygon (Herring 1991). Fuzzy classification methods that recognise the variability within and between classes have been proposed (Hall and Wang 1992; Burrough *et al.* 1992) and implemented in numerous disciplines, such as forestry (Palubinskas 1994) and earth science (Du and Lee 1996).

Others have noted the inadequacy of vector-based categorical data structures to model many types of natural resources. Vector structures were the only alternative when computing power was relatively minimal relative to the amount of data in a large inventory. Today, as many data sources are raster based, the advantages of continuous data structures are becoming obvious. Researchers such as Mark and Csillag (1989) point to the advantages of a native raster structure to represent both spatial and attribute uncertainty. Implementations include soil databases (Rogowski 1996), model propagation (Mowrer 1995) and fire modelling (Delavar 1997) among others.

2.2.3.3. TEMPORAL UNCERTAINTY

Other than subsurface geologic maps, most natural resource databases are dynamic to some degree. The nature and extent of temporal uncertainty will vary with the resource being mapped. The principal sources of temporal uncertainty include: 1) gradual change, such as tree growth, succession, or urban expansion; 2) cyclical change, such as variations in deciduous canopy between summer and winter; and 3) uncertainty due to measurement, where measurements may be spread over time, or analysis is delayed relative to measurement. The first of these—gradual change—is of primary interest in forestry and most other types of GIS analysis. Forestry is a particularly good example of temporal change in a heterogeneous environment.

Gradual change in forest inventory is normally accounted for using a periodic inventory cycle. In British Columbia (BC) the cycle is approximately two years; cycle times in other areas vary (e.g., the Province of Quebec uses ten years). Natural forests are spatially heterogeneous, making stand delineation an uncertain and often unrepeatable exercise. Estimates of map accuracy from photointerpretation of forested areas indicate that disagreement is as high as fourty to fifty percent (Edwards and Lowell 1996). The natural heterogeneous forest also changes in different ways, and at different rates, making it a challenge to model.

Temporal uncertainty models for forestry must focus on several issues. Of particular importance is the variability in stand boundaries over time, due to both change in the forest and photointerpretation uncertainty. Also important is the uncertainty in model results. For example, a growth model is based on data collected at a particular date from point samples within a forest stand polygon. The output of this model develops greater uncertainty as time passes; the uncertainty being based solely on model precision. While this uncertainty is commonly reported, it is rarely integrated with uncertainty in data collection, uncertainty in volume estimates, uncertainty in other stand attributes, and boundary variability. The resulting change in uncertainty over time is a function of spatial and attribute variability over time, as well as built-in model uncertainty.

Although pure computer research deals with temporal uncertainty databases separately, as in data structures (Kanazawa 1994), or computer vision (Rohrer and Sparks 1993), natural resource research focuses on integration of spatial and temporal elements. Examples include forestry work (Lowell *et al.* 1996), soils (Or and Hanks 1992), and fisheries (Hougard and Valdez 1994).

2.2.4. Subdivision by Source

Spatial, attribute and temporal uncertainties are clearly linked in complex ways. Although this three-way subdivision may be useful for basic research, from an operational perspective it is more beneficial to subdivide them by the *sources* of uncertainties. Where uncertainties are currently addressed in spatial database management it is usually in this manner. From a source perspective there are three main areas: inherent uncertainty, uncertainty in data collection and input, and uncertainty in data manipulation.

2.2.4.1. INHERENT UNCERTAINTY

Natural vagueness, also referred to as conceptual error (Veregin 1989), inherent uncertainty (Lanter and Veregin 1992) and inherent property error (Maffini *et al.* 1989), occurs in data that possess no standard for comparison of measurement. Beyond a certain point, increases in sampling density or in the precision of instruments used will not result in any increase in information content. Natural vagueness may be due to natural variations in the source or due to an inability of the chosen data model to fully encompass all properties of the source.

Soil polygons are a commonly cited example of this problem (e.g., Burrough 1986b; Kollias and Voliotis 1991; Burrough 1993) due to the inability of the polygonal structure to address gradual changes over space. Other examples of natural vagueness include mobile species, seasonal fluctuations, and quantities that simply cannot be measured with available techniques. For example, radar images of an ice pack may show well-defined lines, yet the constant movement of the ice introduces uncertainty.

2.2.4.2. DATA COLLECTION AND INPUT UNCERTAINTY

Two field scientists, independently studying the same resource, will rarely generate the same data. This is a well-recognised phenomenon, and many data gathering techniques are designed specifically to minimise this observer bias. However, it remains an issue in uncertainty management because there is rarely sufficient time or money to complete the number of samples needed to minimise such bias. When 'judgement calls' are made, they introduce a subjective element that is very hard to quantify without repeating the entire process. The problem of observer bias will vary in severity between different classes of resource survey and different resources. Rapid reconnaissance surveys, such as shore-zone typing (e.g., Howes *et al.* 1994), will be particularly susceptible to this effect.

2.2.4.3. DATA INTERPRETATION

The precision of sampling instruments is rarely a problem in natural resource surveys. Data interpretation uncertainty arises when the data generated by those instruments are manipulated into a form suitable for storage and analysis. Field samples often require the application of a variety of inference techniques to estimate the distribution between sample sites. For example, soil or forestry point samples on the ground are combined with spatial parameters inferred from remote sensing to derive a polygonal distribution. Although the point samples may be precise to the n^{th} decimal place, their purpose is to describe a local average of conditions. A good sampling scheme will pick up the main components of this variation; however, *between* the samples the data are always inferred. Once the procedures are complete, an inferred data point is indistinguishable from a sampled one.

The data gathering and interpretation stages are also subject to mistakes in the sampling process. Although commonly termed 'errors', the narrow definitions employed in this research field require a separate term to describe accidents or mistakes, rather than deviations from a known value. The term most commonly used is 'blunders' (although this departs from the dictionary definition which focuses on 'gross mistakes'). Such blunders can occur in the field (mislabelling, misreading, wrong position, etc.) or in lab analysis. The statistical likelihood of such blunders can be estimated, providing further information about uncertainty in source data.

When remotely sensed data are the primary data source there are a number of other inference uncertainties to account for. Air photos commonly have several geometric problems that must be corrected prior to use, including tilt displacement, radial displacement and topographic displacement. They are also subject to distortions such as atmospheric refraction, lens irregularities, film or print shrinkage and image motion. The process of correction is often one of inference based on other information (e.g., correcting topographic displacement with elevation data) and can introduce its own uncertainties into the process.

Satellite images can only measure the reflectance of an object to radiation at various wavelengths, and therefore introduce uncertainty in the interpretation of these data. There are also a number of factors that intervene between the source and the sensor, such as atmospheric haze, path radiance, or variations in solar angle. Satellite and other digital remote sensing devices are also subject to many of the geometric distortions mentioned above. Once again, while some of these distortions can be corrected, each correction introduces uncertainty into the data collection process.

2.2.4.4. DATA ENTRY

When data are entered into a system using some manual means, there are a number of opportunities for uncertainty to appear. Digitising, one of the more studied sources of uncertainty (e.g., Chen and Finn 1994 or Youcai and Wenbao 1997), can include registration skewing, variability in line location and outright blunders such as mislabelling. The source map being digitised is also a source of uncertainty due to printing registration problems, stretching of the medium or thickness of lines. For example, the area covered by lines (i.e., underneath) on a map represents an area of uncertainty. In one study, Burrough (1986b) notes that as much as ten percent of the total map area of a 1:25,000 soil map consists of lines.

2.2.4.5. DATA MANIPULATION UNCERTAINTY

The process of manipulating data can include measures as simple as classifying a cardinal measure, or as complex as the simulation of an ecosystem. While the process of digital data manipulation itself produces rounding errors, such problems are dwarfed by uncertainties introduced through simplification and combination of different types of data.

The process of classification is rarely simple—it involves a number of subjective elements. The choice of the appropriate classification procedure is not always straightforward; different procedures often produce different numbers of classes and different class boundaries. Even if the class divisions are mathematically obvious, there is still the question of appropriate representation of the source. The purpose of the classification is also a factor. Class divisions are often chosen based on their appropriateness for a particular purpose, such as soil classes for slope stability analysis or forest classes for maximising profitability. However, such special-purpose classified data are often made available for other purposes. The secondary user is then faced with suboptimal class definitions and therefore a heightened level of uncertainty.

A related problem (although also related to issues discussed under 'data gathering') is that of cell value averaging. An individual pixel of a remotely sensed image represents the reflectance of all surface features within its bounds, as well as some from adjacent pixels due to factors such as atmospheric haze and viewing angle. If the resolution of the sensor is appropriate for the phenomena being measured, then this is not a crucial issue. All too commonly, however, the pixel is larger than the target. This basic variability in reflectance leads to many of the classification problems discussed above. Although it can be reduced by processes such as spectral unmixing (e.g., Mathieu *et al.* 1994), it still remains an issue for uncertainty management.

2.2.4.6. PROPAGATION

Much of the functionality of a GIS lies in its ability to combine two or more maps for the purpose of analysis. The complexity of this combination ranges from Boolean functions between raster layers, through topological polygon overlay, and right up to the integration of environmental simulation models within or closely linked to a GIS. Even one of the most basic GIS functions, topological overlay, generates data uncertainty that is difficult to both understand and quantify.

Sliver polygons that result from such overlays may be spurious or may actually represent information. Determining which is often a difficult task.

Attempts to estimate cumulative errors (e.g., Newcomer and Szajgin 1984) have led to the general conclusion that, at best, the accuracy resulting from digital overlays is less than the accuracy of the least accurate input layer. This upper bound occurs when all uncertainties spatially coincide. At worst, when they are not coincident, accuracy can be much lower. At the upper end of the complexity scale, spatial analysis and simulation models often perform an elaborate series of operations in order to make their projections. Trying to derive statistical estimators of the uncertainty propagated through such models is generally an intractable problem (Mowrer 1995; 1997).

2.2.4.7. GENERALISATION ISSUES

Uncertainty is also a function of the difference between the scale of the source data and the scale of use. A substantial difference between the two can, and often does, lead to problems at the analytical stage. The definition of 'substantial' depends upon the nature of the data, the type of analysis, and the traditions of the discipline. Some data are relatively scale invariant. The 49th parallel that divides much of the U.S. and Canada is a line that will appear the same no matter what the scale of map. In contrast, a coastline or a road system would be displayed with a greater or lesser degree of complexity depending upon the operational scale.

This type of scale variability is not only a feature of data display, it is also important in analysis. A soil or forestry map at a 1:5,000 scale would present different attributes than a map at 1:250,000. Although the resource itself remains the same, the type of analysis performed on a handful of forest stands would be substantially different in nature from one performed on an entire forest district.

Data gathered at one scale can often be utilised at another if generalisation procedures are performed. Normally, one would only move from large to small scale; however, specialised procedures may allow some movement in the other direction. Generalisation involves a complex set of data manipulation procedures that can act on both spatial and attribute data, and as such generate data uncertainty. However, as the scale of analysis changes so does the tolerance for such uncertainty. Generalisation may, therefore, result in a decrease in data uncertainty in some cases.

2.2.5. Measures of Uncertainty

Measures of uncertainty are dependent upon the type of data under consideration. Spatial data uncertainty will typically be measured using standard circles (assuming x and y are dependent) or error ellipses (x and y independent). The z dimension is normally reported separately, as the data source is usually independent of the others. A more complex analysis might include the distribution of error—typically assumed to be a normal curve. Unfortunately, many of the simpler measures do not lend themselves to further analysis. For example, the USGS type of spatial standard (also employed by the BC government for their terrain data) uses statements such as "Ninety percent of all well-defined planimetric features shall be co-ordinated to within 10 metres of their true position" (SRMB 1990 p. 4-5). In the absence of additional information one could assume that the other ten percent might be located *anywhere.*

Analysis of continuous thematic data might involve statistical measures such as dispersion, measures similar to those employed for spatial data, or implicit measures such as monthly precipitation graphs indicating climatological variability (Buttenfield 1991). Categorical data often utilise an index of classification accuracy computed from a classification error matrix. The matrix consists of a cross-tabulation of estimated and actual thematic values for a sample of points. In such a matrix, element c_{ij} represents the number of points belonging to class i that actually belong to class j. Accuracy indices include: a) the kappa (or khat) statistic, which accounts for correct classifications that occur by chance alone (Hudson and Ramm 1987, for example see Stehman 1996; Naesset 1996); b) user's and producer's accuracy, which account for the accuracy of individual thematic classes (Aronoff 1982); and c) the PCC statistic (proportion of points correctly classified) which may be viewed as the probability that a point selected at random from a layer is correctly classified (Lanter and Veregin 1992). Alternatives to the matrix approach include area comparisons between polygons and ground survey results, or computation of positional error in polygon boundaries arising from classification error (e.g., Hord and Brooner 1976).

The principal problem with all of these measures is that they simply report a generalised account of uncertainty at a certain stage of the process. The numbers may be useful for ascertaining whether data or results are useful for a particular task, but do not offer much help in examining the spatial variability of uncertainty, what occurs when data are combined, or what the implications of the uncertainty are for a particular task or decision. Such tasks are only possible when uncertainty becomes part of the data modelling process.

2.3. Uncertainty Modelling

The term 'uncertainty modelling' is used rather loosely in the research literature. It is important to distinguish between the modelling *of* uncertainty, and uncertainty integrated *into* the modelling process. The modelling of uncertainty is the same as other types of scientific modelling: it is an approximation of how some aspect of the world works. In general, the modelling of uncertainty refers to the concepts, methods, algorithms and data structures that allow uncertainty to be represented in a useable format, compressing the complexity of the real world. The measures of uncertainty introduced above are models of uncertainty; however, they constitute a very high degree of data compression—often a single number represents an entire layer of data. More complex methods of modelling uncertainty will recognise spatial, attribute and/or temporal variability. The goal of uncertainty modelling is the appropriate representation of uncertainty within a data structure. It is difficult to specify an 'appropriate' model if the context is not specified. This problem represents one of the major drawbacks of 'pure' research into models of uncertainty.

In contrast, the integration of uncertainty into environmental modelling maintains a focus on the environmental model itself. Uncertainty models are an essential part of the process, however, the choice of model(s) is based on a number of other factors. Uncertainty estimates and models are utilised with an eye to their integration with other types of data, their ability to function within the modelling software, and the possibilities for propagating the information through to the environmental model's results. Other important constraints include the overall purpose of the modelling exercise (which determines the degree of uncertainty tolerance) and the overall complexity of the process relative to the computing facilities available.

This section will focus on both the structures used in uncertainty models and the process of environmental modelling as it relates to uncertainty management.

2.3.1. Modelling

Most branches of science concentrate on the specific physics and chemistry required to develop functional models. Earth sciences have been no exception; hydrologic, erosion and other models—simple or complex—have focused primarily on the processes, not the distribution of properties. Only recently has it become expedient to extend such 'lumped' process models to distributed models that attempt to include spatial distribution or transport across the landscape. Huge increases in data availability through satellite imagery coupled with exponential increases in computing power have enabled this recent shift.

However, these specialists do not necessarily understand all the ramifications of spatially distributed data. The spatial distribution of properties can be complex. Scale changes can affect these properties substantially. Changes over time also affect the properties of distributed models. Although surveyors and spatial analysts and some hydrologists have studied a number of these topics, traditional divisions between the branches of science have slowed the cross-fertilisation necessary to properly develop distributed models.

2.3.1.1. MODELLING WITH GIS

Distributed models will typically be linked in some way with a GIS. The model may be run externally, using the GIS as a data source and method of display, or internally, utilising standard analytical functions. The former method is advantageous when the model is complex, has been previously developed in a particular programming language, or requires specialised hardware. The latter—running the model inside the GIS—has the advantage of minimising data translation problems; however, most GIS programs only provide a range of general-purpose functions. It may be difficult to express a complex model in terms of such functions, and, even if such expression is possible, the generalisation may also lead to suboptimal computation time.

One of the crucial issues in this migration of models to a distributed GIS-linked form is the lack of techniques for determining model reliability. Often, the only criterion of quality in GIS-linked models is the cartographic display of the results (Burrough 1993). Even standard non-distributed models suffer from this problem. Users—particularly non-specialists—often accept the simulated results without adequate validation. In fact, there are few standard methods for validating models, and for some validation is difficult or impossible. Even if a simulation model is validated it may, through reliance on empirical relationships, not describe the underlying process correctly.

This dearth of techniques for analysing, propagating and reporting error in distributed models has been addressed by a recent surge of research. Some methods utilise derivations of standard techniques, while others borrow from related fields. The following section details some of the more common techniques available to address the modelling of uncertainty for the purpose of determining the reliability of natural resource models.

2.3.1.2. METHODS OF MODELLING

In sections above a number of methods of modelling spatial uncertainty in basic GIS entities have been discussed. Uncertainty in points, lines and polygons—the traditional reductionist entities—can be addressed with a number of statistical measures. Often a single measure refers to all entities within a data layer; however, the standard circular error, epsilon band or other measure could be stored as an attribute of each object. The methods of modelling uncertainty presented thus far focus on describing the variability between where an entity is in space and where its database representation places it.

Attribute uncertainty is a different matter entirely. There are two basic types of attribute value: classified (nominal or ordinal) values such as soil classes, or continuous (cardinal) values such as elevation. Classified values are the more intractable of the two for uncertainty management. Although some classifications refer to easily-defined, sharply-bounded areas such as bedrock zones or lakes, in many cases of ordinal uncertainty there is no 'true' value for comparison, and problems introduced earlier such as internal purity, class boundaries, and sampling error increase the complexity of the problem of representation of a complex reality. This section describes some methods that are available to address ordinal attribute uncertainty. A number of these methods are drawn from expert system analysis, which focuses on 'ill-structured problems' using non-dichotomous structures (i.e., more-or-less structures rather than yes-or-no). The method of fuzzy set theory, which forms the basis for the model presented in the following chapter, is discussed in detail. The section concludes with a presentation of several alternatives for propagating uncertainty models.

The quantification of uncertainty has been studied primarily in reference to expert systems. Under development in many different fields, expert systems utilise a series of carefully formulated rules to come to a specific conclusion or offer a set of alternatives. Uncertainty metadata is required to navigate some of the more complex decision-making functions, such as determining the strength of rules, when to apply them, and how to resolve conflicts between the rules (Winston 1984).

The methods developed during the evolution of expert systems can also be applied in quantifying and manipulating classified attribute uncertainty in resource data. There are four approaches that have been commonly used to generate such metadata: Bayesian probability, Dempster-Shafer theory of evidence, non-monotonic logic, and fuzzy-set theory.

2.3.1.2.1. Bayesian Probability

Probability theory is the earliest formal approach applied to quantifying uncertainty and, therefore, has received the most attention in expert system design. This theory translates uncertainty into a rigorously formal definition easily utilised by expert system designers. The probability of a hypothesis represents a number between zero and one indicating the belief in that hypothesis. If we have an observation and wish to compute the probability that a hypothesis is true given that observation, we can do so if we have two items of information:

1. the likelihood that the observation will occur if the hypothesis *is* true; and
2. the prior probability of that hypothesis being true.

This is termed the 'conditional probability,' and is calculated using the following formula - Bayes Theorum (Stoms 1987):

$$Pr[H|D] = Pr[D|H] \times Pr[H]/Pr[D] \tag{2.1}$$

where H is the hypothesis and D is the observation.

The mathematical manipulation of probability with Bayesian methods has several drawbacks when applied to uncertainty. First, the theory assumes that probabilities can be assigned with great precision by experts in a consistent way; often an unreasonable expectation. Second, there exists no consistent and fully objective method of rating the probability assignments; some may be based on thorough research while others may simply be guesswork. When we know nothing the theory requires us to assume equal probabilities. A third criticism, voiced by Gordon and Shortliffe (1992), is that committing partial belief to a hypothesis commits the remaining belief to its negation—which can be counter-intuitive. Bayesian probability theory is therefore best at dealing with uncertainty due to randomness or variability rather than vagueness or imprecision (Stoms 1987; Zimmerman 1990).

2.3.1.2.2. Dempster-Shafer's Theory of Evidence

Dempster and Shafer's theory (Shafer 1976) focuses on the quality of evidence rather than truth of hypothesis. A zero to one rating is applied relating to the chance that evidence demonstrates the truth of the hypothesis. Evidence is accumulated to narrow down the hypothesis set using convergence of evidence. Two functions are applied: Bel[H] measures the probability that the evidence implies H; it therefore is the lower bound on the probability that H is true. A plausibility measure, PL[H] = 1 - Bel[not H], represents the upper bound—the degree to which the evidence fails to refute the hypothesis. The range between the two, [Bel[H],PL[H]], indicates the incompleteness of evidence for H due to uncommitted support. Utilising the notation [lower bound, upper bound], this implies the following:

- H[0,1] - no knowledge at all concerning H;
- H[0,0] - H is certainly false;

- H[1,1] - H is certainly true;
- H[.25,1] - evidence provides partial support for H;
- H[0,.85] - evidence provides partial support for not-H;
- H[.25,.85] - simultaneous partial support for H and not-H.

Support may be distributed among several hypotheses when evidence does not support a single one. The reliability of the source may be accounted for by discounting evidence for all hypotheses.

Dempster's rule of combination allows pooling of multiple pieces of independent evidence, focusing on the intersection of their independent conclusions. For example, a set of airborne multispectral data has been gathered regarding vegetative reflectance of a specific tree species. The probabilities for the cause of a specific anomaly have been estimated as follows:

- p(a) = .25 that the trees are water-stressed;
- p(b) = .15 that the trees are nutrient-stressed;
- p(c) = .40 that the trees are insect stressed; and
- p(d) = .20 that the cause is unknown (represents distributed support for the above).

The belief value for p(a) (the trees are water-stressed) = [.25,.45] (the second measure being derived from 1 - .15 [nutrients] - .40 [insects]). The calculated belief values could then be combined with other values generated from evidence such as soil samples or rainfall data utilising Dempster's Rule.

This theory has been criticised for its lack of attention to content of evidence. If two pieces of evidence conflict there is no mechanism for addressing the reasons for the conflict. Instead, conflicts generate indeterminate results or a diffusion of support among multiple hypotheses.

2.3.1.2.1. Non-Monotonic Logic

Non-monotonic logic utilises non-quantitative reasoning modelled along the lines of certain human decision-making processes. When evidence is lacking, a logical conclusion is to expect default values. Multiple inferences are allowed, generated from a set of default axioms (Cohen *et al.*

1985). A list of observations is also required, separated into those that would prove the assumption true, and those that would prove it false. These act as the rules that drive the system.

A system utilising non-monotonic logic begins the process of reasoning to a conclusion by making decisions based on rules. General rules such as 'increased timber production reduces wildlife habitat' might be applied at one particular branch point in a land-use decision. However, if this inference is proven false at some later point by an observed increase in wildlife, the system backtracks to the branch point that led to the false inference and continues searching until a consistent set of assumptions and facts are found.

This method has several drawbacks that limit its application in real-world situations: it contains no method of deciding which assumption to reject from a set of contradictory ones, nor does it recognise degree of conflict—a direct contradiction is addressed in a manner similar to minor inconsistencies. However, it does provide a method of dealing with incomplete evidence by making best use of defaults.

2.3.1.2.2. Fuzzy Sets

Fuzzy sets have a superficial similarity to Bayesian probability; they represent an uncertainty gradient using numbers between zero and one. However, fuzzy sets are considerably different in concept and, therefore, in application. The numbers represent a *degree* of membership in a set rather than the *chances* of probability theory. The implications of this 'degree of membership' mirror the nature of imprecise data, making fuzzy set theory a prime candidate for inclusion in an uncertainty model.

In 1965 L.A. Zadeh introduced fuzzy set theory to a sceptical audience of mathematicians. It has since blossomed into an industry that produces billions of dollars worth of fuzzy products. At the heart of the difference between classical and fuzzy set theory is something Aristotle called the law of the excluded middle. In standard set theory, an object either does or does not belong to a set; there is no middle ground. The number seven belongs to the set of odd numbers and not at all to the set of even numbers. This principle preserves the structure of logic and avoids the contradiction that an object both *is* and *is not* a thing at the same time.

Multivalent or fuzzy sets allow *degrees* of membership. Items belong only partially to a fuzzy set. They may also belong to more than one set. Fuzzy set theory does not contradict classical set theory, but acts as a generalisation in situations where the class boundaries are not, or cannot be, sharply defined. Applications are numerous: they allow a mathematical way to express vagueness in language, a structure for acting on imprecise information and, key to this discussion, a method of combining and manipulating imprecise input sets. For example, the concept of 'moderately well-drained soil' does not require a strict class allocation, but might be better served by a quantitative judgement that allows partial membership. 'Fuzzy logic' refers to the rules of manipulating these non-standard class functions as defined by the mathematics of fuzzy set theory.

As with class intervals in crisp sets, the choices governing membership functions in fuzzy sets determine the utility of the model. The function utilised should ensure that the grade of membership is maximised at the centre of the set and falls off in an appropriate way to the regions outside the set. Burrough (1989) utilises a common function that can be adapted to specific requirements:

$$\mu_B(x) = \frac{1}{\left\{1 + a(x-c)^2\right\}} \text{ for } 0 \leq \text{x} \leq \text{P} \qquad (2.2)$$

where a is a parameter governing the shape of the function and c defines the value of the property x at the function's centre. By varying the value of a, the form of the function and the position of the crossover point (usually 0.5—where the Boolean-style maximum likelihood would shift from one class to another) can be easily controlled. In Figure 2.4, the difference between fuzzy and crisp sets as well as variations in membership function parameters are illustrated. The first three models (a-c) show several interpretations of symmetric Boolean and fuzzy function comparisons. The latter two (d-e) show asymmetric functions. Other fuzzy concepts, such as 'very low' or 'close to' might be represented by decaying functions.

Fuzzy logic has found its primary application in control circuitry. The most famous is a subway car controller used in Sendai, Japan. Utilising fuzzy rules, each of which is defined as a membership function (e.g., 'apply more brake pressure when the train is moving downhill'), the system

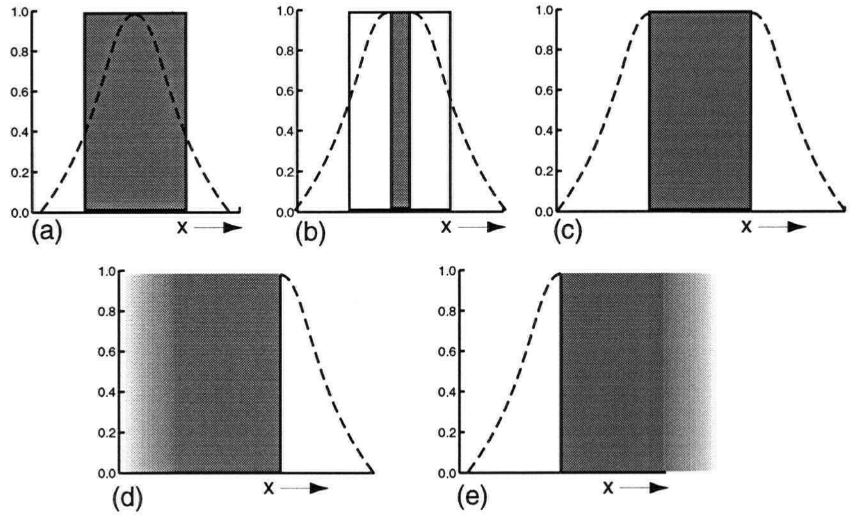


Figure 2.4. Boolean and fuzzy classification models. The attribute value (*x*) is graphed against the fuzzy membership function value. The broken lines show the envelopes of the fuzzy classes for each model, while the solid lines and shaded areas indicate the related Boolean sets. (Source: Burrough *et al.* 1992)

operates trains more smoothly and with greater energy efficiency than human operators (Kosko and Isaka 1993). Another area of application is in speech recognition: teaching computers to interpret the fuzzy concepts inherent in human language (e.g., Zadeh 1970).

In the context of natural resource management fuzzy logic has been primarily applied to two areas: 1) answering complex queries that combine Boolean maps using fuzzy rules and produce fuzzy output; and 2) using fuzzy membership functions to reclassify existing data and submitting the results to simple or complex queries. The former is illustrated by a number of land classification applications, including Zhang *et al.* (1988) and Wang *et al.* (1990). Fuzzy sets are combined with a multi-criteria methodology by Banai (1993) in reference to land classification. Mendoza and Sprouse (1989) describe the generation of forest planning alternatives utilising fuzzy rules, while Kollias and Voliotis (1991) use fuzzy rules to retrieve soil information.

The second type of application is demonstrated by Burrough's (1989; 1991) soil survey and land evaluation and McBratney and Moore's (1985) climatic classification model. Suryana (1993) applies fuzzy reclassification to erosion hazard management and land suitability mapping.

2.3.1.2.3. Linking Fuzzy Sets With Attribute Data

Published papers that link fuzzy set theory with geographic information can be divided into two broad classes. Most deal with fuzzy representation and analysis of site attributes. Classes may be assigned using fuzzy classifiers, supplied with fuzzy limits, or grouped and analysed with fuzzy logic. However, a few authors have ventured to broaden the application of fuzzy sets into the representation of the spatial distribution of geographic phenomena. Fuzzy boundaries, neighbourhoods, and contiguity of results are examples of the spatial applications touched on by these latter authors.

While there are undoubtedly numerous possible methods of deriving fuzzy membership functions, two distinct groupings have arisen in the geographic literature. Robinson (1988), drawing on terms coined by Buckles and Petry (1985), defined these two groupings as the Similarity Relation (SR) and the Semantic Import (SI) models.

The first is similar to cluster analysis and numerical taxonomy in that the value of the membership function is a function of the classifier used. Robinson termed this the 'Similarity Relation Model'. Fuzzy clustering, introduced by Ruspini (1969), provides a way around some of the representational difficulties of conventional clustering where, for example, stray points or 'bridges' between sets can cause problems. For example, the fuzzy-*c*-means method introduced by Dunn (1974) gives points membership values in inverse relation to their distances from cluster centres. Operational examples include McBratney and Moore's (1985) utilisation of the fuzzy-*c*-means method to perform climatic classifications, as well as the implementation of such clustering algorithms in the GIS package 'IDRISI' (Clarke University, Worcester, Maryland).

Such classification methods have been found to be most useful when performing exploratory data analysis; when the researcher has little information on classification training the SR model offers an alternative method of automatically grouping data. For further details on this computationally complex method the reader is referred to Bezdek *et al.* (1984) or McBratney and Moore (1985).

A mathematically simpler approach is to use an *a priori* membership function with which individual items may be assigned a membership grade. This is known as the 'Semantic Import Model'

(Robinson 1987; Burrough 1989). The concept of semantic import refers to this model's ability to represent semantic classifiers such as 'close to', 'nearly', or 'rarely' in numerical form. Natural resource scientists are often aware of such classifiers, but have become accustomed to translating them into precise cut-offs for typical numerical analysis.

A membership function can be structured in several ways. A symmetric function (Figure 2.4 a-c) might be useful for a situation where we want to distinguish between 'light' and 'moderate' rainfall. The parameters are adjusted to create proper centring and a smooth crossover between the two functions. An asymmetric function (Figure 2.4 d-e) might be applied when the function can be truncated on one side. 'Insufficient' versus 'sufficient' rainfall for crop growth might qualify as two asymmetric functions.

2.3.1.2.4. Combining Fuzzy Classifications

Logical models that assess complex issues such as land suitability for agriculture require that data from a variety of sources be combined in various ways. Boolean classifiers and Boolean logic have traditionally been utilised in everything from simple overlays to complex models of runoff and erosion (e.g., De Roo *et al.* 1989). For example, the Structured Query Language (SQL) interface built into many information systems allows class combinations using the operators AND, OR, NOT, etc.

Unfortunately, when data contain some degree of uncertainty, considerable information loss can occur in these strict combinations. Several studies (Marsman and de Gruijter 1984; 1986; Drummond 1988) provide examples of comparisons between derived attributes and ground checks. The 'quality' estimates in these studies were often abysmally low. However, the degree of misclassification was rarely serious because the attribute values causing the misclassification were often only slightly outside the defined class limits. The information existed; however, the map's quality was underestimated by the Boolean matching process (Burrough *et al.* 1992).

A multivariate fuzzy set does not produce such strict boundaries. Data that have been transformed from original observations to fuzzy class values can be combined with a single 'joint fuzzy membership function' or JMF (Burrough *et al.* 1992) as follows:

$$\text{Result(JMF)} = \text{MIN}(\text{MF}_A, \text{MF}_B, \text{MF}_C \ldots\ldots\ldots) \tag{2.3}$$

where MF_A, MF_B and MF_C represent the membership functions of three different spatially-concurrent attributes. The *minimum* value of each single membership function for each attribute value gives the JMF.

The great majority of existing papers in this field focus on applications of this classification/JMF technique. Two streams of research have emerged—the first focuses on physical applications such as land or crop suitability analysis (Drummond 1988; Burrough 1989; Suryana 1993), erosion hazard assessment (Suryana 1993) and soil-property analysis (Burrough 1989; Burrough *et al.* 1992). The second emphasises spatial decision support modelling, including generating decision-making alternatives (Mendoza and Sprouse 1989; Banai 1993) and linear programming (Mendoza and Sprouse 1989).

2.3.1.2.5. Cardinal Values

Resource data may be modelled using either discrete or continuous data structures. Furthermore, the data may be stored as ordinal classes or as cardinal values. Fuzzy sets, as well as the other methods presented above, are of use with either type of data structure, but focus particularly on ordinal classes. Data that are stored as cardinal values cannot be referred to using membership values or single probabilities. Uncertainty in these values is a numerical distribution, and must be simulated using a strictly numerical method such as probability distributions. Ordinal and cardinal values must be dealt with differently when uncertainty is propagated through modelling procedures. Although reduction to ordinal classes is a possible solution, this reduces the potential to model the data mathematically.

2.3.2. Propagation of Uncertainty

Study of the propagation of uncertainty through GIS and spatial modelling processes is an important, even crucial, task. There is considerable justification for such a statement. Goodchild (1991:121) notes that "currently we lack comprehensive methods of describing error, modelling its effects as it propagates through GIS operations, and reporting it in connection with the results." Lanter and Veregin (1992:825) note that most research into error modelling has been carried out "in isolation from the broader context of error propagation modelling in a GIS environment". For example, Fisher's viewshed error analysis (e.g., Fisher 1991a; 1992; 1994) or Veregin's (1996) buffer operation propagation error modelling concentrate on very specific, spatially oriented operations. Work on isolated operations provides necessary input; however, generic propagation modelling for environmental models requires more universal methods. A model of the propagation of uncertainty through spatial data processing may utilise two basic approaches (Joy *et al.* 1994). An analytical approach, such as that employed by Heuvelink *et al.* (1989), uses mathematical functions. However, standard propagation theory (Taylor 1982) restricts such mathematical analysis to operations that are continuously differentiable. The alternative is the Monte Carlo method.

2.3.2.1. ARITHMETIC PROPAGATION

The simplest way of propagating uncertainty through a model is through basic arithmetic relations. However, this only holds true for errors in cardinal values that are both random and independent (for examples see Burrough 1986a). A simple model such as A+B will yield significantly different error values than A-B, particularly if A≈B. However, when variables are correlated and the model involves a product or quotient, partial differentiation of a Taylor expansion is required (Taylor 1982). The rules of thumb for mathematical model error propagation (Alonso 1968) include suggestions to:

- avoid intercorrelated variables;
- try to avoid multiplication or division; and
- avoid as far as possible taking differences or raising variables to powers.

These limitations, though not absolute, place severe limitations on the development of environmental models. Nonetheless, arithmetic propagation is utilised in some studies. For example, Burrough (1993) uses a second order Taylor series to model heavy metal sediment levels in the Netherlands. The model used consists of one environmental variable and one terrain variable. Similar studies have been carried out by other members of Burrough's research group (e.g., Heuvelink 1995).

2.3.2.1. MONTE CARLO

A more universal method of uncertainty propagation is found in Monte Carlo simulation. This method, though computationally intensive, is totally independent of the uncertainty models used and the nature and sequence of GIS or modelling operations employed. It is generally applicable to error propagation problems in a GIS context (Openshaw 1989; Fisher 1996; Heuvelink and Burrough 1993).

The Monte Carlo method was introduced by von Neumann and Ulam during World War II as a code word for the secret work at Los Alamos. The method was applied to simulating random neutron diffusion in fissionable material. Later it was expanded to evaluating complex integrals or solving certain equations that were not amenable to analytical solutions (Rubenstein 1981). Monte Carlo methods now are applied in a variety of disciplines in order to solve complex problems—from radiation transport to river modelling. Recent leaps forward in computer processing power have made this 'brute force' method increasingly appealing.

Monte Carlo techniques allow for *replication* of an experiment. Replicating implies re-running the experiment or simulation numerous times with selected changes in the input parameters. Input data uncertainty is assumed to be characterised by an error model that represents reasonable estimates of the possible values. A single simulation involves randomly selecting a value from each of the input error models, completing a series of analytical operations, and storing the results. The process is repeated M times, and the M result maps are summarised to present some sort of confidence interval around the mean of all the simulations. The basic algorithm is as follows:

---

1) determine the types, levels and error characteristics of each source data set;
2) replace the observed data with a set of random variables drawn from an appropriate probability distribution designed to represent the uncertainty of the inputs;
3) apply a sequence of (GIS) operations to the data—uncertainties in models and equations may also be simulated by randomisation, if possible;
4) save the results;
5) repeat steps 2 to 4 M times; and
6) compute summary statistics.

An important consideration is determining the appropriate *M* value. Most authors suggest rather small values, on the order of 20-30. Openshaw (1989) notes that one should not place too much emphasis on significance tests of the results as classical inference really is not appropriate due to multiple significance testing problems and the absence of a formal experimental design. He suggests that any significance tests be used as a guide to action rather than a precise test of a hypothesis. If one makes the broad assumption that the target value (the Boolean result) is the most accurate, then he suggests stopping the simulation once the target result is ranked higher than fifth. It may also be appropriate to simply watch for the appearance of the target value in one of the tails of the output distribution, indicating a skewed result.

As a Monte Carlo simulation is totally independent of the uncertainty models and data manipulation occurring, it is an ideal platform for addressing propagation through complex GIS processes. Openshaw *et al.* (1990) ran such a simulation on a study of radiation waste dump sites. In this case, they choose to perturb spatial components of the data such as node and vertex locations. Fisher (1991c) focuses on soil inclusions: randomising grid cells in a soil coverage to simulate the differences between reality, where small soil inclusions occur, and the smoothed soil model used by standard Boolean processing. He offers uncontrolled (totally random) and controlled algorithms; the latter requiring knowledge of inclusion probabilities for the various soil types present. Emmi and Horton (1995) use Monte Carlo procedures to examine uncertainty in risk assessment for earthquakes, while Kunkel and Wendland (1997) use similar techniques to model groundwater residence.

The Monte Carlo technique is simple in principle; proper implementation is more difficult. It should be noted that, in the studies summarised above, a considerable amount of the simulator's effort has gone into choosing the randomised variables in a manner appropriate to the application.

2.3.3. Uncertainty in Continuous Data

The various error models discussed earlier can be utilised in either a discrete or continuous spatial data model. In a discrete model a single measure would typically be applied to each object.

For example, one number would describe an entire polygon's spatial uncertainty, while another number would describe its classification uncertainty. However, many environmental models implicitly assume a continuous spatial variation, as do interpolation techniques such as kriging. The continuous data model used in most GIS is a raster. Although raster structures discretise the continuous nature of the reality/models, they do so to a far lesser degree than the standard discrete entities of points, lines and polygons.

Fuzzy set theory has been utilised in three primary ways in geographic data analysis: fuzzy class memberships, fuzzy rules for class combinations or queries, and fuzzy spatial boundaries between entities. This latter application represents an important bridge between discrete and continuous models. There has been some research in this area, although, as Heuvelink and Burrough (1993) point out, there is little experience regarding how to deal with these transition zones between source polygons.

Polygon boundaries are represented as lines on categorical maps. This belies the fact that these lines are fundamentally different from all other geographic linear features. Works such as Thomas Poiker's classic "A Theory of the Cartographic Line" (Peucker 1975) focus on the relationships between linear elements on maps and their computer representations. Cartographic generalisation focuses on complex or dynamic linear features such as shorelines or rivers. The lines that represent boundaries between the classes in categorical maps have received considerably less attention. These types of boundaries exist in the real world to varying degrees. For example, forest types are divided by transition zones whose widths vary widely (Joy *et al.* 1994). Even before addressing positional error, the 'sharp' boundary between a clearcut and mature forest is a 'corridor of transition' that is a minimum of several metres wide. Soil type divisions other than bedrock boundaries or fault lines can exhibit considerably more variability in their boundary widths.

The inclusion of these spatial constraints in attribute classification is a necessity in a spatial uncertainty model. The traditional separation of 'attribute' and 'geometry' is consistent with an entity-relationship model of phenomena in which geometry defines objects which then have attributes and relationships (Mark and Csillag 1989). However, the two are tightly intertwined. The line geometry is still an artefact of the attribute classification process.

In Figure 2.5 four models of probabilistic functions are presented. The first two have traditionally been applied to line position; however, Mark and Csillag (1989) extend this model to include membership probabilities. They focus on developing the third model (c). This work is even more applicable to fuzzy membership functions, as a 'probability' of 0.25 admits a 75% possibility of *anything* else, whereas a similar fuzzy membership refers to a *degree* of class membership. Mark and Csillag assume that the probability of membership at any given point near the boundary can be approximated by some family of parametric curves. They admit the strong likelihood of an asymmetric function that varies with location along a boundary, however, they utilise a simple cumulative normal function as a first approximation. The fourth model presented in Figure 2.5 is the 'corridor of transition' model developed in Davis (1994) and utilised in the following chapter.

The shift from probabilistic line functions to a fuzzy 'possibility' is a simple one. On a multiple-category map the lines may be said to be the areas of *least* spatial attribute certainty. In theory, they represent a series of points where likelihoods collide. Utilising the fuzzy data model it is apparent that this line represents a series of points in space where, rather than there being an

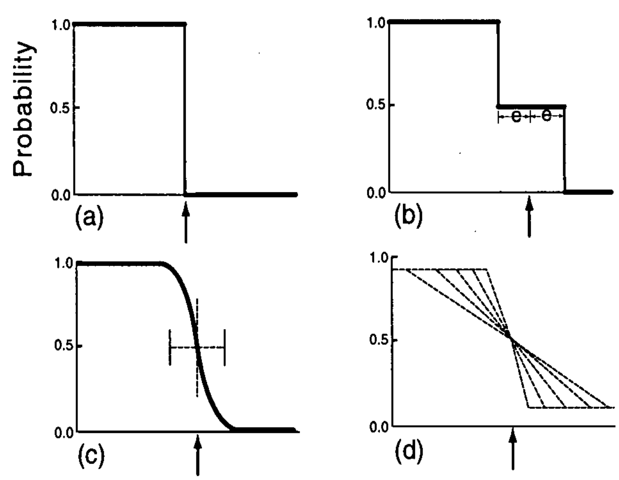


Figure 2.5. Four probabilistic functions of spatial boundary uncertainty: a) the sharp Boolean boundary; b) the epsilon model; c) the membership probability model; and d) the corridor of transition model. The arrow indicates the same spatial position (the Boolean boundary) in each. (Source: Davis, 1994; Mark and Csillag, 1989)

equal probability of one or the other being present, the membership values of the two equalise. Therefore, rather than defining a standard curve for the boundary model, the fuzzy classification techniques discussed earlier in this chapter can be applied to create an attribute-oriented fuzzy boundary model.

2.3.4. Summary

The above section has included discussions of modelling and propagation using discrete entities, continuous entities, spatial location, and the use of both ordinal and cardinal attribute data. This multitude of factors that make up uncertainty modelling and propagation serves to highlight the multidimensional aspect of this field of inquiry. There is no single answer—no 'uncertainty button' that can be tacked on to a GIS. The nature and context of the data and the target application define the type of uncertainty modelling and propagation necessary.

2.4. Communication of Uncertainty

Harking back to the big picture—the management of uncertainty for decision-making—the final step in the process is communicating this uncertainty to the target audience. When dealing with lumped models the uncertainty can be communicated with a simple numerical summary. It still may be a challenge to understand and communicate the *implications* of a standard deviation, fuzzy, or other similar measure; however, the dimensions of the problem are relatively limited in comparison to spatial data. A distributed uncertainty model will probably have a variable spatial distribution of uncertainty; higher in some areas than others, or higher for certain features. A graphical method of communication is necessary to communicate this information. This section briefly describes some of the background and current research in this field.

Research into uncertainty visualisation has gradually extended the graphic variables offered by Bertin (1985) into new realms opened up by computer displays. The first option is the use of static variables such as changing the 'focus' of a particular object to represent uncertainty (MacEachren 1992; 1994). Use of a colour variable such as decreasing colour saturation values or changing hue is another possibility. Other potential static variables such as fog or texture change are discussed by Goodchild, Buttenfield, and Wood (1994).

Dynamic cartography (map animation) offers advantages over static displays in terms of information density. Simple dynamic variables such as 'duration' in a flashing symbol map or shifting pixel map (Dibiase *et al.* 1992; Fisher 1994) may be used to express degree of certainty in displayed information. Allowing the user to toggle between the actual data and uncertainty information is another possibility (MacEachren 1994; van der Wel *et al.* 1994). A third alternative is displaying individual realisations of a model to express variability (Goodchild *et al.* 1994). This concept may be extended to a dynamic display of the full range of realisations. Displaying such realisations has the advantage of drawing attention to the possible effects of the uncertainty—a key issue for the user that lacks understanding of the statistical basis for the uncertainty model. For example, watching the slope stability values for the area above a road change from 'safe' to 'unsafe' has a more dramatic impact than simply reporting standard deviation values numerically.

A variety of variables are available for manipulation. Static visualisation might include the common technique of pairing uncertainty maps with the data they represent, or using Bertin's (1985) variables in combination to depict data and uncertainty in a single map—the best candidates being value, colour and texture (Goodchild *et al.* 1994). MacEachren's (1992) concept of defocusing symbols fits this category, as does Fisher's (1994) use of interactive sound. Another of Bertin's variables, arrangement, could be utilised to represent uncertainty as the z-axis in a perspective view.

In dynamic visualisation, some of the available choices include display animation using various realisations, Fisher's (1993; 1994) use of randomisation of pixels, the use of sound, text and images in a multimedia role to support map presentations (Beard and Buttenfield 1991), or interactive manipulation—where the user is involved in altering parameters and viewing results in real time (MacEachren 1994).

The exploratory nature of many of the studies cited above has typically led to exclusion of the human factor. Cartographic research has made steady progress in understanding how humans interact with static maps. Rapid advances in visualisation technology require equally rapid advances in understanding the psychology of these different displays. Some such advances are summarised in Hearnshaw and Unwin (1994). However, applications that focus on the display of uncertainty information have been rare. Numerous studies such as Evans (1994) are required to compare the efficacy of the many possible realisations of uncertainty display.

2.5. Uncertainty in Forestry Data and Models

Forestry inventory data, growth models and decision models are each subject to many of the uncertainties and issues presented above. A number of the uncertainty models have been applied to various components of the forestry resource sector; however, conservative attitudes and a certain amount of inertia in forestry agencies and businesses has meant slow adoption of new techniques. Uncertainty modelling has the potential to allow more timely and informed decisions, growth and yield models that better reflect reality, and inventory data that incorporates less abstraction and is easier to maintain.

Using the classification of uncertainty sources and issues presented in the previous section, the following is a summary of forestry-related issues and relevant research:

Positional uncertainty – Field sampling relies on accuracy in the absolute position of sample sites and plots. Some forestry research (Bilodeau *et al.* 1993; Lowell 1997) has examined alternative sampling strategies that potentially lead to reductions in attribute uncertainty. The spatial position of buffers and boundaries of forest stands are subject to uncertainties in the various data gathering techniques. The effect this has on area calculations has been examined (Magnussen 1996).

Temporal uncertainty – Inventory and sampling take place at variable spaced intervals. Between these times uncertainty gradually increases in items such as volume estimates via growth models. Therefore, the accuracy of the current inventory is dependent on frequency of update. The more sophisticated decision models predict change in social and economic systems. Uncertainty in these values also increases over time.

Attribute uncertainty—There is variability in the size-volume relations used in estimators. Considerable work has been undertaken in determining the nature of this variability (e.g., Cunia and Wharton 1986). Polygons are not homogeneous; there is also variability in species within a polygon. Accuracy assessments have been carried out by some agencies (California Dept. of Forestry 1992; BC Ministry of Forests 1995). Methods such as fuzzy classification have been proposed as a solution (Capra *et al.* 1995). Work on stochastic models such as Joy *et al.* (1994) deal with multiple uncertainties through Monte Carlo simulation,

Boundary uncertainty—The standard polygonal system of forest inventory distorts a more continuous reality. Although this simplifies data gathering and maintenance, accuracy of inventory suffers. There is also the issue of unrepeatable boundary delineation due to interpretation uncertainty. Research on fuzzy representation of boundaries has also taken place in the forestry sector. For example, Lowell (1993a) creates fuzzy boundaries using Voronoi area-stealing techniques, and evaluates them through comparisons with existing maps and samples.

Satellite classification uncertainty—As satellite images become a more substantial data source for inventory, uncertainty in classification becomes an issue. Work such as Capra *et al.* (1995) propose alternatives such as fuzzy classification techniques for forestry management.

Digitising uncertainty—During the transition from paper to digital maps digitising uncertainty was an issue. It is less so presently as more and more data are collected digitally and are available in digital form.

Propagation—Growth models and decision support models have grown considerably in complexity over recent years. Propagation of uncertainty through these models is an issue. Alternative models such as that proposed by Thompson and Vertinsky (1991) include uncertainties through recognition of spatial constraints and long-term dynamics. VanKooten *et al.* (1990) utilise Monte Carlo techniques to propagate uncertainty through a growth model, however, they found that a lack of data on variability hampered this type of modelling.

Generalisation—Data are generalised for various reasons in the course of inventory, producing summaries and setting up models, although there is little forestry-specific research on this issue.

Modelling with GIS—Growth models that were formerly 'lumped' (i.e., non-spatial) are gradually being incorporated into GIS as techniques become available (e.g., Thompson and Vertinsky 1991). This spatial orientation of forest models is becoming important for many reasons, from financial to ecological. Similarly, management decisions must be increasingly sensitive to spatial concerns. For example, the decision to harvest a particular stand may depend upon the stand's proximity to streams, areas of ecological concern, or urban areas. Outside of a GIS, incorporating a single spatial factor can be a major issue (e.g., Liu and Herrington 1996). Adjacency—the management of activities occurring in nearby stands—will also affect decision making. Critiques of simplistic decision-making models in use in the industry (Marshall 1986; Mendoza and Sprouse 1989) have led to studies and implementations that deal with economic issues (Cleaves 1994; Liu 1995), planting decisions (Reed 1991), land allocation (VanKooten *et al.* 1996) and overall risk management (Fight and Bell 1994; Mulder and Corns 1995).

Visualisation—The complexity of the current forest management decision-making environment makes data visualisation an important issue. The visualisation of both data and its associated uncertainty is often the only way to make sense of complex issues. Work such as Orland (1994) demonstrate this importance to forest planning and decision making

2.5.1. Uncertainty in Soil and Terrain Modelling

Both soil and terrain models are also crucial components of forest inventory data. Uncertainty in these data may affect growth models (through soil parameters, slope and aspect), profitability and harvest decision models (through accessibility factors), and slope stability, among other factors.

2.5.1.1. SOIL

Soil data are natural candidates for uncertainty studies. Pedologists have long been dissatisfied with traditional mapping systems, as soil polygons are poor representations of the continuous variation found in reality. Soil types are rarely delineated by sharp boundaries. There is also considerable uncertainty in sampling, since most data are inferred from vegetation types. Shifting analysis over to the standard discrete entities of GIS did little to ameliorate this problem. However, the digital environment offered new alternatives for both conceptual and physical data structures. Uncertainty modelling techniques could be used to represent spatial variability, attribute uncertainty, sample uncertainty, and uncertainty in classification. Soil scientists paved the way for many other applications that also deal with uncertain data.

Burrough's publications dealing with fuzzy sets for soil data (1986b; 1989; 1991) were some of the first applications of this methodology to GIS data. This work led to recognition of the utility of fuzzy methods in classification (Burrough *et al.* 1992; Odeh *et al.* 1992) and to experiments in visualising soil data uncertainty (Fisher 1993; Maclean *et al.* 1993). Soil sampling schemes were also improved through explicit recognition of variability (Domburg *et al.* 1994; McBratney 1994).

As techniques developed for the propagation of uncertainty in GIS-based models, soil data became a favourite example. Both attribute (Goodchild 1994; Stein 1994) and spatial characteristics (Rogowski 1996a; 1996b) of soils have been modelled using a variety of uncertainty models.

2.5.1.2. ELEVATION

Digital elevation data are used for a large and rapidly growing number of applications. In particular, elevation and its derivatives—slope and aspect—are used in forestry for growth models, slope stability models, models to visualise alternative strategies, viewshed modelling, and others. There are several options for gathering elevation data. Ground surveys are the most expensive, but also (potentially) the most accurate. Ground surveys may utilise traditional surveying techniques or, more commonly, GPS derived readings. Elevation data may also be gathered from stereo photogrammetry, or read directly from synthetic aperture radar (SAR) data.

Uncertainty in elevation data and its derivatives is almost always in reference to cardinal data, and since there is a 'true' value for reference (subject to geoid variability), the most common problem is numerical error. This assumes that the spatial location of an elevation datum is known absolutely; however, spatial uncertainty can also create variability. This latter problem may be addressed as with other types of spatial uncertainty (e.g., Monckton 1993).

Error typically propagates through a series of readings and operations on its way to a final elevation, slope or aspect value. For example, in surveying elevation, errors in each reading of a transect are cumulative, instrument error must be added in, and these errors are added to the error of the reference benchmark. However, these errors are minor relative to the errors generated through interpolation between sample points.

In photogrammetry, or its digital equivalent, error is introduced when the stereo correspondence produces mismatches. These can result from a variety of conditions, including low contrast, clouds, relief distortions between images, periodic terrain textures, or the presence of vegetation. Poor alignment of the images, the hardware, or the image software can also generate errors. Some of these problems can be reduced through post-processing techniques. However, post-processing can introduce its own subtle errors due to assumptions in the algorithms used (e.g., Hannah 1981). SAR data are particularly dependent upon registration and algorithmic accuracy; over the past decade algorithms have improved and errors have been reduced. It may soon join the ranks as a regular source of topographic data.

Elevation data are commonly interpolated into DEMs which can then be manipulated via GIS and used for a variety of applications. The interpolation and the manipulations also introduce uncertainty into the data. Interpolation uncertainty has been studied for a variety of procedures, including contours (Wood 1994), and satellite-derived data (Sasowsky *et al.* 1992). The accuracy of the DEMs themselves has been the subject of many studies. Typically, due to the difficulty in accurately ground-truthing the models, different sources are compared (e.g., Brown and Bara 1994; Garcia 1994; Felicisimo 1994), although the advent of accurate GPS data has led to some ground-truthed studies (Adkins 1994).

The surfaces derived from DEMs are typically used as input to various environmental models. These uncertainties have been studied by themselves (Skidmore 1989), and in relation to viewshed modelling (Fisher 1991a; 1992; 1994), feature extraction (Lee *et al.* 1992), hydrologic modelling (Bruneau and Gascuel-Odoux 1995), and fire modelling (Delavar 1997) among many others. Monte Carlo simulation has been applied to DEM error propagation modelling (under the new name of 'stochastic imaging'), in which constrained random values are applied to the entire surface (e.g., Flamm and Turner 1994; Journal 1996).

2.6. Research Gaps

There are a number of specific aspects to uncertainty management that have been insufficiently researched. Overall, there is a glaring lack of integrated studies that examine 'cradle-to-grave' uncertainty—from sources through to final modelling such as decision support. It is rare that a specific modelling routine will host only one type of uncertainty. Spatial, temporal and thematic uncertainties of both cardinal and ordinal data interact in unpredictable ways in complex environmental or decision models. It is important to study these interactions in real-world scenarios.

2.6.1. Uncertainty Modelling

Uncertainty modelling research is still in its infancy, and justification for research in this field is not difficult to come by. In a general sense, the importance of the accuracy issue in spatial data is epitomised by that issue being the subject of the first of the 12 research initiatives undertaken by the National Center for Geographic Information and Analysis (NCGIA, 1989). More specifically, Burrough *et al.* (1992) point to the lack of work on the problems of fuzzy spatial mapping: boundaries, neighbourhood and contiguity analysis. In an earlier work, Burrough (1986b) noted that:

> It is remarkable that there have been so few studies on the whole problem of residual variation and how errors arise or are propagated in GIS processing, and what the effects of these errors might be on the results of the studies made. (p. 103)

In his summary paper for the 'bible' of spatial accuracy ("Accuracy of Spatial Databases", Goodchild and Gopal 1992), Stan Openshaw (1992) states that:

> There is clearly an urgent need for basic research to resolve many of the [GIS error and propagation] issues. [This includes] developing a better understanding of error propagation through spatial databases, identifying and classifying operations most sensitive to error, and providing basic tools to handle error in a variety of situations. (p. 264)

He goes on to indicate the necessity of utilising the latest high-speed processing technology to investigate the extent of the problem in real-world applications.

2.6.2. Validation

One aspect is almost completely lacking in uncertainty research: validation of uncertainty models. General evaluations such as Eberbach (1993) and Liu and Herrington (1996) that look at the consequences or specific costs of uncertainty base their conclusions on the propagation of uncertainty values. These may be derived in numerous ways, including expert evaluations via semantic analysis, variability data recorded in field surveys, or simply rough estimates. Mathematical propagation models such as those discussed in Burrough (1986b) and implemented in Heuvelink (1995) can produce incredibly large uncertainty values if subtraction of inputs is part of the model. Other propagation models are based on numerous assumptions, such as the validity of stochastic landscape simulation (e.g., Journal 1996) or the Gaussian distribution of error terms. It is rare that uncertainty in the model inputs is validated; rarer still that outputs are confirmed.

The primary difficulty is a lack of techniques for sampling uncertainty. Dealing with cardinal values is a straightforward matter, though certainly time-consuming. One simply compares reality with samples and generates an error distribution. However, if both spatial and attribute uncertainty are addressed, this sample point may actually be located elsewhere, complicating matters considerably. Ordinal data are even more complex. For example, how does one confirm a fuzzy membership value in a soil class? How can boundary uncertainty values be confirmed under both spatial and attribute uncertainty?

If the results of uncertainty models are to be considered useful in decision support, there must be some methods available to indicate that they tell the 'truth'. Although it is accurate to state that a particular uncertainty model or representation is more 'honest' than standard Boolean methods (e.g., Lowell 1993b), it would assist the credibility of this research field if it were possible to compare the degree of 'honesty'.

2.6.3. Linking Uncertainty Management with Decision Making

Although uncertainty *in* decision making is a common research area, links *between* data uncertainty and decision making are rarely investigated. There are two major aspects: developing methods of summarising uncertainty for decision makers (and evaluating these methods), and developing methods for evaluating the effectiveness of decisions made using such inputs.

The first of these aspects focuses primarily on the visualisation of uncertainty. MacEachren (1992) indicates that cartographers have spent little time investigating methods of presenting uncertainty information. This is also indicated by a 200 page report on the "Visualization of Spatial Data Quality" (Beard and Buttenfield 1991) as part of another NCGIA research initiative. The bulk of the position papers in this work indicate a need for applied research in this area, the major impediment being that uncertainty models are required before they may be visualised. The models themselves are perhaps the most neglected area of research, yet the models, the propagation algorithms, and the visualisation methods can only be effectively developed in concert with each other. Evaluation of their effectiveness is also a crucial issue.

The second aspect—evaluating the effectiveness of decisions made using uncertainty information—is also a neglected area. This may be principally due to the immaturity of this research field; however, the lack may also be due to the immensity of the problem. It is difficult to compare the implications of decisions made under differing information environments without performing substantial double-blind experiments. Before committing resources to such experimentation, it is important to have a thorough understanding of uncertainty management at all levels. Therefore, while this current research examines some of the implications of decision making using uncertainty management, verification and evaluation is not implemented for the reasons discussed above.

2.7. Summary

This chapter has attempted to summarise a number of related research fields that, together, focus on uncertainty modelling in spatial data and natural resource inventory. A number of research gaps have been noted, including the need for better understanding of uncertainty behaviour and propagation in spatial databases, the need for applied research to allow interpretation of uncertainty, and a requirement for validation and applied tests of uncertainty models.

The following chapter makes the first inroads into these research gaps by presenting an uncertainty model developed in previous research undertaken by the author. The basic components of this model are used as data sources during the new research described in upcoming chapters. Rather than introduce these components piecemeal into the discussion of new research methodology, a summary of the previous research is presented in the following chapter to facilitate understanding of data lineage.

Chapter Three

Modelling and Storing Measures of Uncertainty in Inventory

3.1. Introduction

There are a number of uncertainty models that could be used in the process of developing sampling, evaluation and verification techniques (e.g., Fisher 1991, Heuvelink 1995). Much of the work undertaken in this dissertation makes use of an uncertainty model developed by the author during previous research (Davis 1994; Davis and Keller 1997a). The purpose of that research was to develop the basic components of an uncertainty model that could parallel a standard process model. The theoretical data model used was fuzzy sets, implemented in a raster environment. The principal focus of the research was on dealing with the troublesome conversion of standard vector and raster data, based on a Boolean concept of reality (something either *is* or *isn't*), into a fuzzy representation. The process model utilised in the test case of the research was the infinite slope stability model.

Upcoming chapters will make use of various components of the uncertainty model. Building on the groundwork established in the previous research, Chapter Four will focus on the development of uncertainty sampling and evaluation techniques (applicable to a variety of uncertainty models), while Chapter Five will include evaluation of uncertainty model output. Throughout the remainder of this dissertation, the fuzzy-set data concept (introduced in the previous chapter and further discussed herein) is used extensively.

In these upcoming chapters, the specific details of the uncertainty model used are not of great importance during discussion of the sampling, evaluation and verification concepts and of the development of theoretical techniques (i.e., various models could have been used). Nevertheless, when these concepts and techniques are *implemented*, an understanding of the examples presented in upcoming chapters will require some degree of understanding of the details of the underlying data model. Therefore, this chapter provides a capsule overview of the previous research. Sufficient detail is provided for the reader to understand the purpose and major components of the model. Minor components are, for the most part, omitted. The reader is referred to the original document for clarification (Davis 1994).

To summarise, the previous research included two major stages:

1) Making use of information such as expert opinion and published statistics, metadata was generated that describes uncertainty in a number of data layers (including soil type, percent slope, and ground cover). These metadata focus on the spatial variability of uncertainty, rather than simply summary measures, and are therefore described as an 'uncertainty model' of input data.

2) By utilising a combination of fuzzy set analysis and Monte Carlo simulation, an uncertainty model for slope stability was developed. The uncertainties modelled in stage (1) were propagated through a slope stability modelling procedure.

In order to address the various *types* of uncertainty, the research used a variety of techniques, including constrained DEM randomisation, the coding of expert opinion as fuzzy classifiers, fuzzy set manipulation, and uncertainty parameter estimates from compiled laboratory data. Two major shortcoming were noted in the conclusion: 1) the visualisation of the results is a key component of understanding such a complex dataset—work is required on designing and implementing effective visualisation tools for this type of uncertainty data in order to feed into work on decision making under uncertainty; and 2) although a number of alternative uncertainty models had been generated by various researchers, there existed no way of evaluating or comparing these models. Both of these items require considerable research work—both in laboratory development and field data gathering—and were felt to be well beyond the scope of the original project.

3.2. Methodology

The central focus of the previous research was the production of spatially variable uncertainty estimates in the output of a typical resource modelling procedure: slope stability assessment. This procedure was chosen due to the variety of data types required as input: soil and forestry polygons, DEM and slope surfaces, and laboratory-derived soil attribute data. The methods developed emphasised three new elements: 1) an asymmetric spatially-variable polygon boundary model termed the 'corridor of transition model'; 2) refinements to and applications of a theoretical DEM randomisation procedure (proposed by Goodchild 1980); and 3) the combination of fuzzy values and variability data in the same modelling procedure.

The first stages of the modelling procedure required that each of the major uncertainties in the inputs be identified and numerically modelled. These are identified as follows:

1. Classification uncertainty in classified values such as soil type or forest type;
2. Data collection uncertainty (e.g., 10% of polygons misclassified);
3. Spatial uncertainties (e.g., certainty in classification decreases near polygon boundaries);
4. Error envelopes around derived items (e.g., soil cohesion for Type 1 = $X \pm A$); and
5. Error envelopes around continuous mapped values such as elevation.

These five general groups can be split into two major types. The first three focus on uncertainties in classified values, while the latter two focus on error—where cardinal numbers are available (or can be derived). The two groups differ conceptually, and so must be addressed in different ways.

The first three deal specifically with classes and polygons. One of the principal issues that had to be addressed revolves around the polygon data structure. Polygons are somewhat appropriate in describing information such as forest cutblock boundaries, but less useful in describing the boundaries of tree stands, and even less useful in delineating the distribution of soil or slope. In the process of reducing a continuous reality to a polygon data structure, a great deal of uncertainty is introduced into the data. The first task in generating an uncertainty model for polygon data is to model exactly how much uncertainty is generated in this process and where it is located.

3.2.1. The Corridor of Transition Model

Uncertainty in resource data is often spatially variable, and in such cases would be best represented by a continuous surface. However, the most common method of storing such resource data (polygons) is a representation that contains abrupt transitions between homogenous areas. The 'corridor of transition' (COT) model is a term developed to describe a procedure of estimating both the level and spatial location of uncertainty generated by the data reduction process leading to polygon formation. Essentially, this procedure takes a polygonal surface, information about what assumptions were used in forming the polygons, and information about the data gathering and classification system, and generates information about the level of uncertainty in each type of polygon at every point in the mapped area.

3.2.1.1. THE SEMANTIC IMPORT MODEL

Ideally, most of the required information would be available from the original data used to generate the polygons. However, it is rare that such information would be available to the typical resource modeller—who normally has to make due with a map generated by others, and possibly some minimal metadata. The modeller could go out in the field and perform extensive resampling of the data; however, as before, this scenario is unlikely given the costs of current sampling techniques. While less accurate, a reasonable substitute can be found in expert knowledge regarding the resource (note that in the following discussion, soil data are used as an example; the procedures discussed are equally applicable to ground cover data or any other spatially variable phenomena).

Soil scientists, familiar with the area being modelled, can be polled for information regarding the intermixing of soil types, the characteristics of soil inclusions, and other generalised data. Burrough (1989) utilised such methods in a study of fuzzy soil classification. His 'Semantic Import Model'

(Robinson 1988) was utilised in the previous research to translate semantic classifiers into numerical form. Similarly, textual soil survey results often contain such non-mathematical data; but such information is lost when the soil maps are produced. Using the techniques of Semantic Import, this quantitative information can be translated into a fuzzy index or Certainty Factor (CF), where $0 <= CF <= 1$. This technique has been used in a number of soil survey studies, including the Robinson and Burroughs work noted above, as well as by Suryana (1993), who modelled crop suitability using expert opinions on the certainty of a variety of soil factors.

For example, the soil class 'sandy/silty morainal blanket' is described as 'very easily confused' with the class 'silty morainal veneer'. In this case, the term 'very easily confused' would be translated as a low certainty factor: '0.2' when quantifying the classification certainty of one relative to the other. A scale of phrases linked to numerical values is utilised (for more information on the method see Robinson 1988 and Davis 1994). In the same way, the level of certainty in boundary delineation is also captured. For example, there is generally high certainty in delineating the boundary between a bedrock extrusion and another soil class, but low certainty when the boundary is between two surficially similar soil types. Note that the data derived from Semantic Import were not used to replace existing quantitative data, but to enhance detail regarding the level of certainty to be assigned to these numbers.

The concept of 'certainty factor' is simply another name for fuzzy set theory. Fuzzy sets are a way of quantifying *degrees* of membership in a set. If the 'set' in question is the class 'sandy/silty morainal blanket', then this number would refer to the degree to which a particular sample is *like* the ideal class. It may also refer to our degree of certainty that the value at a particular point would fall into the bounds of this class. Fuzzy set theory is introduced in greater detail in Section 2.3.1.2.4.

The behaviour of fuzzy thematic indices (the 'certainty factor') in the 'transition corridor' between polygons can be referred to as a thematic spatial interaction model. Spatial interaction models, such as the gravity model, have a long history in human geography. The concept has application on the physical side of the discipline as well. The transition corridor model is defined as follows.

3.2.1.1.1. Polygon Centres

In addressing the areas that bound a polygon, it is necessary to focus on zones both inside and outside the polygon. In defining such zones, a very precise definition of a polygon's spatial structure is necessary—with particular emphasis on the polygon centre. One basic assumption of this transition corridor model is that polygons are most similar to their classified type at their physical centres, while least 'pure' at or at some distance beyond their borders, subject to the constraints of the Semantic Import (SI) tables (such as Table 3.1). This contention is supported by Burrough *et al.* (1992) in reference to soils, while Joy *et al.* (1994) support such a model regarding forest stands. The physical centre of the polygon is an important part of this model, so adopting it requires defining what the centre actually is. Circular polygons have an easily determined centre; however, the more common irregular polygon does not.

The most common measure of polygon centrality is the *centroid.* The centroid of a polygon does not necessarily define a practical spatial centre—a non-circular polygon can easily have a centroid located outside its borders. A method used by Lowell (1993a) in which centroids were digitised at the visual centres of a polygon (by eye) is impractical in a real-world data set—too many polygons exist for such a manual technique. In automating the procedure it became necessary to redefine the polygon in a way that accounts for irregularities in its shape. For example, in computer image analysis / object recognition research it is often necessary to 'skeletonise' an image segment in order to compare (and recognise) a generalised shape (e.g., see Choras 1993, Brandt

Soil Type	1	2	3	4	5	6	7	8	9	10
1	0.85	0.35	0.25	0.15	0.15	0.10	0.15	0.15	0.15	0.35
2		0.85	0.25	0.15	0.15	0.10	0.15	0.15	0.15	0.35
3			0.90	0.15	0.35	0.10	0.15	0.15	0.15	0.25
4				0.80	0.15	0.10	0.40	0.40	0.35	0.25
5					0.90	0.10	0.15	0.15	0.15	0.20
6						0.90	0.10	0.10	0.10	0.25
7							0.80	0.40	0.35	0.25
8								0.80	0.35	0.25
9									0.85	0.25
10										0.90

Forest Type	1	2
1	0.95	0.1
2		0.95

Table 3.1. Misclassification matrices derived from the SI model. The value at c_{ij} represents the possibility that type *I* would be misclassified as type *J*. The chart's trace contains the maximum certainty values.

1994). A skeleton is composed of lines that run parallel or at right angles to the contours defining an object's 3-D shape. Another relevant application is the construction of extended Voronoi networks (Gold 1992). These networks capture the spatial relationships between objects in an intuitive manner. Although current research has implemented them for points, extensions of the concept may be applicable to defining the relationships between the points surrounding a polygon.

In the current application, although the polygons are only 2-D, the effective 'shape' of a polygon can be defined by treating it, in essence, as a 3-D object. An even slope is drawn from the boundary to the central region. Where the various slopes intersect—there lies the centre. The centre now is defined as a line, rather than a point, with varying degrees of centrality. The method utilised in this 'corridor of transition' is as follows.

A spatial model was defined in which the 'centre' of an irregular polygon is delineated by a series of points that are located the maximum-minimum distance from any polygon boundary; that is, each point in the series is as far from an edge as it can be without being closer to another edge. As illustrated in Figure 3.1, a circular polygon has only one point at a maximum-minimum distance from its boundaries. A 'sausage' polygon has a clear line, each point on which is the same distance from its closest boundaries. In the case of the irregular polygon, the centre ridge is more complex. If the distances to the edge are treated as an elevation (using a slope of one), the resulting ridge-like structure can be visualised as shown in Figure 3.2. The elevation of the ridge varies with max-min distance to the nearest boundary.

Utilising a raster surface model, the algorithm used to produce this 'max-min ridge' creates a stepped surface by filling the cell from its boundaries inward. The details are as follows:

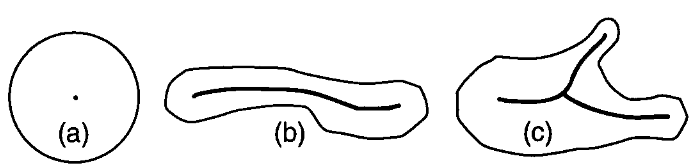


Figure 3.1. Alternative centroid models for variable polygon shapes

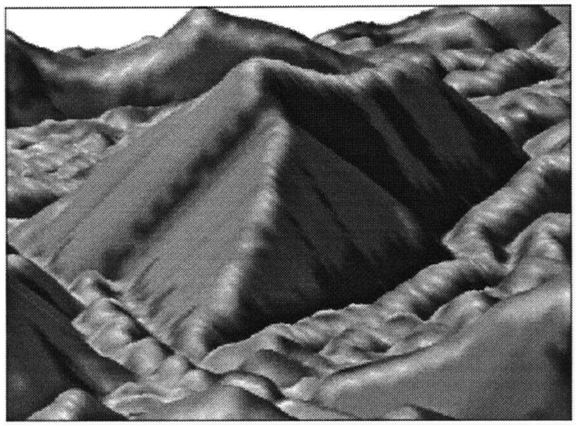


Figure 3.2. The variable ridge model of a polygon's centre, using the z dimension and a perspective view for illustrative purposes. The central ridge, and its relative height, defines the local thickness of the polygon and provides a 'direction to centre' for later algorithms.

---

1) Produce a raster surface of polygon boundaries (set at '1'); all other cells remain empty;
2) Begin a loop through all cells with Current_Depth initialized at '2';
3) If Current_Cell is empty AND adjacent to a filled cell AND the filled cell's value = Current_Depth -1 (i.e., the current cell is the next step up on the stepped surface);
   - 3a) Fill Current_Cell with Current_Depth;
   - 3b) If an adjacent cell meets the criteria of (3), move and fill as above;
   - 3c) Repeat until criteria cannot be met; (note that special processing is required for very narrow structures)
4) Repeat for all cells in surface;
5) Increment Current_Depth and repeat loop (i.e., next step up in the surface);
6) Continue until all cells have been assigned a value;

Although similar to a raster distance function (i.e., 'distance from every pixel to the nearest boundary'), this algorithm provides a more evenly stepped surface as required by later procedures, particularly in areas with narrow polygon extensions.

3.2.1.1.2. Polygon Boundaries

With the polygon centre defined, the next stage is to incorporate the 'expert opinion' information into the boundary model. To reiterate, the boundary between two polygons is assumed to be the point where the possibility of a sample falling into one or the other category is equal. If one polygon is Soil Type 5,and the other Type 2, a sample at their bounds would have equal possibility of being 5 or 2. The boundary model deals with what the possibilities are in all *other* areas. If, for example, the two types are very dissimilar, and could be easily distinguished during a soil survey, then it is likely that the boundary is well placed, and a sample taken a short distance within the Type 5 polygon would likely show Type 5 soil. In contrast, if the two soil types are quite similar, then the opposite would be true. By applying the 'expert opinion' data coded via semantic import, the entire surface of the (originally) polygonal map can be coded with the 'certainty factor' that quantifies this possibility for every soil type. This procedure must also capture other information into the boundary model, such as overall classification error.

Standard boundary models (Figure 2.5) are inadequate to the task of incorporating the certainty factors—also termed 'fuzzy metadata'—defined by the Semantic Import (SI) work. Even Mark and Csillag's (1989) parametric functions (Figure 2.5-c) suffer from an (admitted) series of broad assumptions—chief of which is symmetry. There is no reason to assume that two classes will blend evenly at their polygon's boundaries with a similar membership function slope on both sides.

In defining the membership function(s) that occur around the polygon boundaries, four primary items are of interest.

1. What are the certainty factors involved in classifying each cover type? This will indicate the maximum certainty that can be associated with a particular type. For example, even in the centre of a polygon there would still be some degree of uncertainty in the class due to variability on the ground.
2. What are the minimum certainty factors; that is, what is the likelihood that Soil Type B has been misclassified and is actually Soil Type A? In this case the focus is on misclassification rather than variability. As this will vary for different soil-type relationships, a matrix of values is required.
3. How do two spatially adjacent types interact in the transition corridor? If the two are similar there may be a gentle gradient, while dissimilar types may have sharper barriers. For example, on the boundary between two similar soil type polygons, there might be a large area

where a random sample would have similar possibilities of showing one class or the other. This information also requires a matrix of values.

4. The attribute 'blurring' in the transition corridor between polygons is likely dependent upon polygon size. An estimation of the size factor is also required.

Utilising such SI-derived data, the following model was defined (Figure 3.3). Keep in mind that the 'slopes' referred to in the definition are changes in the certainty factor of a classification, and not a measure of actual intermingling (although the two might coincide).

For a given point on a polygon boundary the directions of internal and external slope were determined from the max-min ridge map. The maximum intrusion distance was determined from the values derived from the SI matrix. One effect of this limit is to cause small polygons (with a maximum width below this distance) to be 'influenced' (shift in the internal fuzzy structure) by their larger neighbours.

A final intrusion limit was also determined from the SI data. For example, a value of 0.9 sets the final intrusion distance as 90% of the distance to the maximum intrusion line. The SI data were

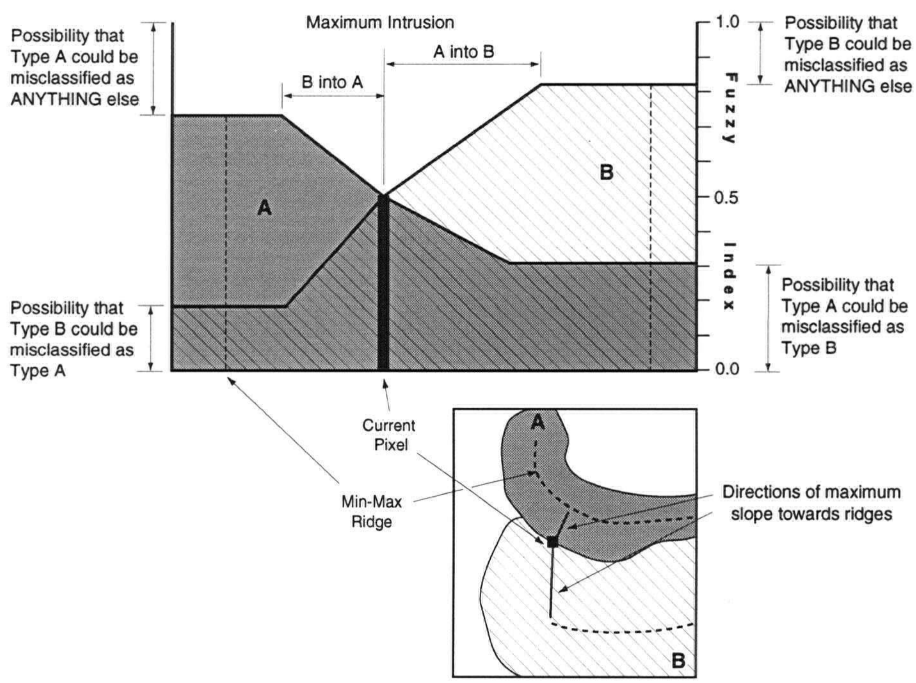


Figure 3.3. The 'corridor of transition' model for spatial boundary uncertainty – cross-section and plan views.

also used to set the extrusion distance in a similar manner. Note, however, that the spatial constraints on the extrusion result from the characteristics of the adjacent, rather than the current, polygon.

The internal and external slopes were calculated and applied to the fuzzy surface using a 0.5 index at the polygon boundary. At this boundary an idealised sample taken from the surface should be 'equally' similar to the polygons on both sides of the line. This is termed the 'coinciding of possibility' by Mark and Csillag (1989), or, in Boolean-style probability terms, there is a 50% chance of a sample belonging to either polygon.

The following algorithm was devised, making use of 'soil type' as an example.

---

1) Create a surface for each soil type, where 0 <= Cell_Value <= 1;

2) Initialize surfaces using SI values. For every cell in every surface:

   2a) Assign the related value from the SI table. For example, if Current_Surface is Type B, and Current_Cell was originally assigned Type C, then use SI value for possibility of C being misclassified as B. If Current_Cell was originally assigned B then use max. certainty factor (SI trace).

3) For every surface; For every cell in Current_Surface:

   3a) If Current_Cell is a polygon boundary, assign it 0.5 and determine number of adjacent polygon types. For each Adjacent_Type:

      (i) Calculate distance to internal 'ridge' from Current_Cell using the line of max slope from the max-min ridge surface.

      (ii) Calculate internal slope, where rise/run = maximum internal fuzzy value / SI-derived max intrusion distance (see Figure 3.3).

      (iii) Loop around Current_Cell, assigning SI values based on slope and distance to Current_Cell; When inside a Current_Surface polygon, overwrite if < current value, when outside a Current_Surface polygon, overwrite if > current value;

In summary, this 'transition corridor' procedure has taken the original Boolean model of polygons, where there are sharp bounds between class A and class B, and—using a number of values derived from expert opinion—created a constrained smoothing between classes. If the 'certainty factor' or fuzzy surface is visualised using a perspective view, the results appear as in Figure 3.4

and 3.5. These views represent the certainty in one particular class. In the first figure a standard (Boolean) soil class map is shown. The ridges are areas where the polygons are classified as soil class 1 (sandy/silty morainal blanket), while the valleys are areas that are *not* soil class 1. In the second figure the transition corridor model has been applied to the original polygons. High degrees of classification certainty only appear in the central regions of the original polygon structures. Note that this type of surface is generated for *each* of the soil classes.

3.2.2. DEM Randomisation

Classification uncertainty and how this uncertainty varies over space has been dealt with above. The next issue in addressing uncertainty in the inputs to the slope stability model is the latter two of the five general classes of uncertainty—the error envelopes for spatially distributed values such as slope, and values that are attributes of classified values, such as soil cohesion.

In this section the principal issue is *error*—a far more common problem than the modelling of uncertainty dealt with above. Error propagation can be addressed in two basic ways: mathematical functions (e.g., Heuvelink *et al.* 1989), or a Monte Carlo method. Standard propagation theory (Taylor 1982) restricts mathematical analysis to functions that are continuously differentiable. Though computationally intensive, Monte Carlo methods are considered applicable to error propagation problems in a GIS context (Openshaw 1989; Heuvelink and Burrough 1993).

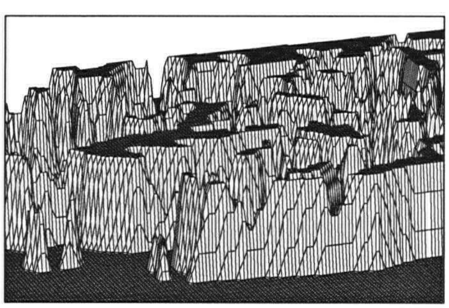


Figure 3.4. Perspective view of the fuzzy surface representing soil type 1 prior to implementation of the transition corridor algorithm.

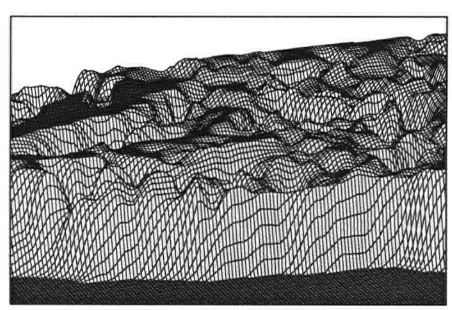


Figure 3.5. Perspective view of the fuzzy surface representing soil type 1 after applying the transition corridor algorithm.

In Monte Carlo simulation the model is run using the standard set of inputs, and then run again using a new set of inputs that have been randomised within the error envelope for each type of data. Rather than deriving a single output, the output becomes a distribution, and the overall variability can be ascertained. Attribute values such as 'soil cohesion' can be randomised in a straightforward manner. However, spatially distributed values such as slope must be treated slightly differently due to autocorrelation. The following procedure is an extension of one originally proposed by Goodchild (1980).

A typical continuous digital elevation model (DEM) is derived from spot height data. The errors associated with these spot heights are normally available as accompanying metadata. If Kriging is utilised to generate the DEM, then variance values for every cell in the model can be saved. Combining this variance with the original spot height error, a final error value can be generated for every cell. It is assumed that this error follows a normal error curve (Goodchild 1980; SRMB 1990).

By making use of this error curve and a constrained randomisation procedure, it is possible to generate an 'equally likely' elevation surface in which each cell is provided with a new height (within its error bounds). However, there is one major problem. Working with individual cells ignores the autocorrelation present in an elevation model. If two adjacent cells are assigned values from opposite ends of their error envelopes, artefactual roughness has been created, decreasing the overall autocorrelation index. This problem was addressed by Goodchild (1980). However, the original algorithm suggested by Goodchild brings the original and new autocorrelation indices together through a constrained random swapping of cells. However, this algorithm was intended for datasets other than DEMs. The procedure described here utilises a series of constrained smoothing passes over the dataset to gradually reduce short-range variability. In so doing, it lessens the artefactual peaks and troughs generated in the randomisation procedure. In the end, a new elevation surface is created in which each cell's value falls within the original elevation value's error envelope, and the autocorrelation index of the entire surface matches the original surface to within a stated tolerance.

This process uses the following algorithm:

---

1) Generate a variogram from the original DEM and fit a curve;

2) Generate a blank DEM. For every cell:

   2a) Determine distance (D) from cell to closet elevation spot height.

   2b) Substitute D in variogram curve, and derive std dev for Current_Cell (SDc);

   2c) Add SDc to SD of closest spot height, giving actual SD for cell's elevation (SDe);

   2d) Generate a normal-constrained random number based on SDe, using the original DEM value as a mean;

3) repeat for every cell in the DEM;

4) determine a spatial autocorrelation index (Moran's Index: I) for the original DEM (I_{orig}) and for the 'equally likely' surface (I_{el});

5) If $|I_{orig} - I_{el}|$ > specified tolerance T;

   5b) Smooth the new DEM (using 0.1 * | Old_Value - 9-cell-window mean |) and repeat (5);

3.2.3. Combining Error and Uncertainty

Application of the procedures developed in the previous two sections resulted in data structures that carry considerably more information content than the original polygonal maps and raster DEM. This information might be used to provide uncertainty estimates in simple GIS queries such as 'area of class A' or 'what is here?'

The fuzzy datasets generated through the transition corridor procedure can be combined using a fuzzy math function known as the 'Joint Membership Function', described in more detail in Section 2.3.1.2.4. The error values can be combined using a Monte Carlo procedure as described above. However, more complex resource modelling procedures require the development of methods of combining fuzzy class membership data and cardinal error data. These procedures are presented in the context of slope stability modelling.

3.3. Slope Stability Modelling

The infinite slope stability equation is a commonly used measure of the stability of surficial materials. This model utilises data with a variety of potential appended uncertainties, namely soil cohesion and other soil properties, forest cover and root depth, and slope of the soil plane based on an elevation model. The resulting 'factor-of-safety' value is a relative number only—comparable only within a particular application.

$$\mathrm{FS} = \frac{C_r + C_s' + \cos^2\alpha\left[q_0 + \gamma(D - D_w) + (\gamma_{sat} - \gamma_w)D_w\right]\tan\phi'}{\sin\alpha\cos\alpha\left[q_0 + \gamma(D - D_w) + \gamma_{sat}D_w\right]} \tag{3.1}$$

FS	=	factor of safety	α	=	slope of ground surface
D	=	total soil thickness	D_w	=	saturated soil thickness
C_r'	=	tree root cohesion	q_0	=	tree surcharge
C_s'	=	soil cohesion	ϕ'	=	internal angle of friction
γ	=	moist soil unit weight	γ_{sat}	=	saturated soil unit weight
γ_w	=	water unit weight			

3.3.1. Combining and Summarising

Fuzzy set membership values for soil and forestry classes can be easily combined with the fuzzy 'AND' of the JMF function. The cardinal DEM and soil attribute error data can be propagated through the equation using Monte Carlo methods. Combining the two requires a multiple stage simulation procedure, resulting in a number of output maps representing degrees of certainty in each particular realisation. Rather than simply generating factor-of-safety *data*, these results can be utilised to present *information* relevant to the particular application of the model.

Each cell in the map is a member of all soil classes and all forest classes, with varying degrees of membership—some close to zero. Each realisation requires a different set of soil parameters to be applied in the FS equation (3.1). The Monte Carlo procedure must, therefore, be repeated for every possible combination of forest and soil class. The following algorithm summarises this procedure:

1) For each soil type; for each forest cover type;
2) Generate an 'equally likely' DEM based on error estimates (as discussed above);
3) Derive a slope map from the DEM;
4) For every cell, randomise all the derived variables based on the current soil and forest cover types;
5) Apply the factor-of-safety formula to every cell;
6) Repeat #2-5 *M* times; and
7) Compute summary statistics for the *M* maps.

The number of Monte Carlo runs (*M*) required to properly represent the distribution of the uncertainty is a subject of debate in the literature. In this case, a significance test does not really apply as there is no formal experimental design. The only value to test against is the Boolean result, which, technically, should be the mean of the resulting Monte Carlo frequency distribution. However, as proposed by Openshaw (1989), there is no reason why such a statistic might not be used as a guide rather than a precise test. Hope (1968) showed that only 19 realisations were required to yield statistically useful results. Openshaw refers to an *M* of 20-30 if only summary statistics are required.

For the purpose of the following case study, a conservative value of *M*=50 was chosen. To test the consistency of the Monte Carlo algorithm at this *M*, the 50 run algorithm was repeated 20 times for a limited subset of test data and the results plotted (see Davis 1994). The curves were consistent, and the Boolean value never approached the tails, indicating that *M*=50 was sufficient for the operation. Too small an *M* value would be indicated by exceedingly random lines, while identical curves would indicate an unnecessarily large *M* value. This visual method was suggested by Openshaw (1989).

3.4. Case Study

In order to demonstrate these techniques a case study was implemented. As a test-of-concept, the purpose of this case study was not to verify the actual numbers involved, but to demonstrate how these uncertainty and error management techniques could be used to extract additional information from existing data and knowledge. The case study served to illustrate how slope stability data can be turned into information useful in a decision support context, but did not actually involve any additional interpretation of the output for decision support.

An 8500 hectare study site was selected. The area is located on Louise Island, on the east side of Moresby Island in the Queen Charlotte Group, British Columbia, Canada, at 53°N, 132°E (see Figure 3.6). The area is a forest company test site, and was selected for: 1) the availability of data, and 2) the availability of experts with experience in the region.

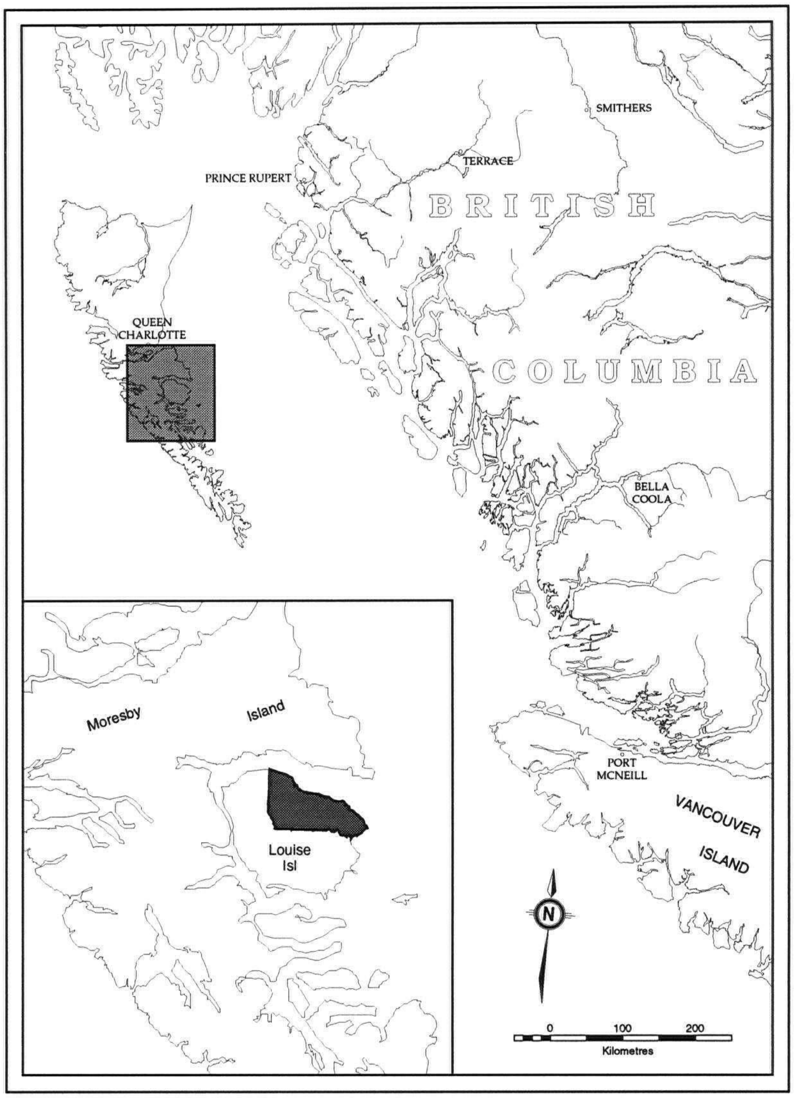


Figure 3.6. Location of the Louise Island study site

3.5. The Boundary Model and Attribute Uncertainty

Soil data were imported from existing maps digitised during previous Boolean slope stability studies. Though a factor, positional uncertainty of linework was minimised and then ignored in order to simplify this test of concept. Ten classes of soil were defined. Forestry data were gathered from digital forest cover maps and reduced to three classes: cut within 3-10 years, forested and other. Maps were rasterised at a 25m resolution.

Several soil and forestry experts were consulted regarding SI data for the transition corridor model. A sample of the resulting values for soil types is presented in Table 3.2. The values shown represent the estimated likelihood of misclassification based on surface indicators. The trace represents the maximum certainty values. A similar table was generated for class-to-class intrusion (overlap). Tables for the considerably simpler forest classes were also generated.

The transition corridor algorithm detailed above was then implemented for each soil and forest class, resulting in a fuzzy surface similar to Figure 3.5 for each. Although many small polygons appear in the foreground, note the 'plateaus' apparent on the larger structures, indicating central areas of polygons not affected by the boundary model. In contrast, the equivalent Boolean map (Figure 3.4) shows nothing but plateaus and cliff-like transitions. Cross-sections of two different boundary types are presented in Figure 3.7.

3.6. The Monte Carlo Procedure

The 'non-spatial' items required for Equation 3.1 were gathered from an extensive literature review undertaken by the US Forest Service Intermountain Research Station (Hammond *et al.* 1992) while developing their slope stability modelling system. Soil types found in the Louise Island study site could be successfully matched with the classification system used by the USFS, and means and standard deviations of the relevant data were calculated. The USFS study found that, for the most part, the values for each variable were normally distributed, with the exceptions of soil cohesion and root cohesion which were log-normal. The values are presented in Table 3.2.

The Monte Carlo simulation process must be repeated for every possible combination of soil and forest cover. In the limited 'test-of-concept' there were only 10 soil and 3 forest types, requiring 30 different simulation runs. However, more complex models could require significantly more simulations. In such a case, it would be prudent to determine in advance what classified data value

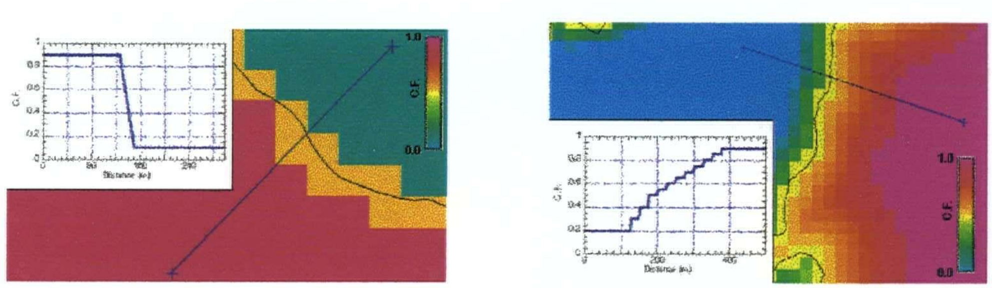


Figure 3.7. Detail of transition corridors between (a) bedrock and soil type 1; and (b) two similar soil types (1 and 3).

Class	Desc.	UCS	Dry Weight				Cohesion				Ang. Friction				Soil Thickness	
			gm/cc		mean	sd	kg/m2		mean	sd	degrees		mean	sd	mean	sd
			low	hi			lo	hi			lo	hi				
1	Sandy/Silty Morainal Blanket	SM	1.121	2.051	1.586	0.155	0	3197.7	1598.9	533.0	32	46	39	2.33	1.5	0.1
2	Silty Morainal Blanket/Veneer	ML	0.961	1.922	1.442	0.160	1221	1855.2	1537.8	105.8	23	43	33	3.33	1.5	0.1
3	Rubbly, Silty Colluvium/Morainal	ML-MH	1.378	1.474	1.426	0.016	390.6	1464.6	927.6	179.0	29	32	30.5	0.50	1.5	0.1
4	Gravelly, Silty, Fluvial	GM	1.762	2.083	1.922	0.053	488.2	2099.3	1293.7	268.5	33	43	38	1.67	1.5	0.1
5	Rubbly Colluvial	GW	1.570	2.051	1.810	0.080	0	0.0	0.0	0.0	28	39	33.5	1.83	1.5	0.1
6	Bedrock	-			0.000	0.000			0.0	0.0	88	89	88.5	0.17	0	0
7	Gravelly Silty Fluvial	GM	1.762	2.083	1.922	0.053	488.2	2099.3	1293.7	268.5	33	43	38	1.67	1.5	0.1
8	Gravelly Silty Fluvial	GM	1.762	2.083	1.922	0.053	488.2	2099.3	1293.7	268.5	33	43	38	1.67	1.5	0.1
9	Silty Fluvial	ML	0.993	1.089	1.041	0.016	0	0.0	0.0	0.0	22	30	26	1.33	1.5	0.1
10	Silty Morainal w. Bedrock	MH	1.121	1.442	1.282	0.053	390.6	1464.6	927.6	179.0	27	47	37	3.33	1	0.3

Table 3.2. Soil characteristics and estimated standard deviations (Primary source: Hammond *et al.* 1992).

combinations are incompatible or very unlikely (e.g., bedrock and mature forest) and eliminate these from the procedure.

Elevation data consisting of a semi-regular grid of elevation spot heights (British Columbia TRIM data; SRMB 1990) produced from stereo-photogrammetry were utilised. The average spacing between data points is 28m, therefore a raster grid spacing of 25m was chosen to minimise the interpolation required.

The stated error parameters for spot height data are (SRMB 1990):

- 90% of all determinate DEM points vertically accurate within ±5 metres.
- 90% of all indeterminate points vertically accurate ± 20m 90% of the time.

The vertical accuracy of the primary elevation points was calculated as follows: assuming that the error at each point is normally distributed (an assumption supported by Fisher 1989, 1991b), 90% of the area under a normal curve is contained within ±1.28 standard deviations of the mean. As this 1.28 refers to a 5m elevation difference regarding determinate points, one standard deviation of a specific point is calculated as 5/1.28, or 3.91m. One standard deviation for indeterminate points can be calculated as SD = 20/1.28, or 15.62m.

Block kriging using Surfer 4.0 was used to interpolate from these points to a regular grid. Kriging was chosen as the interpolation method for two reasons: its high accuracy and an ability to produce variance maps of the derived values. The published error values were combined with the variance for each data point to produce a final variance value for each interpolated cell.

The variance values were utilised in running the 'equally likely' DEM algorithm; in most cases, three to five smoothing iterations were required to bring Moran's Index within 0.001 of the original. Slopes were then derived from the DEM.

3.7. Results

The entire run of 1500 Monte Carlo simulations (50 runs x 30 type combinations) at a 25m resolution cannot be feasibly stored on most contemporary GIS platforms. A series of initial test-runs of the slope stability model were performed on a representative section of the data in order to determine the shape of the output curve. When fitted to a curve, the results indicate that a normal function would suffice to properly describe the resulting realisations of the slope stability values. Three maps were generated for each type combination: certainty factor, factor-of-safety mean and factor-of-safety standard deviation.

For example as illustrated in Figure 3.8, a particular cell might be assigned the following values:

1. For realisation Soil = 3 and Forest = 2, CF = 1.2, FS = 2.1, SD = 1.4

2. For realisation Soil = 4 and Forest = 2, CF = 6.5, FS = 6.2, SD = 1.5

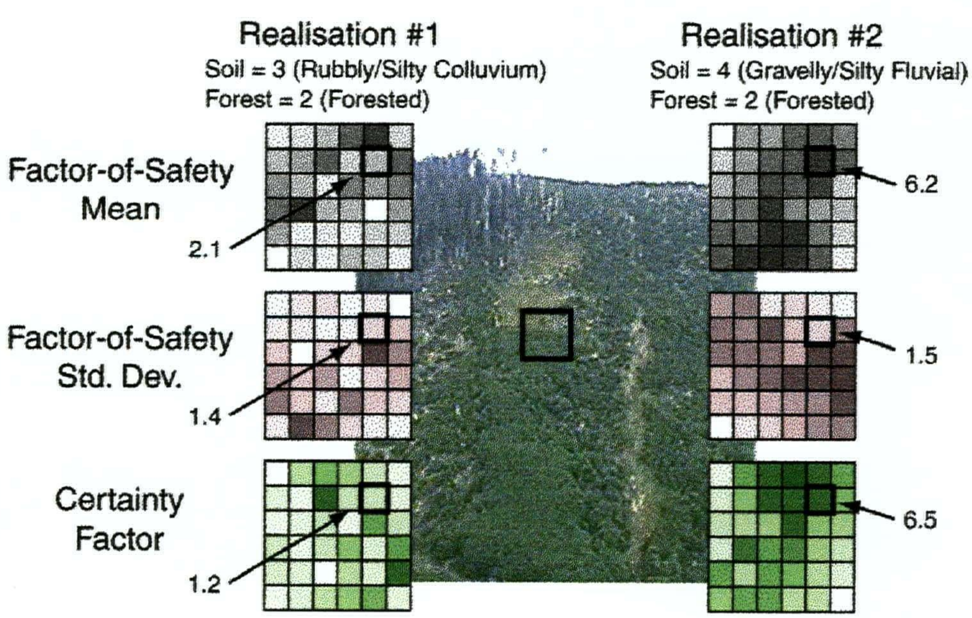


Figure 3.8. An illustration of the three types of surfaces resulting from the uncertainty modelling routine. Two of the many realisations are pictured. (Note that these are typical planimetric grids - the image is for illustrative purposes only)

This means that, in the first case, the likelihood that the soil type is 3 and the forest type is 2 is rather low (CF=1.2); if this actually is the case then the factor-of-safety is relatively low (2.1 – indicating a high likelihood of failure), and there is a reasonably high certainty in this factor-of-safety prediction (standard deviation = 1.4).

The second realisation, where the soil is Type 4 and the forest is Type 2, is much more likely. In this case, the factor-of-safety is relatively high (i.e., safe), and the standard deviation indicates a relatively high certainty in this FS value. It is necessary to examine all realisations to determine the most likely combinations, and it is important to note that certainty factors and standard deviations are relative to the entire study area.

The standard slope stability results (i.e., Boolean results) may be extracted using a maximum likelihood filter, in which only the highest certainty factor for each cell, and its associated values, are retained. Essentially, this maximum likelihood summary tosses out all the additional information generated by the new procedures discussed above, and returns to the Boolean representation. The maximum likelihood map for the entire region is displayed in Figure 3.9.

By incorporating the standard deviation information into the analysis (Figure 3.10), it is possible to use the maximum likelihood information in different ways. For example, as demonstrated in Figure 3.11, areas in which slope instability are highly possible (low standard deviation and low factor-of-safety) are highlighted.

The uncertainty model's real utility is in its retention of information about realisations that do not quite 'make the grade' in the maximum likelihood filter. For instance, using the example values presented above, if the CF for realisation one was 6.2 rather than 1.2, the low factor-of-safety associated with this very likely realisation would be important. If this realisation should represent reality, then the cell is particularly unstable and perhaps should be avoided for road construction or harvesting. Realisation two would give the opposite results. This type of 'less-than-maximum likelihood' analysis is illustrated in Figure 3.12. Here, the lowest factor-of-safety with a *reasonable* likelihood is retained, rather than *maximum* likelihood. This type of data summary might be termed a 'worst case analysis', and would be useful when potential danger is the issue.

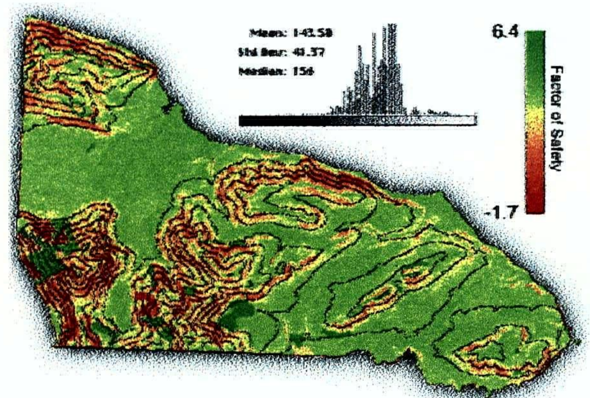


Figure 3.9. Maximum likelihood summary of slope stability factor-of-safety.

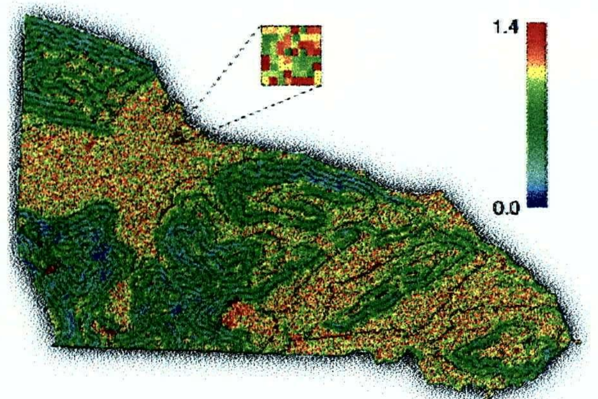


Figure 3.10. The spatial distribution of standard deviation of slope stability factor-of-safety

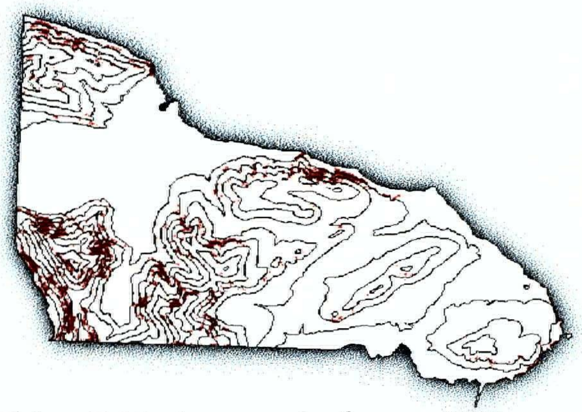


Figure 3.11. An example of an application-specific data summary: highlighting the areas in which slope instability is highly possible.

Figure 3.12. The worst-case-scenario summary, in which the 'most dangerous' realisation that has a reasonable likelihood of occurring is utilised (on a cell-by-cell basis).

These multiple surfaces can be used in many other ways. Although the research in this dissertation stops short of incorporating the information into management schemes, some of the relevant issues are introduced in Chapter Six.

3.7.1. Problems and Work Required

3.7.1.1. VERIFYING PARAMETERS

A typical modelling procedure results in an unequivocal answer. It may be right, it may be wrong, or it may be somewhere in between, but the interpretation of the values is normally straightforward. In the uncertainty model, however, equivocal values are more the norm. In this case study, soil classes, forest classes, and model results all have certainty factors assigned to them during the procedure. Standard deviations were also assigned to each cell of the results. What do these values actually mean on the ground?

The field of uncertainty modelling clearly lacks procedures for verifying or comparing the results of the models developed. A straightforward Boolean soil classification can be verified with a sampling scheme, and perhaps summarised in a classification error matrix. If errors are too high, then the classification scheme may need to be modified. However, there are no existing procedures to verify, for example, a certainty factor of 6.5 in a particular soil class. Procedures are needed to demonstrate the utility of uncertainty modelling in a quantitative format, rather than simply relying on generic statements such as "a truer representation of reality" (Ramlal 1996) that have little value in practise. Given such procedures, it would then be possible to re-tune model parameters to better represent the uncertainty in a particular area, or to modify the model's inputs derived from semantic import for the same purpose. In the following chapter such procedures are developed and tested through field verification of the uncertainty model.

3.7.1.2. PREDICTION

A second problem with this particular modelling procedure is the difficulty in verifying its output (as opposed to verifying input as discussed above). Slope stability models are inherently difficult to verify unless considerably detailed data are available. A single sample (e.g., an image) of an area gives only one temporal slice. Mass movement takes place over time; areas grow back and obscure the evidence of past mass movement; and some mass movement is delayed for many years beyond the peak initiation time. Both temporal and spatial detail are required to properly evaluate slope stability.

In a similar manner to the previous problem, there are no standard procedures for verifying certainty factors in multiple realisations resulting from the uncertainty model. We need to determine if this model predicted mass movement correctly, but of greater importance is: did it predict mass movement *uncertainty* correctly? Procedures are required for this problem as well. These procedures are developed and discussed in Chapter Five. There, the model developed in this chapter and a second model (based on new parameters) developed in the following chapter are applied to a separate test area. The model results are validated using a high-resolution temporal model of mass movement.

3.7.1.3. REPORTING AND COMMUNICATING UNCERTAINTY

The Louise Island test case provided a platform for demonstrating the complexities of uncertainty propagation, and for developing one particular approach to handling this complexity. One underlying purpose of this procedure was to maintain a maximum amount of information through to the final results of the environmental model; whatever it may be. As one author puts it, "it is a mistake to round inventory data or classify it [prior to] final presentation" (Iles 1994:12).

This multiplicity of results should then lead to a further stage in model processing: summarising the results for the particular application at hand. For example, when slope stability data are used in a harvesting profitability model, the key issue is: what areas are too steep/unstable to cut? When road building is the issue, the question becomes: what areas have a medium to high probability of catastrophic failure? The data required of the slope stability model would be somewhat different in each of these two cases. Summarising procedures will use different parameters in each situation.

It is particularly difficult to understand spatially variable, multidimensional model results using simple summary statistics. Although maps are an improvement, they too are inadequate to the task; particularly if the target audience is not familiar with the underlying science. Decision support models making use of this information have as basic criteria: clear, concise, understandable summaries of many types of data. If uncertainty models are to fit into this framework then it will be necessary to place considerable emphasis on communication of the model results. Examples of simple types of communication were offered above. A more extensive discussion of communicating uncertainty in this particular project can be found in Davis and Keller (1997b); however, as discussed in that paper, a considerable research program would be required to implement uncertainty communication into real world management. The research discussed in this dissertation is but one step in that direction.

3.8. SUMMARY

In this chapter a previously developed uncertainty model has been described, and its application to slope stability modelling on Louise Island, B.C. has been presented. Procedures were developed to address the conversion from a Boolean-polygon to fuzzy-continuous data model, and to apply this fuzzy model to a typical process modelling procedure. At this stage, the results are of limited utility due to a lack of procedures for visualising and therefore enhancing understanding of the data. Its spatial constraints are primarily based on information gathered from semantic import (SI), so there is also a lack of proof that the model's inputs actually describe uncertainty correctly. Furthermore, at this stage there are no data to determine if the prediction of uncertainty in the model outputs is actually correct; only with extensive landslide data could this be addressed. The following chapter focuses on the second of these problems: evaluating the parameters of the uncertainty model obtained through semantic import and other secondary methods.

Chapter Four

Verification of Model Inputs

4.1. Introduction

The uncertainty model described in Chapter Three uses parameters obtained from both numerical analysis of resource attributes and semantic import of expert opinion. The metadata available for the modelling procedures described are somewhat typical of resource modelling in general: metadata are not gathered during resource surveys and so must be extracted through analysis, or estimated from other sources (e.g., Burrough 1989; Livingstone and Raper 1994). Only in highly controlled studies is a wealth of metadata likely to be available.

In typical resource modelling, data derived from secondary sources are usually subjected to some type of procedure to confirm their utility in the current modelling scenario. For example, in standard slope stability modelling, the slope values derived from photogrammetry might be spot checked in the field to determine their accuracy. However, one of the principal problems with studies of *uncertainty* in resource models is the difficulty in obtaining these confirmatory data. In essence, uncertainty modelling uses additional data (either retained or from new sources) to increase knowledge about the potential variability of databases, modelling procedures or decision models. However, the use of an inappropriate uncertainty model or propagation procedure can lead to under- or over-estimating this variability. Such mistakes can potentially be as significant as ignoring uncertainty altogether. Even if the appropriate procedures have been used and the estimates of variability make sense, there is no easy way of confirming this fact.

There are two principal areas where confirmatory procedures are currently lacking: a) fuzzy numbers, and b) classification uncertainty.

Fuzzy Numbers: There is some difficulty in understanding and utilising fuzzy numbers outside of database manipulations. For example, what does a 0.7 certainty factor for a sandy/silty morainal soil actually look like? In theory, it refers to a sample that is 'somewhat like' the ideal class. Membership values make sense in manipulating data; however, confirming this number with a sample is more difficult. There are no established procedures to compare such a fuzzy classification with a confirmatory sample.

Classification Uncertainty: Fuzzy classes and fuzzy classification methods are utilised in many resource management disciplines. However, the majority of the effort in resource management uncertainty analysis involves either creating fuzzy classifications from a series of samples (e.g., remotely sensed images), or assigning fuzzy class memberships to samples based on a training dataset. When faced with the typical situation of *existing* definitions of class structure, and *existing* polygon-based resource databases, fuzzy class membership routines do not necessarily make sense. The data used to establish these classes are not available at that point. In comparing samples with classes one is faced with a 'black box', where attribute *I* of sample *A* falls between parameters *b* and *c*, and so sample *A* belongs to class *X* (and only class *X*). There is often insufficient information to support notions of class 'purity' required in fuzzy classification.

In addition to these two problems, there is also the issue of tuning an uncertainty model. In standard discrete natural resource models, confirmatory sampling might be used on a random (or systematic) basis to determine if polygons were classified correctly. Parameters could then be redefined to fit reality. In a standard distributed model, one might perform transect samples to determine if the spatial structure of the attribute(s) being modelled are accurate. For example, a transect between two forest stands could establish if the spatial distribution of species between the stands matches the model (e.g., gradual change or abrupt change).

However, in a distributed uncertainty model this transect involves a series of changing membership values. If we focus on a complex system such as soil class—which is based on multiple attributes—the model might indicate that "a sample at point (x,y) should belong to class 'gravelly, silty fluvial' with CF (certainty factor, also known as fuzzy membership value) of 0.7, and class 'sandy morainal' with CF of 0.5". There are no established methods for sampling such a transect and then comparing it with these modelled values. Fuzzy classification systems provide a starting point; however, their assumptions do not necessarily apply.

This chapter addresses this issue of sampling in order to calibrate an uncertainty model. The methods explored are extensions to existing fuzzy classification techniques—adapted and expanded to address confirmatory sampling. Existing techniques are reviewed, new extensions are developed, and the implications are addressed for uncertainty modelling in general. A subset of the techniques are then applied to the model developed in Chapter Three. Samples taken within the study area are used to re-calibrate the most crucial parameters of the uncertainty model. The differences between the new data models and the originals are then discussed (implementation of the data models in the *process* model and comparisons with the original will occur in the following chapter). The sampling and allocation issues discussed herein have considerable relevance to uncertainty models in general, particularly those utilising expert opinion as input.

4.2. Background

The principal questions addressed in this chapter are: 1) how can fuzzy classification structures be compared with confirmation samples?; 2) how well did expert opinion function as an input to generate the distribution of uncertainty represented by the fuzzy structures? (the transition corridor model); 3) how well does metadata gathered from published statistics represent the actual uncertainty on the ground? (focusing on major model inputs); and 4) how can these confirmation data be used to recalibrate the model?

Although most of the physical effort involved in answering these questions is concentrated on numbers two and three, it is the first question that consumes most of this chapter. Fuzzy classes represent a unique and often highly appropriate way of looking at the world. However, when the focus is on specific attributes, such classes are also a considerable abstraction. Given this focus on fuzzy classification, the first point of business is to delve into the topic in greater detail than provided in the introduction to fuzzy sets in Chapter Two.

4.2.1. Fuzzy Classification

The model utilised in this work is based in part on the theory of fuzzy sets (see §2.3.1.2.4. for background details). Four main application areas of fuzzy sets have appeared in resource analysis. These are:

1. **Fuzzy rules**: Rather than encoding the steps in decision making as a series of IF-THEN statements, fuzzy rules are a set of parameters that are applied all at once, and the decision is made through a weighting system. This more closely emulates the human decision process, and is the most common application of fuzzy set mathematics (e.g., Bouille 1992). Fuzzy rule applications are numerous. For example, the Tokyo subway system uses a braking system based on fuzzy rules. The speed of braking is determined through a fuzzy decision-making process based on simultaneous evaluation of dozens of separate inputs (speed, weight, weather, etc.). In cases such as this, the fuzzy system has been found to provide smoother, more efficient operation than standard computer-assisted hardware. The fuzzy decision-making process can be encoded in hardware, speeding the process by exponential factors.
2. **Fuzzy class definitions**: Fuzzy classifications allow a blurring between standard classes by defining a class boundary as a function, rather than the hard boundaries of an IF-THEN statement (e.g., Burrough 1989). A sample might belong to two (or more) classes to varying degrees. Figure 2.4 demonstrates how the edges of 'hard' classes are blurred by a fuzzy classification system.
3. **Fuzzy queries**: Standard spatial database queries involve hard numbers. For example, determining the suitability of an area for agriculture requires queries such as 'IF RAINFALL < 200mm AND DRAINAGE = GOOD THEN...'. Fuzzy queries apply fuzzy set theory to the analysis of spatial data, allowing the fuzzy semantics of a query such as 'What areas are

NEAR the river, NORTH of the town and SUITABLE for agriculture?' When humans ask such a question, they are actually setting a series of fuzzy constraints on the query. For example, the word 'near' has different meanings depending upon the scale and the purpose of the analysis. We do not mentally picture a sharp cut-off when using this word. Fuzzy queries serve to translate this type of meaning into an actual spatial data query.

4. **Fuzzy classification systems**: A variety of methods have been developed to segment complex environmental sample sets into classes. Fuzzy set theory has been applied through algorithms such as fuzzy-*c*-means (Bezdek *et al.* 1984), in which classes and sample memberships in classes are determined through iterative minimisation of a fuzzy function. This system has proven useful in several areas, including remote sensing (Du and Lee 1996; Foody 1996) and soil classification (McBratney and DeGruijter 1992).

In the work discussed in this chapter the focus is on using the latter item—fuzzy classification—as well as the second item—fuzzy classes—to verify the uncertainty inherent in the major inputs to the slope stability model. In essence, it involves defining how far a particular sample is away from its modelled class in fuzzy attribute space. In this section, soils are used as the principal example as forestry data are represented by a much simpler classification system and the other major inputs to the slope stability model are based on cardinal data.

The model discussed in the previous chapter uses fuzzy values to define to what degree a particular point on the ground (actually, a cell in the raster structure) belongs to each of the soil classes. In this case, the fuzzy value refers to how much we expect a ground sample at that site would be *like* each ideal class. For example, a value of 0.8 for class 1 indicates high similarity, while (in the same cell) a value of 0.2 for class 4 indicates low similarity. We would expect that an average sample taken in the cell would be similar to the ideal definition of class 1, and dissimilar to the ideal definition of class 4.

However, the key question is, how is it possible to make this comparison? In a normal confirmation sampling situation one would gather and analyse samples, classify them, and then compare modelled class with sampled class. If the comparison did not fall within a classes parameters, then the cell is deemed misclassified.

In a more complex situation the classes might be defined using a fuzzy system (we are now referring to 'fuzzy class definitions'—number two in the above list). In this case, small differences such as the sand content falling marginally outside the class bounds would not disqualify the sample. The comparison between the sample and the class would not be a yes or no, but a fuzzy number—a *degree* of belonging. However, the definition of a class typically involves a number of different attributes. This comparison must therefore summarise how the sample and the class compare relative to *all* of these class components. If the comparison is made using a graphical method, the graph must have as many axes as there are attributes. This is termed *p*-dimensional attribute space, where *p* is the number of attributes used to define the class. Figure 4.1 illustrates a simplified view of this space, showing just two attribute axes, three classes and two samples. Note how the fuzzy class definitions blur the class boundaries.

Essentially, the uncertainty model discussed in the previous chapter has generated a prediction for this fuzzy number. The purpose of the verification is to see if the number corresponds with reality.

Soil systems are particularly suited to fuzzy classification. As Fridland (1974) notes (quoted by Odeh *et al.* 1992:506): "in terms of classification, the soil cover is liable to be either continuous (with gradual transitions between soils, though closely related soil forms) or discrete (with sharp transitions between soils and very dissimilar neighbouring soils)." This complexity is apparent at all scales of soil analysis (Webster 1985). Soil science was one of the first natural resource-based applications of fuzzy classification (e.g., Burrough 1989), and this discipline continues to be a favourite area of application for these techniques.

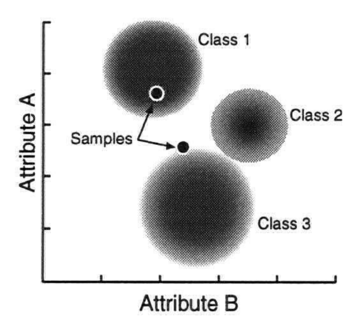


Figure 4.1. Simplified (p=2) view of *p*-dimensional attribute space, fuzzy classes and samples.

In the sections that follow, two terms will be used to address uncertainty verification: *classification* and *allocation*. 'Classification' refers to building a new set of classes based on detailed data (e.g., a series of samples using cardinal values,) while 'allocation' refers to fitting a new sample (with its cardinal values) into previously defined classes. The term 'allocation' is the more correct of the two in this context, as classes will have already been defined when verification commences. However, the methods available for allocation are all drawn or extended from classification procedures; therefore, much of the preliminary discussion below will focus on this latter term.

There are a number of possible ways of setting up classes or deciding to which class a new sample belongs. Most of these methods assign one class (and only one class) to a particular sample. There are, however, several techniques that allow multiple class memberships. The following sections will introduce or elaborate on classification methods that, in the process of setting up classes, utilise some type of multiple class technique that is potentially useful in the process of allocation.

4.2.2. Maximum Likelihood

The maximum likelihood (ML) classification algorithm is a decision rule that assigns a set of measurements to a class based on probability. It is commonly used as a supervised classification procedure in remote sensing, yet it is equally applicable to assigning a sample to a class in attribute space. The ML rule is normally used to assign a pixel/sample *X* to a single class; however, if the decision stage is removed, a series of probabilities can be assigned to *X* indicating probability of membership in every class.

The ML classification algorithm is as follows:

Decide that X is in class c if, and only if,

$$p_c \geq p_i \text{, where } i = 1,2,3\ldots,m \text{ possible classes} \tag{4.1}$$

and

$$p_c = \left[-0.5 \log_e \left(\det(V_c)\right)\right] - \left[0.5(X - M_c)^T \left(V_c^{-1}\right)(X - M_c)\right] \tag{4.2}$$

where M_c is the mean measurement vector (i.e., the set of measurements in attribute space to class centroids), V_c is the covariance matrix of class *c*, and $\det(V_c)$ is the determinant of the covariance matrix (Odeh *et al.* 1992). To retain probabilities for all classes (rather than create hard boundaries) the first decision rule is removed.

In remote sensing classification the primary problem with this routine is the assumption that each class has an equal probability of occurring in the terrain. If this is not the case then the decision rule can be changed by weighting each class by it's *a priori* probability.

There are two primary problems with this algorithm when applied to the allocation of new individuals to existing classes. First, the ML routine assumes a Gaussian distribution of all statistics. This assumption is sensible for remote sensing applications, since in supervised classification the training sites may be chosen with this restriction in mind. However, there is no indication that all attributes of soil or forestry classes are distributed this evenly. Indeed, there is some evidence otherwise (Odeh *et al.* 1992; discussed in the following section). Secondly, the assumptions of 'probability' are not the same as fuzzy 'possibility'. Though similar in notation, the two are considerably different in application. The former deals with yes/no answers, while the latter deals in similarity. For a full discussion of the difference the reader is referred to Yoshinari *et al.* (1993) or Bezdek (1992).

4.2.3. Continuous Classes - Fuzzy Clustering

Starting with the work by Ruspini (1969), Dunn (1974) and Bezdek (1974), several methods for constructing continuous classification systems have been developed, where the reduction to a single class membership per sample does not (necessarily) occur. Collectively, these methods are referred to as fuzzy clustering. The most popular and well studied method is known as the fuzzy-*c*-means or fuzzy-*k*-means algorithm (the two are identical but use different notation). Fuzzy-*c*-means is a direct generalisation of hard-*k*-means (Hartigan 1975). This method is based on minimising the within-class sum-of-square error function. The details of both methods are presented in Appendix A.

4.2.3.1. BACKGROUND - FUZZY CLUSTERING

Fuzzy clustering was developed for geo-statistics and soil science due to problems encountered in restricting class boundaries to regions (in attribute space) with a small probability density. The gradual changes found in reality were poorly represented by hard classes. For example, individuals could be very close to each other in all attribute values but be split into different classes due to hard-and-fast rules. Figure 4.2 details some of the possible distributions of soil samples (individuals) in simple 2-D attribute space; hard boundaries are only useful in some of the cases.

Although ordination methods (e.g., principle components analysis or multidimensional scaling) offer ways to represent data using a continuous model and a simplified dimensional space, these methods are less than ideal for non-linear class structures such as those found in soil analysis (McBratney and DeGruijter 1992). Fuzzy clustering using continuous classes is better suited to non-linearity.

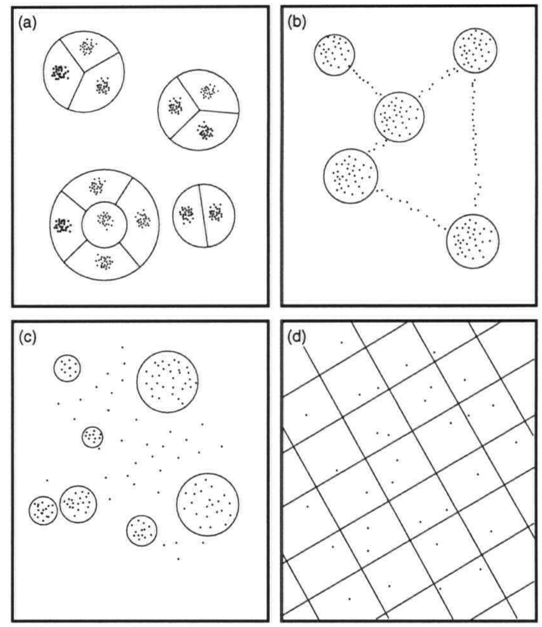


Figure 4.2. Notional distribution of individuals in attribute space: (a) hierarchical, (b) clusters with directed lines, (c) weakly clustered, (d) equal density (Source: McBratney and DeGruijter 1992).

Continuous classes provide better representation of individuals located interstitially between classes (intergrades) than do standard discontinuous classes. Instead of trying to expand the nearest class to include them (or simply calling them 'exclusions'), intergrades are given partial membership in all nearby classes based on the distance in attribute space to each class centroid (Figure 4.3). The key difference between partial (or fuzzy) memberships and other types is this *series* of memberships. A particular sample is not locked out of all other classes once its maximal membership value is determined. As with other fuzzy values, a particular class membership indicates to what degree the sample is 'like' the idealised or chosen class, rather than a probability of membership.

The concept of using a distance metric to allocate an individual sample in attribute space is demonstrated in Figure 4.4. The spheres around the classes represent the boundaries used in 'hard'

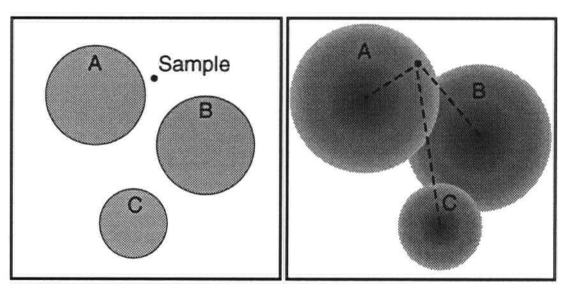


Figure 4.3. Hard classes and continuous classes. In (a), the interstitial sample belongs to no class, while in (b), its membership is defined based on centroid distances.

classification to exclude all others. In fuzzy classification, such boundaries become gradients.

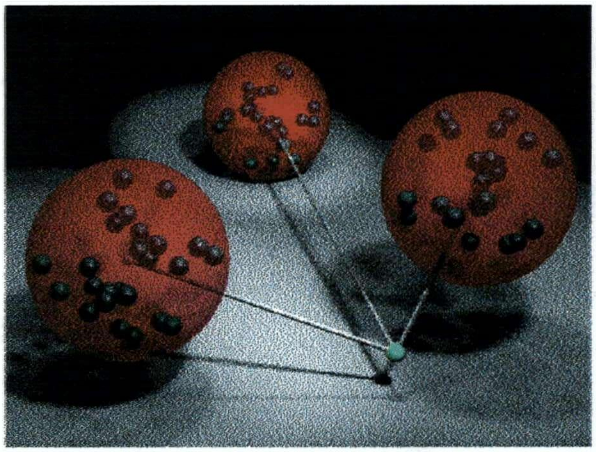


Figure 4.4. Mahalanobis distance in a 3-D attribute space. The spheres indicate Boolean class divisions that are discounted in fuzzy clustering.

The classification algorithms (such as those discussed in Appendix A) do not define any specific method of calculating the distance to the class centroids. This distance can be calculated using several possible metrics. One option is the Euclidean norm which gives equal weight to all axes and ignores any dependencies among them. For example, in a two attribute soil class, the distance from a new sample to the class centroid would be measured using a simple Pythagorean equation, as demonstrated in Figure 4.3. Additional attributes would simply increase the number of dimensions in the calculation.

However, this simple concept of a circular (2-D), spherical (3-D) or hyper-spherical (>3-D) class would require that all attribute variables be linear. It is possible to normalise the various axes in order to approach the spherical class shape, but non-linearity will distort the imaginary spherical class. In fact, studies have shown that such regularly shaped classes are a rare occurrence when modelling soil systems (Odeh *et al.* 1992). Non-linearity in variables will requiring a class shape that might be termed a 'hyper-polygon'. In Figure 4.5 the concept of a hyper-polygon class is demonstrated in three dimensions.

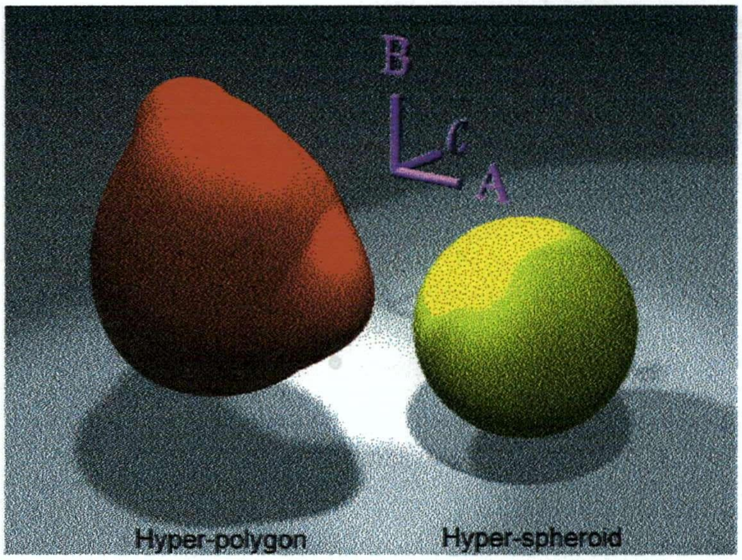


Figure 4.5. Classes viewed as structures in (A,B,C) attribute space. Classes are not necessarily spherical groupings (e.g., yellow class); they may be represented by more complex objects, here termed 'hyper-polygons' (e.g., red class).

The other problem with the Euclidean norm is that, in practise, many of the axes will be dependent on one another to some degree. Therefore, in most cases a more appropriate distance metric is the Mahalanobis norm. This metric is utilised in many natural resource studies (e.g., Abel *et al.* 1992; Leese and Main 1994). It is capable of compensating for a non-spherical shape of the class in attribute space and, additionally, can account for dependencies in the variables. It makes use of the pooled within-classes variance-covariance matrix.

$$d_{ij}^2 = (x_i - c_j)\sum^{-1}(x_i - c_j) \tag{4.3}$$

where x_i is the vector of attributes, c_j is the vector of centroids, and $\sum^{-1}$ is the variance-covariance matrix (Odeh *et al.* 1992).

4.2.3.2. APPLYING FUZZY CLUSTERING TO CONFIRMATORY SAMPLING

Functionally, the work discussed in this chapter differs from the fuzzy clustering derived from fuzzy-*c*-means algorithms. We are not interested in defining classes; these were defined during the original soil survey. The focus is instead on allocating new samples (individuals in attribute space) to existing classes. The methods will draw on the algorithms described above; however, classification (in the sense of defining classes) is no longer the issue, so iterative procedures are not necessary.

There has been relatively little research performed on alternative methods of allocation. Some of the basic methods are reviewed by Sneath (1979), Payne and Preece (1980) and McBratney (1994). These authors note three primary methods of allocating new individuals to existing classes: diagnostic keys, diagnostic tables, and distance in attribute space.

Diagnostic Key: Keys force a user to make a sequence of tests, each having different possible outcomes. After a series of tests the unknown individual will be fitted into a known class. Keys make use of a tree structure of decision making, although the tree can have a variety of topologies. Keys are normally restricted to standard 'hard' classification systems such as a standard soil taxonomy.

Diagnostic Table: This is a two-way table used for identifying the class of an unknown individual. For example, the rows would represent the class and the columns the range of attribute values required. An unknown may belong to more than one class, or none. This system is often used in biology for nested classification systems (e.g., order, group, subgroup).

Distance in Attribute Space (taxonomic distance): These methods use some measure of the distance from the individual to a class centre in attribute space. Methods such as discriminant analysis and pattern recognition can be considered part of this group. For example, neural network-based methods (e.g., Skidmore *et al.* 1997), though they use a network structure rather than points in attribute space, allocate individuals based on example rather than a set of predetermined rules.

A fuzzy classification system such as fuzzy-*c*-means (and its extensions) utilises this type of distance measure as an integral part of the analysis. The Euclidean or Mahalanobis distance metrics, minus the iterative steps, can be used to determine the distance in attribute space.

4.2.3.3. STRUCTURE OF THE CLASSES IN ATTRIBUTE SPACE

The algorithms used in classification procedures group the samples (individuals) in a wide variety of ways. However, once classes are established (even as an intermediate step in an iterative procedure) the most common method of measuring how 'close' an individual is to a particular class is to use the class centroid as a target (as illustrated in Figure 4.4). This makes perfect sense when shuffling centroids in the iterative classification process, particularly when variances are equalised and classes are somewhat spherical. However, in the process of allocation—particularly regarding soil or other non-linear sets of attributes—the centroid may not be the best target for a quantified degree of membership.

In a simple example using spherical class structures, Figure 4.6a shows a new individual located in an intergrade position between homogeneous classes. Though it is located equidistant from each centroid (and therefore has equal memberships in both classes using Euclidean or Mahalanobis measures), it is clearly 'closer' to class A than class B. Whether this fact is of functional utility depends upon the application.

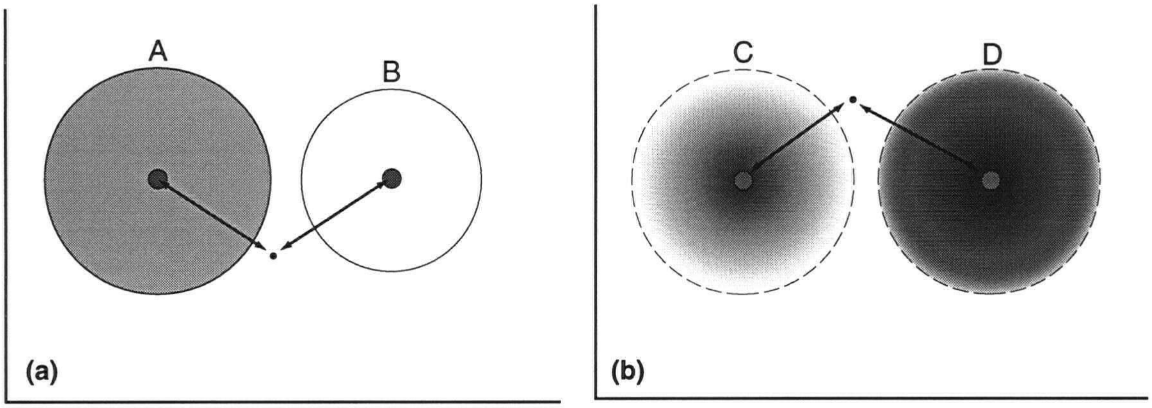


Figure 4.6. (a) A new individual at an intergrade position between class A and B is given equal membership if MD values are the same, despite the variations in class size. In (b), MD is smaller for class C, but the fuzzy structure of the classes indicate that D may be the better choice (the dotted line represents the location of class bounds *if* the fuzzy classes were to be 'hardened'.

In another example (Figure 4.6b) the intergrade individual is located between two classes with *fuzzy* attribute definitions. It is clearly closer to the centroid of class C; however, class C is defined internally with a gentle membership slope. It may be more appropriate to assign the individual to class D due to its higher internal 'density' (the dotted line indicates where the class bounds would be located if the fuzzy classes were 'hardened').

Another issue is the classification system itself. Using soils as an example, standard classification uses a general purpose taxonomy where soils are divided into classes based on many different attributes. When a classification subset is used for a specific purpose such as agriculture or slope stability modelling, only certain attributes may be of interest. The classes thus defined may have no 'pure' centre. There may be no ideal combination of attributes that define the perfect 'sandy, morainal blanket'; the definition is simply a range of values (with or without a fuzzy boundary). The 'pure' centroid as a target makes little sense in this situation.

Yet another problem with centroids is the non-linearity of some environmental classification systems (notably soils). In spatial data analysis the spatial centroid of a polygon is only an appropriate summary device if the polygon is regular. Various distortions in polygon shape can lead to a centroid that is highly inappropriate (Figure 4.7). A similar situation exists with hyper-polygons in attribute space. The class centroid may, therefore, be a poor measure of centrality.

What is clearly required is a more complex measure of 'belonging' to a class than a simple centroid. However, the sample that we are trying to allocate may also be represented by more than a simple point in attribute space. The nature of the sample will also determine what methods are required.

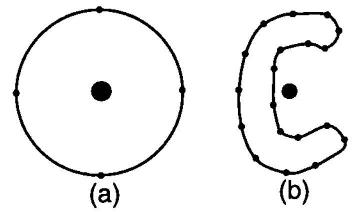


Figure 4.7. Centroids are not an ideal measure of polygon location when polygons are non-circular.

4.2.3.4. NATURE OF THE SAMPLE

To reiterate, simple measures that generate classification or allocation statistics for a sample in attribute space may be insufficient due to complicated class structure. A more complex method may also be required due to the nature of the sample itself. This individual (the sample) may not simply be a point in attribute space, but instead a region of higher geometry (i.e., a hyper-spheroid or polygon rather than a point). The nature of this region is determined by two major factors: 1) sample uncertainty; and 2) the nature of continuous models.

Sample Uncertainty: The single measurement of each attribute represented by a point sample would normally appear as a zero-dimensional point in attribute space. This is the standard way of dealing with samples in most classification schemes, including those that incorporate fuzzy clustering. However, the precision of the tools used, the possible errors in laboratory analysis, and the nature of the attribute being measured all contribute to sample uncertainty (see §2.2.3.1. and §2.2.3.3.). This uncertainty in essence blurs the sample point in attribute space. This will complicate any measure of attribute distance to a class, for the sample itself may be represented by a region or by fuzzy boundaries.

Continuous Model: The second problem is the resolution of a continuous model. A raster-based model of a continuous data layer (a common method of representation) has a specific cell size that defines the model's resolution. In a remote sensing application the reflectance of a pixel is an average of all occurrences within (and some from neighbouring pixels). In a raster data model the value of a cell is normally the 'typical', most common, or the effective contents of the cell (e.g., a raster model of transportation might assign a cell a value of 'road' even if only 10% of the cell contained a road). There are a number of other methods of determining how a cell should be coded, based on what is inside or what its' neighbours contain (for a full listing see Chrisman 1997). On the ground, a simple point sample within this cell would not be an appropriate method of testing a model; rather, a series of samples or a sample in a *typical* location within the cell would be more sensible. Such a sampling scheme would generate something more than a point in attribute space. In extreme cases, where variability is high or inclusions are common (e.g., soil inclusions, heterogeneous forests) a sample hyper-spheroid or polygon would be appropriate. In Figure 4.8 a sample hyper-spheroid with three classes in attribute space is demonstrated. Note how the centroid-to-centroid membership measures become increasingly inadequate as the complexity of the situation increases.

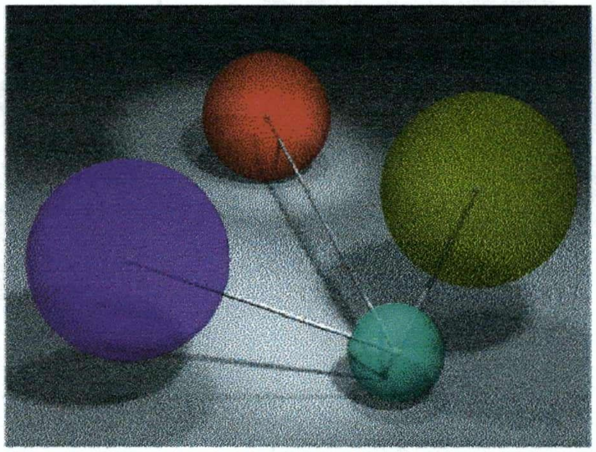


Figure 4.8. A sample hyper-polygon (lower right) and three classes in attribute space with Mahalanobis measures between their centroids.

4.2.3.5. METRICS AND MEASURES FOR MEMBERSHIP VALUES

If the data are available, the situation in attribute space can be modelled in quite complex ways. The metrics used to make measurements can be based on the Euclidean or Mahalanobis types discussed herein, or could also be based on numerous others generated in fields such as remote sensing, pedology, biology, and most other natural sciences (for e.g., Manhattan or Minkowski metrics). The measures and axes being measured may be complicated by transformations such as those used in principle components analysis. The classes could be defined by anything from a black-box with rigid boundaries to a hyper-polygon composed of a complex function. Samples could be points, blurred points, or equally complex hyper-polygons with discontinuous structures.

Therefore, even when using a relatively simple metric such as Mahalanobis, the measurements that are used to characterise class memberships may not be optimum if the standard centroid-to-centroid method is used. There are a number of alternatives; the choice of which would depend upon the amount and type of data available, the nature of the feature being modelled, and the purpose of the modelling. In the case of soil modelling these might include a) class boundary inclusion, b) alternative vector measures, and c) various combinations.

Inclusion of class boundaries: The boundary of a class in attribute space provides additional information regarding a sample's potential membership in that class. Depending on the class model utilised, the boundary may be either a rigid yes/no line or a function that tapers off. The rigid boundary provides an obvious measurement point; the tapering boundary has many.

In the case of classes that differ in size (on one or more attribute axes) the use of class boundary is a more accurate measure of relative membership than the centroid (Figure 4.6). When the class shape does not approximate a hyper-spheroid it may also be a more accurate measure.

Measurement to a hard boundary could be accomplished as follows:

Each class is represented by a membership function, which (for comparison) in hard partitioning is in the form of:

$$m_A(x) = 0 \text{ for } x < b_{min} \text{ or } x > b_{max} \tag{4.4a}$$

$$m_A(x) = 1 \text{ for } b_{min} <= x <= b_{max} \tag{4.4b}$$

where b is the set of values defining the class, and $m_A(x)$ is the 'grade of membership' of x in A (either a *yes* or *no* when using hard partitioning). For one attribute, the distance to the border is simply the minimum of $|x\text{-}b|$. If x is within the bounds of the class then the distance equals zero. The grade of membership ($m_A(x)$) will be defined as the inverse of this distance function. With multiple classes the Euclidean distance function becomes:

$$d_{xb}^2 = \sum_{v=1}^{p} \left[\min\left(x_{xv} - b_{bv}\right)\right]^2 \tag{4.5}$$

where p is the number of classes, x is the sample value and b is the boundary location for the class. Adding the covariance matrix to the equation would incorporate the Mahalanobis distance.

In the case of continuous or fuzzy classes, the class membership function is more complex than defined in Equation 4.4. A typical fuzzy class membership function might be defined as (Burrough 1989):

$$\mu_A(x)=\frac{1}{\left\{1+a(x-c)^2\right\}}\text{for } 0\leq x\leq P \tag{4.6}$$

where a is a parameter governing the shape of the function and c defines the value of the property at the centroid. In this case the distance function will be more complex, as a cutoff value is required to 'harden' the boundary. In any case, the resulting number will reflect the distance in attribute space from the class; however, it will not take into account the continuous nature of the class. For this, a more complex measure will be required.

Vector Measures: If the class, the sample or both are considered fuzzy sets, then metrics for fuzzy set distance are appropriate. Several authors have developed such metrics for the measurement of *physical* distance (Preparata and Shamos 1985; Altman 1994). The metric developed by Altman (1994) returns a fuzzy set as a measure of the distance between two fuzzy regions. These authors focus on raster datasets, and so are actually dealing with stepped, or discretised continuous values. The sets being compared do not have infinite memberships, as might be found in a truly continuous structure (e.g., object-oriented data structures with fuzzy sets defined by functions). This restriction simplifies calculations considerably, and is appropriate for the current study.

Preparata and Shamos (1985) note that there are several distance metrics available to measure inter-group distance, such as the Mahalanobis, Euclidean, Manhattan, and Minkowski L_p metrics. However, as noted above, one of the problems with a predefined class is the difficulty in utilising a centroid or group mean, due to non-linearity in the attributes (attribute space) or problems in assuming normality. These authors make use of a nearest-neighbour metric. Altman (1994) extends this metric to return a fuzzy set from the calculation. The nearest neighbour fuzzy distance metric (NNFD) between regions **A** and **B** is defined as:

$$dist(\mathbf{A},\mathbf{B})=\bigcup_{(a,b)\in\mathbf{A}\times\mathbf{B}}\left(\min\left(\mu_A(a),\mu_B(b)\right)/\,d_2(a,b)\right) \tag{4.7}$$

where $d_2(a,b)$ is the distance between elements a and b using one of the metrics defined above (Altman 1994). This NNFD metric results in a fuzzy set of distances and membership values. These values may be summarised graphically or 'hardened' in an application-specific manner.

For example, in Figure 4.9 the nearest neighbour fuzzy distance results for a simple set of two classes and a new individual are illustrated. The area represented by the 9 x 9 raster map is a simplified two-dimensional attribute space (all three raster maps represent the same attribute space—the objects are separated out for clarity). The object in the first view is a class, but defined using a fuzzy representation rather than hard bounds (a hard bounded class would be represented by all ones). In the second view a second, smaller class is shown. The third view shows a fuzzy sample. The two graphs on the bottom show the resulting distance from the sample to each of the classes. Instead of the single value that would result from a standard distance measure, each graph contains a fuzzy set of values. This set represents the distance from *all* of the sample to *the entirety* of the class. For example, the values above zero in class two (top middle raster) are contained with nine cells. The 'new individual' (sample) is contained within four cells. The graph on the lower right contains points that describe the *physical* distance between each of the four

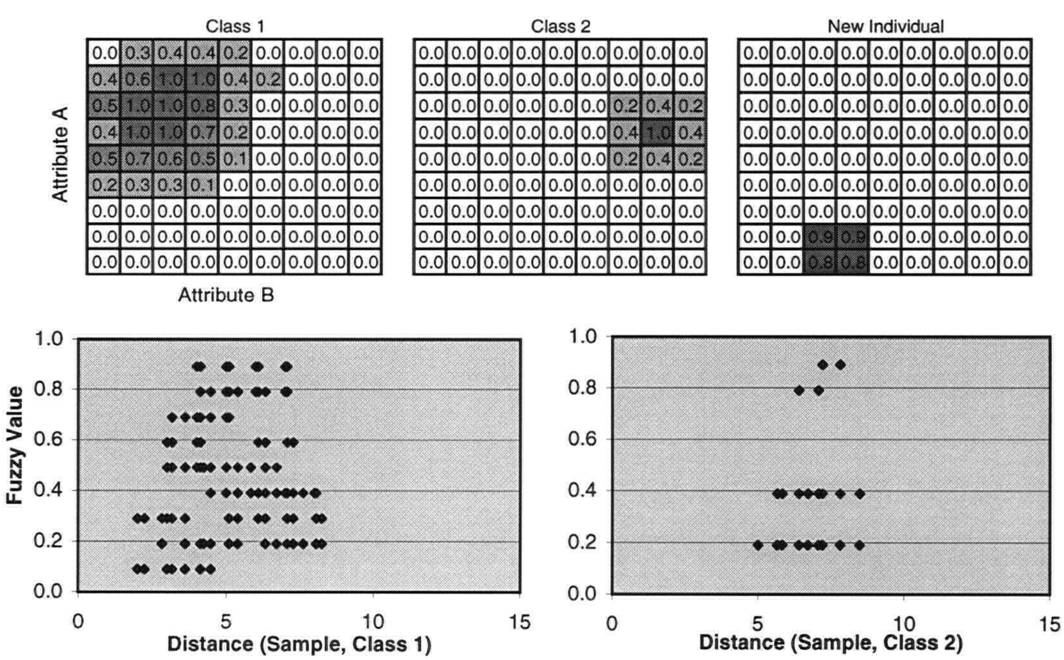


Figure 4.9. Sample to class distance defined using fuzzy sets in a simplified 2-D attribute space.

sample cells and the nine class cells (horizontal axis) and also describe the fuzzy overlay value, min(a,b), for the two (vertical axis).

Other Measures: A variety of combinations of the above measures are possible. For example, the distance from a sample with variability to a well-behaved (i.e., Gaussian on all axes) class could be accomplished with a vector series of measurements to the centroid, summarised (if appropriate, i.e., normal distribution) using a mean distance and a standard deviation. For example, if the second graph in Figure 4.9 was represented by a line around the boundary (and the raster was smoothed at a higher resolution), the result would be approximately a normal curve that could be represented by summary numbers (as long as the class was 'well-behaved').

Weighted measures would be appropriate when the various axes (attributes) contribute to the class definition to greater or lesser degrees. Although normalising of the attribute axes deals with numerical and number scale differences, there should be a way of de-emphasising non-crucial attributes in a similar manner to weighted fuzzy classification rules.

4.2.4. Summary of Theoretical Work

The sections above have introduced the concepts of fuzzy classification systems and the allocation of new members to existing sets using a variety of methodologies. Several possible distance metrics were introduced. Existing methods of fuzzy allocation based on centroids of classes were extended to include recognition of class boundaries using both hard and fuzzy classification functions. Several other methods were discussed that allow a fuzzy sample and a fuzzy class to be compared resulting in either scalars or new fuzzy vectors.

4.3. Application to Parameter Verification and Tuning

The distance metrics described in the previous sections are primarily useful in allocating new individuals in a fuzzy classification scheme. Application of one or more of these metrics will allow an uncertainty model to be verified using confirmatory sampling procedures. This is an important step in determining if the model adequately represents reality (adequate for the purpose of the model). In this section an appropriate distance metric is chosen and used to allocate new samples into the fuzzy classification scheme introduced in the previous chapter. The samples and allocation procedures allow the spatial structure of the uncertainty to be verified. In some uncertainty models the structure may have been inferred from other studies, or given some default smoothing value (e.g., Mark and Csillag 1989). In the scheme developed for the previous research, expert opinions about uncertainty levels and structure were coded using the Semantic Import (SI) Model (see §3.2.1.1). Parameter verification (and subsequent re-tuning) will allow a new uncertainty model to be produced that reflects the actual levels and structure of variability at the site.

4.3.1. Samples Required

The primary purpose of parameter verification in this context is to verify and update the *spatial* constraints of the uncertainty model. Classification uncertainty is not included as an explicit part of the model (e.g., fuzzy class functions, boundaries in attribute space, etc.). The classes utilised are considerably simpler than those used in standard soil taxonomy, for they focus specifically on soil cohesion parameters. The classes are predefined, and have rigid boundaries. However, spatial and classification uncertainty constraints create a model of continuous spatial variation of fuzzy membership values.

The key elements of the model that are subject to verification are the overall levels of uncertainty (maximum and minimum—referring to the possibility of misclassification) for each of the soil classes, as well as the spatial variation in uncertainty across the boundaries of the original polygonal layer. Sampling will therefore focus on these polygon boundaries as well as on the 'purity' of a class in the centres of the original polygons.

For simplicity of sampling and analysis seven attributes are defined, each of which can be represented by a percentage. The target classes are defined using these seven attributes, based on the original subdivision of a standard soil survey (see Davis 1994 for details on the original survey). The attributes include relative percentages of different grain size classes, as well as values for general origins: morainal, fluvial, etc. Due to the huge number of samples required, values were chosen that could be quickly estimated in the field (after calibrating the estimates; discussed further below).

4.3.2. Methodology

The purpose here is to verify the spatial structure of uncertainty used as input to the uncertainty model of slope stability. There are two main types of data (as discussed in §3.2), classified (formerly) polygonal—run through the transition corridor model, and non-classified continuous. Initially, this verification will focus on the classified data, in particular the soil data, as it is the most important and most variable classified input.

The transition corridor model makes use of two data sources: the original polygon locations and expert opinion about classification uncertainty and spatial structure of that uncertainty. The purpose of this current exercise is to verify the predicted spatial structure and, later, to revise the structure based on the results.

The first step in verifying the spatial structure of the uncertainty is to design a ground-sampling scheme that will capture this uncertainty in an efficient manner. The transition corridor model algorithm is designed to work, as much as possible, at right angles to polygon boundaries. Therefore, the most efficient way of sampling would be along transects that follow this direction. In ideal circumstances the transect locations would be chosen randomly within the study area. The transects would then be oriented to pass directly through polygon centres and to cross their boundaries at right angles. This would represent the most efficient way of verifying the spatial structure of the soil uncertainty model (due to the assumptions of the interpolation procedures; see §3.2.1 and Figure 3.3). Also, a substantial number of transects would be used. However, several practical constraints limited the number of transects and their locations. Transportation logistics limited the available areas to those reasonably close to roads. Time constraints and the high intensity of the sampling required to delineate spatial structure limited the number of transects possible. The extreme and often impassable nature of the terrain made it difficult to travel along an ideal transect line; suboptimal opportunistic transects were often required. The presence of a large numbers of bears in the area also precluded work in deep forest cover or valley bottoms. These limitations no doubt reduced the efficiency of the comparative procedures (the limitations are discussed in greater detail below).

The data were collected in a representative section of the test site described in Chapter Three (Figure 3.6). A total of 171 samples were collected using four transects, each crossing a number of original polygons (Figure 4.10). Due to the low incidence of certain soil types in the study area, only three of the original ten soil classes are part of the sampling (Types 1,2 and 3 in Table 3.2). These three comprise 79% of the total study area, and over 92% of the forested land. Spatial locations of the initial control points for each transect were recorded using GPS co-ordinates. Intermediate stations were surveyed using hand-held compass transits and tapes.

To match the resolution of the spatial model, sample locations were separated by 25 metres. Pit locations were chosen based on the typical surficial soil type within the surrounding 25 x 25 metre area. Estimated percentages of other soil type inclusions in the local area were also recorded. The soil was sampled for grain size as a percentage in seven classes, as well as for its relation to three origin classes: colluvial, morainal and fluvial.

There are two potential drawbacks to this type of sampling. First, by matching the resolution of the spatial model (at 25 metres), the samples may suffer from grid mismatching (ignoring spatial

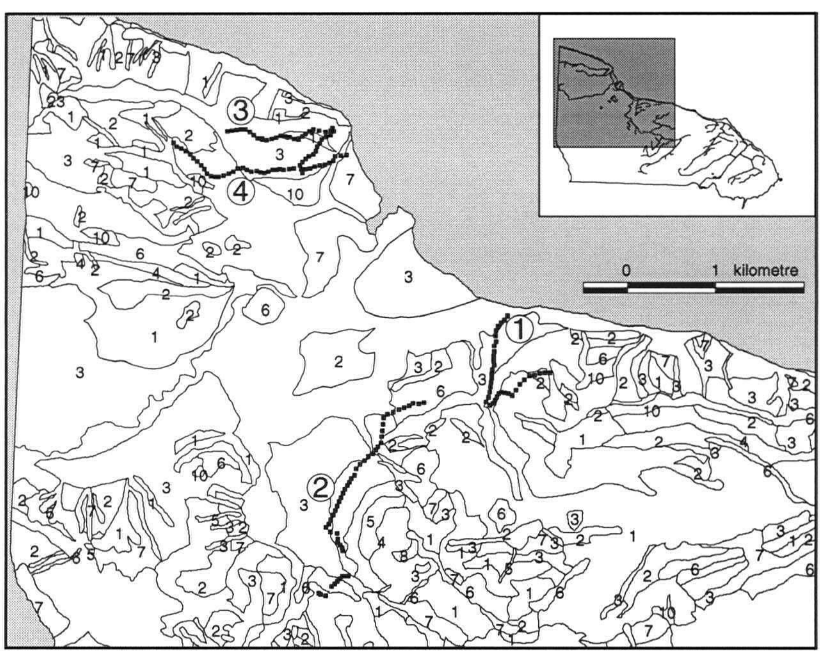


Figure 4.10. An overview of transect locations on the Louise Island test site. The circled numbers represent the transect numbers, while the smaller numbers represent soil types for the original polygons (for reference).

uncertainty for the moment), creating a mixed-pixel problem. Smaller sample separations would not eliminate, but would considerably reduce this problem. Although a separation of 10 metres was originally planned in the sampling scheme, initial test sampling work (prior to the beginning the first transect) indicated that 10 metre separations were highly redundant. The spatial resolution of the resource itself is at least 25 metres in the areas sampled. The sample spacing for the actual transects was therefore increased accordingly. Also, as the target of the verification scheme is the relative structure of the surface (rather than a point-by-point comparison), this mismatching of grids becomes less of an issue; mismatches are subsumed by structural smoothing (discussed in the next section).

The second potential drawback is spatial uncertainty. Tests using a GPS unit in the field indicated that 95% of all readings would fall within ± 110 metres of the true location (Appendix B details the method used; note that differential GPS, though preferable, was unavailable in this region). Spatial uncertainty in the original soil data is not known; however, minimum mapping units used in the survey indicate that resolution is on the order of 25 metres, and so well within the bounds of the GPS uncertainty. Therefore, each of the sample 'cells' could be misaligned with the soil model by up to (approximately) four cells distance. Note, however, that this only applies to the entire transect (i.e., a global transformation). Within the transect itself the uncertainty of location is several orders of magnitude smaller (approximately one to three metres—no further quantification was performed due to the small magnitude of this uncertainty relative to the others).

The direction of the transect relative to local polygon boundaries has considerable influence on the spatial uncertainty. In Figure 4.11 (an idealised demonstration), the modelled values (using the fuzzy model illustrated in Figure 3.3) are represented by the shaded squares, while the original polygon boundary is the wavy line. The transect at 90° to the polygon boundary (B) is only effectively misaligned along its length; misalignment parallel to the polygon border would have little effect on the results. However, a parallel transect's (A) misalignment would have the opposite effect. The only area where the parallel transect would be substantially uncertain is in the neigh-

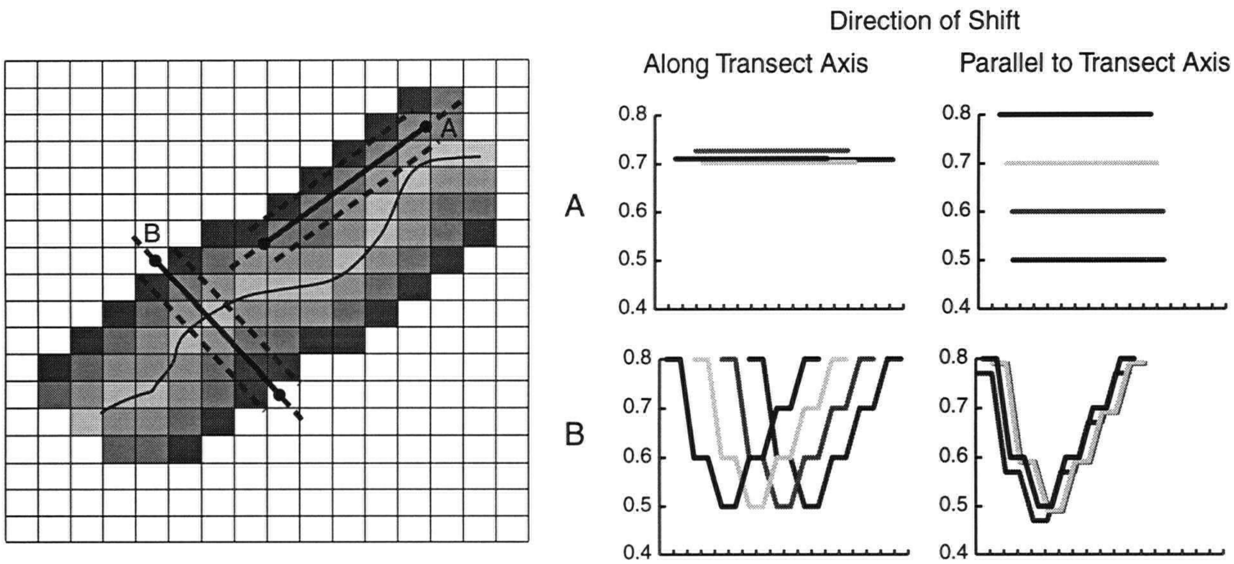


Figure 4.11. Idealised transects and the effects of shifting them within uncertainty bounds. Transect A, running roughly parallel to the original polygon boundary, is minimally affected by shifts along its length, while sideways shifts create changes in magnitude. Transect B, at 90° to the boundary, is significantly affected by shifts along its length, though sideways shifts produce minimal change.

bourhood of the original polygon boundaries because, if this transect were on the *wrong* side of the line, the relevant interpolation algorithm would be quite different, and so the values might shift drastically. If the parallel transect is located well within the current polygon (i.e., > ~100m from the boundary), the sideways variability would be less significant (although still a factor). Moreover, as demonstrated in Figure 4.11, the primary shift would be in *magnitude*. The analytical methods discussed below focus on relative change and are less sensitive to differences in absolute magnitude, so the problems of parallel transects are further minimised. Nevertheless, the issue of transect uncertainty when close to a modelled boundary is an important one. However, it is only relevant to *spatial* uncertainty issues (which may or may not be related to attribute uncertainty - see Goodchild 1991.)

Addressing this spatial uncertainty makes the task of analysis considerably more complex. To address the uncertainty in a truly comprehensive manner it would be necessary to generate numerous (anywhere from 12-25 or more depending upon the assumptions used) realisations of the model in a manner akin to the Monte Carlo techniques discussed in the previous chapter.

However, if suboptimal transects are eliminated, a reasonable solution (i.e., only slightly less accurate and much easier to interpret) is to perturb (i.e., offset) the transect locations only along their *length*. If the transects are at 90° to the boundary then this shift will (potentially) align the models and the sample if they correspond to some degree. Elements of a transect that parallel a boundary will not be affected by the shift (other than in absolute magnitude) because, in the model, paralleling a boundary generates a (roughly) straight line (Figure 4.11, upper graphs). The methodology used to implement this 'reasonable solution' is described below.

4.3.2.1. CROSS-CORRELATION

For comparison of the sample transect and the modelled values across-correlation statistic is utilised. The purpose of cross-correlation is to compare two or more data series and determine the strength of the relation between them. The offset (often termed 'lag' in reference to cross-correlation) at which the two are maximally equivalent can also be determined.

Designating n^* as the number of positions in the series, and Y_1 and Y_2 as corresponding values from sample and model, the cross-correlation for match position m is (Davis 1986):

$$r_m = \frac{n^* \sum Y_1 Y_2 - \sum Y_1 \sum Y_2}{\sqrt{\left[n^* \sum Y_1^2 - \left(\sum Y_1\right)^2\right]\left[n^* \sum Y_2^2 - \left(\sum Y_2\right)^2\right]}} \tag{4.8}$$

The significance test (using standard t-test tables) is (df = N*-2):

$$t = r_m \sqrt{\frac{n^* - 2}{1 - r_m^2}} \tag{4.9}$$

A value of 1.0 indicates perfect correspondence, while a −1.0 indicates perfect negative correspondence. Two random, independent series would generate a value of zero. The match position is varied within the bounds of spatial uncertainty, and the resulting set of values plotted in a cross-correlogram (Appendix C).

There is significantly more short-range variability on the transects than is apparent in the model. In order to better represent the general variability of the model, a series of smoothing passes are performed using a moving average equation based on orthogonal polynomials. Initial tests indicated that the relevant range of smoothing would be ±2 pixels to ±4 pixels (larger amounts of smoothing led to decreases in correspondence). Therefore, the smoothing equations utilised are limited to those that function within this range:

$$
\begin{aligned}
&\text{Avg}(5): \ \hat{Y}_i = \frac{1}{35}\left[17Y_i + 12(Y_{i+1} + Y_{i-1}) - 3(Y_{i+2} + Y_{i-2})\right] \\
&\text{Avg}(7): \ \hat{Y}_i = \frac{1}{21}\left[7Y_i + 6(Y_{i+1} + Y_{i-1}) + 3(Y_{i+2} + Y_{i-2}) - 2(Y_{i+3} + Y_{i-3})\right] \\
&\text{Avg}(9): \ \hat{Y}_i = \frac{1}{231}\left[59Y_i + 54(Y_{i+1} + Y_{i-1}) + 39(Y_{i+2} + Y_{i-2}) + 14(Y_{i+3} + Y_{i-3}) - 21(Y_{i+4} + Y_{i-4})\right]
\end{aligned}
\tag{4.10}
$$

To further clarify this issue, a set of idealised cross correlograms is illustrated in Figure 4.12, with the correlograms on the left, and the data they are based on placed on the right. The two values being compared are a modelled transect of an attribute and a measured transect of the same attribute. In the first pair (a), the model and the measured values line up exactly. This is reflected in the cross correlogram by a steep rise to a value of 1 at 0 (zero) offset. In (b), the modelled and measured values are similar, although the amplitude is different. The cross correlogram gives the same results as it is relatively insensitive to amplitude variations. In (c), the measured value is the exact opposite of the modelled value. Here, the correlogram is reversed, showing –1 at offset zero, indicating perfect negative correspondence. In (d), modelled and measured are similar in pattern, but are offset from each other by 2 distance units. This is reflected in the cross correlogram by a steep rise at an offset of +2. In example (e), a more complex situation is pictured. Although the two are in perfect correspondence—generating a steep rise on the left graph at 1, they would also be in perfect correspondence if the entire 'measured' curve were to be shifted left or right by 6 distance units. This offset correspondence gives rise to the peaks in the correlogram at +6 and –6. The correlogram also dips at +3 and –3, because an offset of +3 or –3 would give rise to high *negative* correspondence. In reality, the situation is rarely this unambiguous. For example, in (f) the effect of offsetting the modelled values would be minimal, therefore the correlogram shows high correspondence at *all* offsets. There is little information regarding *pattern* correspondence that can be gained from this graph.

(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)

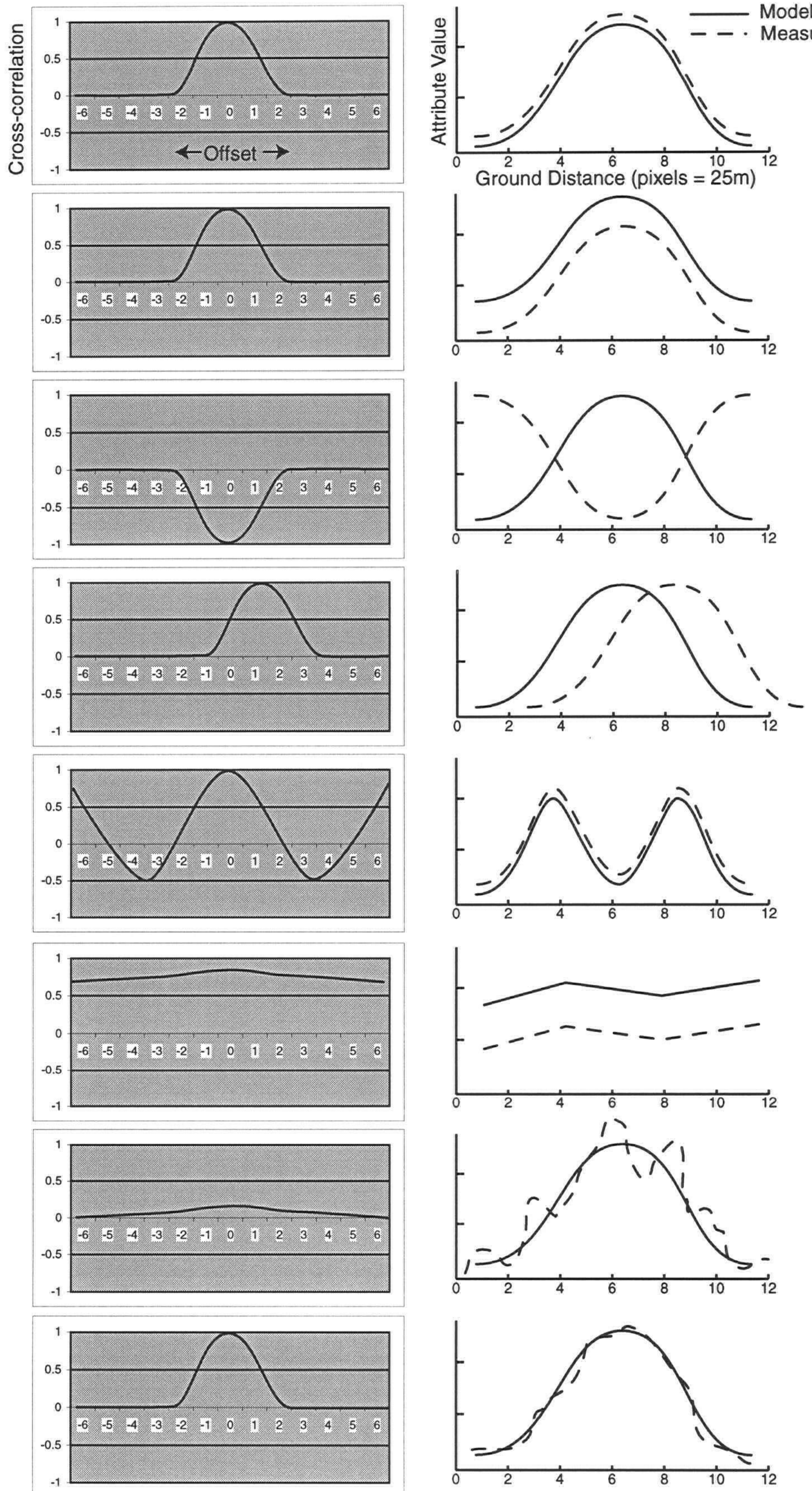


Figure 4.12. Simplified examples of cross-correlograms.

The purpose of smoothing is to eliminate short range variation that causes a reduction in correspondence between measured and modelled values. In Figure 4.12 (g) and (h) the effects of this smoothing are demonstrated. However, smoothing is not automatically called for. In cases where short term variability *is* the norm, then it would decrease overall correspondence.

4.3.2.2. DATA SUMMARY

The Mahalanobis distance (MD) metric is used to determine the fuzzy allocation statistics for each sample. In this instance the centroid of each class is used as a target. While this situation is not ideal, a lack of information regarding the detail of class structure makes it necessary. Due to the generalisations involved in simplifying the original soil classes, considerable overlap in class attribute space is expected. A continuous class structure is therefore utilised, with the fuzzy value $\varphi = 2$. This is the standard value utilised as a first approximation in fuzzy-*c*-means implementations (Appendix A explains the significance of this value). Variations in this value have been explored by Odeh *et al.* (1992) (also see Appendix A). Ideally, this allocation procedure would take place in concert with the primary soil survey. In that case the detailed class definitions would be readily available from earlier calculations.

The MD values assigned to each sample are then inverted (distance is the effective opposite of fuzzy membership), normalised (using the maximum value in the dataset) and scaled (globally) for comparison with the fuzzy membership values assigned by the uncertainty model. Scaling is not

	BR	Cob	Peb.	Sand	Silt	Fluv	Mor.	Coll.
Class 3	0	0	30	20	70	100	0	0
Weighting	0.17	0.17	0.17	0.17	0.17	0.05	0.05	0.05
Sample #17	0	0	20	40	40	0	100	0

$$\text{Membership} = \sqrt{(d_A)^2 * w_A + (d_B)^2 * w_B + \ldots}$$

where d = sample to centroid distance, w = weighting of attribute

$$= \sqrt{(30-20)^2 * 0.17 + (20-40)^2 * 0.17 + (70-40)^2 * 0.17 + (100-0)^2 * 0.05 + (100-0)^2 * 0.05}$$

$$= 35$$

$$\text{Normalise, scale and invert: } = \frac{\text{Max} - \text{Value}}{\text{Max}} * \text{Scale Factor} = \frac{66-35}{66} * 0.8 = \mathbf{0.37}$$

where 'Max' is the largest membership value in the dataset and 'Scale Factor' is determined by visually lining up result graphs (it has no significance in cross - correlation statistics).

Table 4.1. Calculation of the Mahalanobis distance for one sample.

strictly necessary, as the cross-correlation procedures ignore magnitude; however, scaling assists in visual comparison of the model and sample datasets. An example of the original data and calculated allocation values are shown in Table 4.1.

4.3.3. Results

The resulting cross-correlograms for the transects are displayed in Appendix C, as are calculated significance tests of cross-correlation at the 5% level. Smoothed values are also calculated in order to determine if large-scale variability is affecting the correlation coefficients. In most cases the smoothed values perform better than the raw data.

The cross-correlograms summarise a great deal of information, and are used here to develop general observations about the behaviour of the transect samples relative to the model. The transects are individually summarised below. Each of the paragraphs refers to the soil types defined in Table 3.2, where Type 1 = sandy/silty morainal blanket, Type 2 = silty morainal blanket/veneer, and Type 3 = rubbly, silty colluvium/morainal. In the discussion, a 'positive correlation coefficient for Type 1' would refer to a strong positive correlation between a) the Mahalanobis Distance values for Type 1 based on samples along the transect, and b) the values generated from the uncertainty model. The 'spatial offset' refers to the peak (or trough) at which the correlation coefficient is maximised.

There is some difficulty in applying standard terminology in a comparison of MD sample data and fuzzy modelled data. Using a fictitious example, in the original Boolean polygonal soil model, soil Type 2 is not present at location A—whose map co-ordinates are (3,7); however, the fuzzy model would represent it with some value (e.g., 0.2), based on possibilities of misclassification and other data discussed in the previous chapter, indicating that there is a small possibility that whatever is there would be misclassified as Type 2. The ground sample data run through the MD manipulation will also report some value at (3,7) for Type 2, even if the presence of bedrock makes the value zero. Therefore, when a type is termed 'not present', this only refers to the original polygonal data—its' uncertainty value is still present. Unfortunately, it is not possible to use the cross-correlation statistic to directly compare samples and model for these 'not present' types (and therefore compare the misclassification estimates for all soil types). This is because the calculations are not sensitive to absolute values, only relative changes. Therefore, two reasonably flat graphs (such as the graph typically generated by a 'not present' type) will show high correlation at all offsets, even if they differ somewhat in absolute values. This type of correlation is demonstrated in Figure 4.12(f). It may be possible to compare such values using simple averages; however, this extension to the work is not pursued herein (if it were possible to sample all soil types this work would have merit; with only one parameter to test—misclassification—the test would be of little use in model tuning); therefore, the focus of this analysis will remain on the three types deemed 'present' by the original polygonal soil map.

The following observations refer to the graphs (Figure C.3—'Original') of Appendix C. Note that correlations beyond an offset of +4 or -4 are considered spurious, as the spatial uncertainty is ~100m (4 x 25m cells). The graphs are extended to -6 and +6 to assist in trend visualisation. The paths of the transects are illustrated in Figure 413.

Transect 1: Comparison of samples and model along this transect indicates that Soil Types 1 and 2 both demonstrate substantial positive correlation coefficients (see Appendix C), and both of these exhibit a similar spatial offset (~+2). Soil Type 3 shows a weaker positive correlation, but at a very different offset (-4). It shows a strong negative correlation at +2.

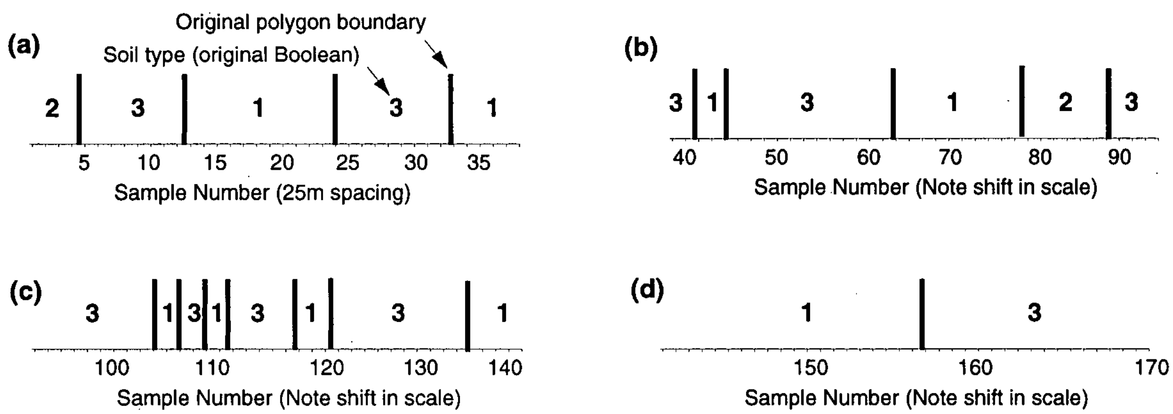


Figure 4.13. The sequence of polygons 'encountered' (i.e., using the Boolean model) on each transect. (a) Transect #1, (b) Transect #2, (c) Transect #3, and (d) Transect #4. The transects have been individually scaled to fit a standard length.

The following general conclusions can be drawn from this transect analysis. The +2 offset common to all three cross correlograms indicates that the transition between 'soil units' (using the term to refer to groups of similar soils) is likely being modelled well, though at a 50m (+2) offset. For soil Types 1 and 2 the model and the samples appear to correspond well; however, there is an apparent significant problem in one of the soil type definitions—possibly Type 3—indicated by the mirrored coefficients. This means that the values modelled are changing in the *opposite* direction than expected, though they are changing at the expected spatial location.

Transect 2: Again, both Types 1 and 2 demonstrate significant positive coefficients between samples and model, and the offsets fall within the same range. The range in which the offsets are high is broad, indicating that the transitions between soil units may not be very distinct. This may be explained by the fact that this transect spends some of its length in the vicinity of a 1/3 boundary. Because it does not cross the 1/3 boundary at the ideal 90°, the transition between the two on the graph is less distinct. Again, Type 3 demonstrates some negative correlation; almost the reverse of Types 1 and 2. This lends further evidence to the apparent soil type definition problem for Type 3.

Transect 3: This transect passes through a number of smaller areas, back and forth between Types 1 and 3 (Type 2 is 'not present', though is included in the discussion as noted above). In this case, correlation for Type 1 is highly negative at an ~0 offset; Type 2 is highly variable, with a small peak at 0, but a small negative peak at +3, and Type 3 has a positive peak (~0.4) around offset 0. As with transect 1, it appears that the common offset is 0, and again there is an apparent definition problem causing a 'mirrored' negative offset of Types 1 and 3, however it is Type 1 that receives the negative value this time. This provides some evidence that the type definition problem may be related to confusion or overlap between Types 1 and 3.

Nevertheless, there is also a polygon size influence occurring. The polygons encountered in this transect are smaller than in others (Fig. 4.13c). The results of cross-correlation may also be influenced by a poor representation of small polygons in the soil model and derived uncertainty model (due to cell size and implied scale of analysis). For example, at some points on the transect a 100m offset is sufficient to 'hop' from one Type 1 polygon into another, providing spurious results, as demonstrated in Figure 4.12e.

Transect 4: This transect crosses a 1-3 boundary between two large masses. Both Types 1 and 2 show peaks at a 0 offset, while Type 3 shows a somewhat flat line. Type 3 is obviously poorly represented in the uncertainty model; the flat line indicates that the single boundary present on the transect may also be poorly modelled.

These results indicate that the definition of the spatial structure of Soil Type 3 may be inadequate. This soil type (rubbly, silty, colluvium/morainal) is the only one defined that contains a wide variety of materials. The other types are much more specific. It appears that, on the ground, Type 3 and Type 1 have greater similarity than their definitions indicate. Classification confusion between the two is higher than predicted by experts. Overall, there are a number of points where the correlation peaks (or troughs) are located at very similar offsets. Although this suggests that the samples and model are in general agreement about the locations of boundaries, there are too many dependencies between the soil types (both samples and model) for this to be considered statistically significant, or the significance to be investigated.

4.3.4. Applying Changes

In general, the results indicate that expert definitions of classification accuracy and spatial behaviour of uncertainty in soil distribution are not ideal. Five of the cross-correlations did not achieve significance on the *t*-test (Appendix C); some of the more specific problems are noted above. In particular, the spatial boundaries and overall misclassification estimates between the models of Types 1 and 3 are apparently in error to some degree; a number of other parameters also may be suboptimal. However, though there is some indication of what direction these parameters should move (e.g., the misclassification values between Types 1 and 3 should increase), there is no direct way of determining the *degree* to which they should move.

The purpose of the routines described in this section is to attempt to re-set the parameters (originally defined by expert opinion) so that they provide a better match between the model and the samples. Given the complexity of the situation—the fact that a change in one parameter would affect others—there is no simple, direct way to effect such a change. One possible way of addressing this problem is an exhaustive search of *all* the possible parameter settings, with a check of cross-correlation statistics at every step. The target is to find a set of model parameters that achieves the highest overall cross-correlation for all transects simultaneously.

An iterative procedure is therefore applied to the soil data subset of the uncertainty model. The parameters defining misclassification and spatial behaviour (Table 3.1) for these three soil types are varied within reasonable bounds. Cross-correlation for all transects are calculated at each stage. When the parameters converge on the highest overall correlation coefficient for each transect the procedure is ended. Essentially, this procedure takes *all* potential values that the 'expert opinion' input could possible have, and builds a new set of cross correlograms for each of the many combinations (approximately 12^9) for each transect. It searches for the set of values that give the highest overall correlation in a maximum number of the graphs (though ensuring that the result is reached through convergence, rather than through erroneous peaks).

'Reasonable bounds' were used to reduce the number of nested iterations required to complete the procedure. They were chosen based on the original expert opinion estimates of the values (Table 3.1). For example, the possibility of Type 3 being misclassified as Type 2 was defined originally as 0.25. The values were therefore varied between 0.05 and 0.6 (an initial guess at possible parameter bounds). The results of the procedure discussed below were checked to determine if these limits were adequate (i.e., did any new misclassification values approach these limits). No changes were required.

The algorithm used is as follows:

---

1) define the bounds of the fuzzy misclassification and intrusion matrices (Table 3.1) (in the absence such 'expert input' the range of 0.1 to 0.9 would be used) and initialise with the lowest value;
2) (re-) calculate the correlation coefficients for all soil types;
3) if a coefficient is > previously stored max value, store this value and append the fuzzy values used in the calculation (retain previous maximum values);
4) repeat (2-3) with the next set of fuzzy values; continue until all combinations have been attempted (i.e., through nested iteration);

5) examine results and discard any maximum that shows no evidence of convergence.

The results of this procedure show considerable similarity for each transect (Table 4.2). For example, the misclassification matrix values calculated for Soil Type 1 on transect 1 and transect 4 (two transects crossing similar boundaries) are as follows:

Transect 1: Maximum correlation coefficient: 0.92 using (0.50, 0.15, 0.40)

(the three values in parentheses represent the relative possibility of misclassification for Types 1-3 respectively as discussed in §3.2.1. They are summarised in Table 4.2).

Transect 4: Maximum correlation coefficient: 0.41 using (0.65, 0.15, 0.65)

For Soil Type 3 the values are also quite similar:

Transect 1: Maximum correlation coefficient: 0.85 using (0.70, 0.10, 0.50)

Transect 4: Maximum correlation coefficient: 0.56 using (0.60, 0.10, 0.50)

In other words, the procedure described above was used to determine the model parameters that best describe the field data. This was done individually for each of the transects. The parameters chosen by the procedure (the three numbers in parentheses above) were quite similar for different transects, indicating independent confirmation of the values.

The results of this procedure were then used to reset the original misclassification matrices in the uncertainty model. Original values and updated values are displayed in Table 4.3. In general, it is

Type	Transect	Max CC	Lag	Avg	Misclass Values	Comment
1	1	0.92	2	9	0.50, 0.15, 0.40	large 1/3 boundaries
	2	0.56	2	5	0.60, 0.50, 0.50	one large 1/3
	3	0.55	-1	0	0.50, 0.15, 0.70	small 1/3, small inclusions of 1
	4	0.41	-4	5	0.65, 0.15, 0.65	one large 1/3
2	1	0.83	3	9	0.70, 0.50, 0.60	one section at the edge
	2	0.76	0	9	0.25, 0.50, 0.10	one large section
	3	0.35	-4	0	0.70, 0.50, 0.55	no 2 present
	4	0.70	0	9	0.30, 0.50, 0.20	no 2 present
3	1	0.85	3	9	0.70, 0.10, 0.50	large 1/3 boundaries
	2	0.50	2	9	0.55, 0.60, 0.50	one large 1/3
	3	0.56	1	9	0.15, 0.10, 0.55	small 1/3, small inclusions of 1
	4	0.56	1	9	0.60, 0.10, 0.50	one large 1/3

Table 4.2. Maximum values of correlation coefficients obtained through iterative calculations.

(a)

Soil Class	1	2	3
1	0.85	0.35	0.25
2	0.35	0.85	0.25
3	0.25	0.25	0.90

(b)

Soil Class	1	2	3
1	0.60	0.25	0.55
2	0.25	0.55	0.30
3	0.51	0.10	0.55

(c)

Soil Class	1	2	3
1	-0.25	-0.10	0.30
2	-0.10	-0.30	0.05
3	0.26	-0.15	-0.35

Table 4.3. Original (a) and updated (b) misclassification matrices, and the difference between the two (c). The value at c_{ij} represents the possibility that type *i* would be misclassified as type *j*. The chart's trace contains maximum certainty values.

evident that all maximum certainty values have decreased substantially, and that classes one and three have been considerably 'blurred'. Note also the lack of symmetry in the second table. According to the optimisation procedure results, the possibility of mistaking Type 3 for Type 1 is different than the possibility of mistaking Type 1 for Type 3.

4.4. Discussion

These confirmatory sampling and fuzzy allocation procedures indicate that, in this soil and slope stability modelling scenario, expert opinion has not provided an accurate assessment of classification uncertainty. However, there are only 171 samples over four transects available to base this conclusion on. Sampling for spatial structure is very time and resource consuming relative to spot sampling; it would be difficult to gather sufficient amounts of transect datasets to even approach statistical certainty. When 40 or 50 individual soil pits are subsumed into one statistic (such as the cross-correlation peak or *t*-test for one transect), it becomes difficult to generate a sufficient *n* to satisfy most statistical tests. Moreover, most standard statistical comparison techniques require independence of samples; by its very nature the samples composing a transect are *not* independent, although the transects themselves are independent from one another.

Here, cross-correlation is used to compare sampled transects with modelled transects. Cross-correlation provides two results: the strength of the relations between the two series, and the offset in distance between them at their position of maximum correspondence. In this section, the parameters of the model generating the modelled transects were reset using an iterative procedure. Resetting the parameters of the model has increased the strength of the relations between the sample transects and the model. Details are provided in Appendix C.

However, this 'increase in strength of the relations' is only a statistical comparison between the actual and modelled versions of a slope stability model input (soil). The *true* test of these confirmatory samples and changes in model parameters will be a comparison of the model's final results—slope stability uncertainty—with some actual slope stability data. Only then will it be possible to see if the changes wrought by this procedure have any real value in increasing the predictive capability of the slope stability uncertainty model. A high-resolution dataset of slope failures is required to perform this task. The following chapter is devoted to analysis of both realisations of the uncertainty model (i.e., pre- and post-calibration) using such a dataset.

4.5. Calibration of Continuous Data

The work discussed above focuses on calibration of classified, polygonal data in a spatially-oriented uncertainty model. Although the focus is on soils, the techniques developed or adapted are of use with any type of data that have been broken into classes and 'shoe-horned' into a polygonal structure (i.e., they have a more continuous distribution in reality than polygons would indicate.) This section focuses on calibrating data of a different nature: continuous data represented by cardinal values. This type of calibration or verification is not typically a difficult procedure, as values can be directly compared between the model and samples. The only major complicating factors are spatial uncertainty causing mismatches between the two, and issues of sampling scale.

The two principal inputs to the slope stability model are soil type and slope; the previous section used the former as an example, while this section addresses calibration of the latter (again, as an example of a typical set of continuous values). In the work discussed in Chapter Three, the level of uncertainty in the slope values was estimated from the published error statistics for the data source and from the Kriging function used to generate the elevation model. These values were propagated through the slope function using Monte Carlo techniques, and a final value and variance were derived for each cell in the model (details are in §3.2.2).

Typically, calibration of a continuous value does not require advanced methods, just a numerical comparison. However, this calibration is of particular interest due to a heavy reliance on elevation data in many sectors of natural resource management. Many decisions are made based on these data (e.g., sight-lines, slope stability, road placement, growth models), and few independent studies have been undertaken to verify the elevation data or their derived products such as slope. There are no known studies from the test region or, in fact, the Queen Charlotte Islands in general. Again, it must be noted that the calibration focuses on values falling within the estimated uncertainty bounds. In the case of continuous values, however, this comparison is much easier to achieve than with classified values.

The soil sampling effort described in the previous sections also included slope measurements at all sample points. Additional slope measurements were made on an opportunistic basis, providing a grand total of 240 sample points. The purpose of this sampling was not to perform a *thorough* statistical analysis of elevation data. This type of work has been undertaken by several authors (e.g., Xiao 1996; Ruiz 1995), who generally conclude that elevation error is specific to regions, terrain, methods used, etc. Instead, the specific purpose is to characterise any significant consistent deviation from the elevation-model-based slope that fall outside of the error estimates for this particular area, and to use this information to perform a correction of slope values. However, if significant deviation *does* exist, this fact implies that other areas where similar data are used may also be subject to deviations outside the published error statistics.

Slope values were sampled using a hand-held inclinometer with a tested accuracy of ± 3°. Sample values were gathered as above—characterising the local 25 x 25 metre zone. By focusing on the average slope over this area, rather than a 'spot' slope measurement, the value recorded for the sample eliminates short-range variations in slope. This fits the assumptions of the elevation model that is being calibrated (where 25 metre cells are used to best characterise source spot heights at an average 30 metre spacing). Positional uncertainty is an issue, as was noted with the offsets used for soil transects above. Initial tests were performed in which the comparisons discussed in the following paragraph were repeated using random shifts in cell location within the bounds of locational uncertainty. However, these tests indicated that the shifts had no significant effect on the results. The averaging nature of slope calculations (where a cell only has a slope *relative* to the eight cells that surround it) is likely the reason for this lack of significant change. Therefore, positional uncertainty was ignored in the following procedure.

The variance in the slope values was derived as discussed in Chapter Three. To reiterate, elevation values were gathered from published spot height data derived from photogrammetry. The elevation model was produced using Kriging, and elevation uncertainty was determined by combining the variance output from the Kriging procedure with the published error statistics for the spot heights. The predicted variance in slope was determined through a Monte Carlo procedure, in which an 'equally likely' DEM was produced, a slope surface derived, and then the procedure repeated ($n = 50$). The predicted slope variance values were saved on a cell-by-cell basis, but are summarised in the figure discussed in the following paragraph using a set of lines (those above and below the zero line).

A comparison of the modelled and sampled slopes (Figure 4.14) illustrates considerable variability between the two outside the bounds established by variance calculations. However, a definite trend is apparent in the data. Gentle slopes are generally overestimated by the source data and slope modelling procedure (i.e., the inclinometer-measured values were consistently gentler than

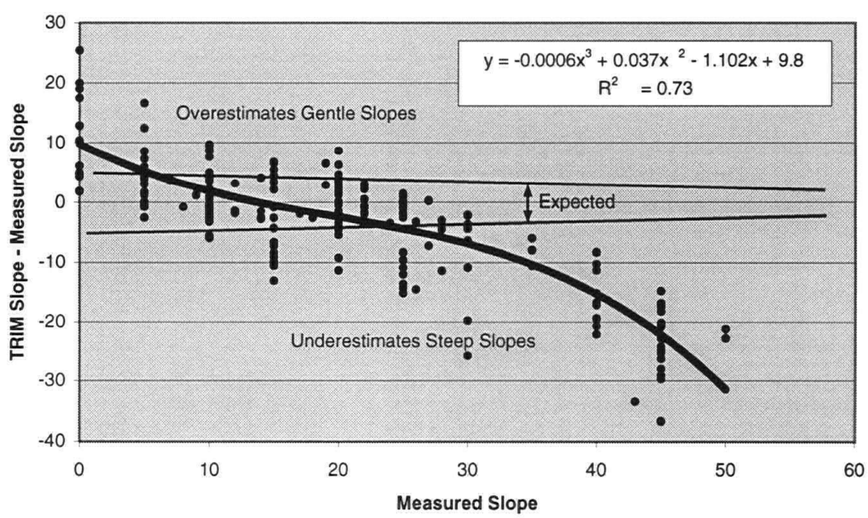


Figure 4.14. Differences between measured slope and TRIM modelled slope graphed relative to measured slope. The 'expected' lines enclose the variance that is expected to be found in the values, based on the published error values and the additional error introduced during DEM generation. Variance is generally larger on gentle slopes.

the slope values derived from the spot height data), while steep slopes are generally underestimated. The trend-line is represented by:

$$Y = -0.0006x^3 + 0.037x^2 - 1.02x + 9.8;\ R^2 = 0.73 \tag{4.11}$$

This formula (4.11) can be used to correct calculated slope values for this particular region. It would be unwise to use this formula to generalise outside of the study area and its environs, as it is unclear whether the extreme terrain, the contractors who produced the elevation spot heights, or the slope modelling routine assumptions are to blame. For example, the photogrammetric equipment and algorithms used to derive elevations for the source data might have been miscalibrated, or designed for less extreme terrain. The abrupt changes between forested and cleared areas may not have been compensated for (stem density tends to be very high in this area). The slope algorithm used in the GIS could also have a bearing on the error; such derived values are rarely checked in the field. This would be an area for further study. Some related work has been performed by Ruiz (1995) and Xiao (1996).

Nevertheless, whatever the cause, within this type of terrain and in areas where the same photogrammetric equipment and routines were used to gather elevation data, this slope correction should provide more accurate values for stability modelling. This correction factor was applied to the dataset, and the corrected values will be employed in the following chapter when implementing a slope stability model on an island near to the current study area. This correction has reduced variability, but the variability of the corrected slope values are *still* greater than the published statistics indicate (demonstrated by the vertical range in Figure 4.14—which still fall outside the variability line after the trend is straightened out). Therefore, the overall variability values are redefined for the calibrated version of the model, based on these post-correction values.

4.6. Conclusions

The calibration and verification of uncertainty models was noted as a neglected area in the development of such models. Two major areas were identified: the verification of classified 'fuzzy' values and the verification of continuous values. Several procedures were proposed, including a number of possible extensions to the Mahalanobis distance metric for dealing with class variability in attribute space. Sample 'fuzziness' was also noted as an issue, and a complex measure of distance were adapted for use in measuring and determining fuzzy sample allocation statistics.

The principal questions addressed in this chapter were: 1) how can fuzzy classification structures be compared with confirmation samples?; 2) how well did expert opinion function as an input to generate the distribution of uncertainty represented by the fuzzy structures (the transition corridor model)?; 3) how well does metadata gathered from published statistics represent the actual uncertainty on the ground? (focusing on major model inputs); and 4) how can these confirmation data be used to recalibrate the model?

Methods for comparing fuzzy classification structures with confirmation samples were developed using the extensions to the Mahalanobis distance metric; other issues were presented, such as sample variability and complex distance issues. A subset of these methods was utilised to calibrate the uncertainty model developed in Chapter Three using field transect data. It was determined that the expert opinion input for classification uncertainties did not adequately describe the necessary values, due in part to the nature of the classes utilised.

The metadata gathered from published statistics were tested against ground data for the major continuous input to the model: slope percentage. Analysis of field data indicated that the continuous values in the slope dataset were outside their expected zone, as determined by stochastic simulation using source data error statistics. However, a clear trend was found to be present in the model-sample comparisons, and therefore global corrections were possible. The model was recalibrated using these data. The soil uncertainty values were also recalibrated based on the confirmatory data using an iterative procedure, utilising cross-correlation analysis incorporating sample spatial uncertainty.

Having introduced the uncertainty model and its application to slope stability modelling in Chapter Three, and having presented methods for field-verifying and updating the inputs and parameters of such a model, at this point there remain several crucial unanswered questions. First, how useful is the uncertainty model? The principal model parameters (though not all) have now been calibrated, but currently the model results are simply predictions of possibility of slope failure and level of certainty in that prediction. The basic question is: does the uncertainty model accurately predict *uncertainty* in the slope stability model? For example, in areas where slides have not occurred, did the model either predict a low factor-of-safety (FS) or predict a high FS with high uncertainty? A second hypothesis follows from the work in this chapter: does the updated version of the model better predict slides and slide uncertainty than the original?

Chapter Five

Evaluation of Uncertainty Model Output

5.1. Introduction

The confirmatory sampling undertaken in Chapter Four addresses the problem of tuning uncertainty model parameters. This type of information is useful for a purely descriptive inventory model, or when inventory information is used as input to a process model such as slope stability. However, much of the work in Chapters Three and Four refers to the *output* of a slope stability process model—one of the essential items of information used to make forest management decisions. The model itself predicts relative levels of stability, while the uncertainty model predicts the variability in these results. This chapter focuses on confirming the latter through use of a highly accurate landslide database.

The work discussed in this chapter is applicable to many types of inventory uncertainty models; however, slope stability model evaluation is important in its own right. Slope stability models are rarely evaluated in a data-rich environment. Typically, they are tested on very limited areas, then applied in a wide variety of situations (Christian *et al.* 1994). Proper evaluation of a model requires detailed landslide data that cover a wide temporal swath. This type of information is rarely gathered at a scale appropriate to mass wastage analysis (Aung 1992). Even though this information is available here, the relative nature of the model's predictions (i.e., there is no 'slide/no slide' cut-off line) precludes a complete evaluation. True 'evaluation' of a model that makes use of relative predictions requires something similar to evaluate it against—typically another model or another version of the current model. Evaluating it against reality is more difficult.

Process models incorporating spatially variable uncertainty are a relatively new development. There are few guidelines for developing methods of evaluating their predictive capability. In the case of the slope stability uncertainty model there are several issues that circumvent simple analysis. These include:

1. **Binary events predicted on a cardinal scale.** The predictions made regarding slope stability are on a cardinal (i.e., approximately one to six) scale. In the absence of prior studies that calibrate these numbers for the region, there is no obvious cut-off line between 'slides will occur' and 'slides won't occur'. However, this information is to be compared with just such yes/no events. Simple summaries and tests of significance are therefore not available to fully evaluate model performance.
2. **Evaluating predictions with variance.** The uncertainty model would no doubt be considered successful if all mass wastage zones were predicted with low factor-of-safety values and tight standard deviations. However, it could be considered equally successful if very few slides were predicted, but associated variance was very high. Although in such a case the predictions for mass wastage would be useless, the uncertainty model would be illuminating the fact that the source data are of insufficient quality to support predictions (in itself an important output). This considerably complicates the process of evaluation.
3. **Grid model.** The raster model used for this procedure involves some different assumptions than the vector model used to gather the mass wastage information. For one, the resolution of the raster model is different than the vectors (25m vs. 1m or less). Slide zones smaller than a 25 metre pixel will therefore be poorly modelled, leading to inaccuracies. Similarly, the smoothing of slide boundaries will also affect model predictive accuracy. Another raster issue is the multiple predictions applied to slide areas larger than one cell. This variability must be addressed in the evaluation. However, a raster model is required by the slope

stability uncertainty modelling procedure, whose focus is the *continuous* variation of uncertainty across the landscape, and therefore these raster-vector issues must be addressed.

4. **Spatially variable variance**. For similar reasons to the point above, multiple cell slides will not only incorporate variance between cell predictions, but also variance *in* each cell's prediction.

5. **Multiple realisations**. The fuzzy model allows multiple realisations of its output; the user must choose the appropriate way of visualising or using these data. Therefore, there is no single answer to the question of model confirmation.

6. **Incomplete data**. Although the mass wastage database used in this evaluation has both high spatial resolution and a wide temporal extent, it does not capture every possible slide. There are undoubtedly still some areas that have yet to slide due to past forestry activity, and pre-logging slides are only partially captured (i.e., they are outside the temporal extent of the model). This will contribute to apparent inaccuracies in model predictions.

7. **Autocorrelation**. As with most spatial models, contiguous spatial units cannot typically be considered as independent samples. In the case of slope stability analysis, if one cell contains a slide, it is very likely that the one below or above it (on the slope) also contains a slide. It is almost as likely that the ones beside it are slide zones as well. This lack of independence violates many statistical test assumptions. It is therefore necessary to rely on a number of descriptive or exploratory techniques in their stead.

This evaluation therefore relies primarily upon techniques of exploratory spatial data analysis (ESDA, see Keller 1994), coupled with some standard statistical tests where appropriate. It provides no single answer to a hypothesis of predictability. Instead, it offers comparisons of a number of model realisations and methods of summarising model predictions. As with many other uncertainty analysis techniques, this lack of a single answer may frustrate those accustomed to seeing the resource analysis world in black and white. However, this is balanced by the increase in information content regarding the process model, the source data, and the field site itself.

5.2. Methodology

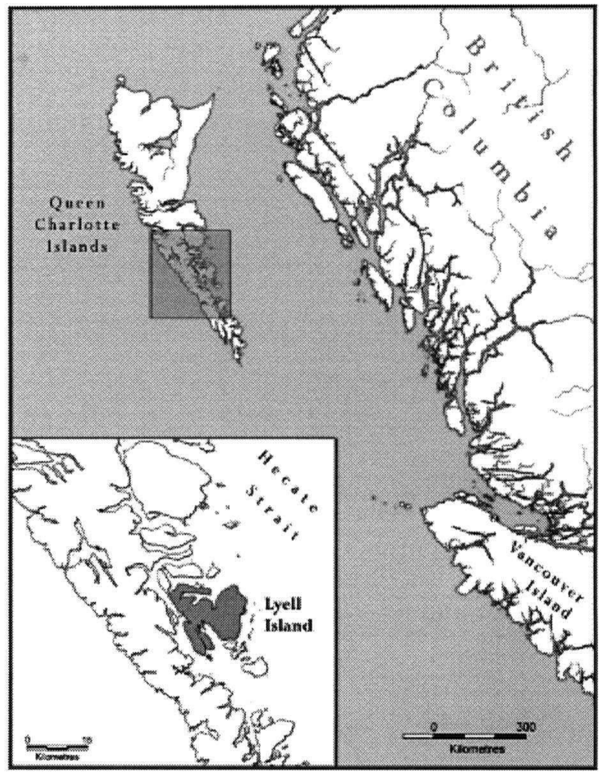


Figure 5.1. The Lyell Island study area.

The slope stability uncertainty model developed and calibrated on Louise Island (Chapter Four) is applied to data of similar source and resolution on Lyell Island, 50km to the south (Figure 5.1). The two islands have similar terrain, similar soils, and were subject to similar resource extraction methods. There is reason to believe that slope processes are similar on both (B. Peters pers. comm.). Although it would be ideal to calibrate the model at the original test site, the second site was used for practical reasons. Funding was available for mass wastage database development for only the latter.

Soil and elevation data are processed in a manner similar to that described in Chapter Three. The calibrated version of the model (calibration includes both soil parameters and updated slope values based on field testing) is run using the methodology described therein, and results are produced at the same spatial resolution (25m). A second set of results is produced using the original (not calibrated) model parameters and slope values to facilitate comparative evaluation.

The mass wastage data are taken from a database developed for this work. The database covers over twenty years of slide history in the study area, has a spatial resolution of approximately 1 metre, and an average accuracy level of under 3 metres. The tools used to develop the database make use of uncertainty visualisation techniques in a data fusion tool that merges oblique frames with planimetric data. The details of the system and its development are provided in Appendix D. This development represents a new way of entering data into a GIS through the use of uncertainty tracking and uncertainty visualisation. It was an integral part of building the database used in this chapter.

The landslide data were gathered for two complimentary purposes. The first is the uncertainty model confirmation discussed herein. The second purpose was to provide a baseline and initial high resolution dataset for on-going monitoring of landslide stabilisation efforts. A summarised version of the case study is presented in Appendix E. The full case study is described in Davis *et al.* (1998). Appendix E also includes (in context) details of construction of the database used in this chapter.

All slides in the database are utilised (going back to those visible in 1976), including those that have since stabilised and grown back. Divisions between spatially contiguous slides are dissolved, and a raster database is produced at the working resolution of 25m.

Two general hypotheses are proposed:

1. H_0: Predictions in slide zone cells are not significantly different than predictions for the population (all cells).
2. H_0: Predictions made by the original version of the model (prior to parameter calibration) are not significantly different than those made by the calibrated version produced in Chapter Four.

These hypotheses are tested using summary values in several different contexts.

The exploratory analysis methodology uses the following sequence:

1. Analysis of data means;
2. Presentation of alternative realisations of the uncertainty model;
3. Description and analysis of data variance;
4. Graph and comparison of expected vs. actual values;
5. Incorporation of spatial constraints.

5.3. Results

The infinite slope stability model predicts where areas of greater or lesser slope stability occur based on local slope, soil type, and ground cover. The model of uncertainty in slope stability introduced in Chapter Three is a generalisation of this model, in which the mean value of the maximum likelihood output corresponds to the original (Boolean) model. The uncertainty model contains a considerable amount of additional information about alternate possibilities and variance in the results. With this additional information comes an increased responsibility to understand the details of the model, the assumptions built in, and the implications of the data in order to properly analyse and communicate this information.

In this discussion of results a number of different methods are used to compare the model predictions with the landslide data. The discussion progresses from a simple non-spatial statistical comparison of means through to a number of different methods of dealing with variance in the results and spatial constraints on the model.

5.3.1. Comparison of Means

An initial step in comparing the predictions and the landslide data is to simply see if the slide areas (mapped as discussed in Appendix D) have predicted factor-of-safety values (slope stability predictions) that are significantly different than non-slide areas. This involves comparing the mean values of the factor-of-safety for the population (all cells) and for the slide cells.

As noted in the Chapter Three, the uncertainty model can be viewed using a variety of 'realisations'. For each cell in the model, all possible combinations of inputs (soil and forest classes) are stored with their associated likelihood. A 'realisation' of the uncertainty model involves choosing from among these possibilities in a structured manner. The realisation utilised initially in this means comparison is maximum likelihood, in which the most likely value for each cell, as defined by the fuzzy overlay value, is fixed. This realisation produces the same numbers as would be obtained using the standard (Boolean) version of the slope stability model. The values are presented in Table 5.1., row #2. Other realisations presented in this table will be discussed in upcoming sections. The values in this table are based on the calibrated version of the model as produced in the previous chapter. They are presented together here to facilitate comparison at later stages of analysis.

	As Cells	Mean	SD	N	Z-test	Probability *
1	Slides: Max Likelihood Realisation, Std. Dev. Surface.	0.040	0.020	1801		
2	Slides: Max Likelihood Realisation, Factor-of-Safety Surface	1.61	0.48	1801	-18.33	0.99
3	Slides: Worst Case Realisation, Std. Dev. Surface	0.030	0.023	1593		
4	Slides: Worst Case Realisation, Factor-of-Safety Surface	1.19	0.65	1593	-44.9	0.99
5	Slides: Worst Case Realisation, Std. Dev. Sfc., Upper 50% of slides	0.025	0.017	683		
6	Slides: Worst Case Realisation, Factor-of-Safety Sfc., Upper 50% of slides	1.09	0.57	683	-33.58	0.99
7	Population: Max Likelihood Realisation, Std. Dev. Surface	0.051	0.024	1825		
8	Population: Max Likelihood Realisation, Factor-of-Safety Surface	1.87	0.61	1825		
9	Population: Worst Case Realisation, Std. Dev. Surface	0.34	0.68	1825		
10	Population: Worst Case Realisation, Factor-of-Safety Surface	1.00	0.94	1825		
11	**As Slide Units**	1.55	0.22	154	-6.65	0.99

* Probability that slide cells are a different population than non-slide cells

Table 5.1. Summary statistics for slope stability predictions, based on mean values. Rows 1 through 10 show various realisations of the uncertainty model, and summarise values for either 'slide' cells--where the model predictions were correct, or 'population' values, for the entire area. The paired rows (e.g., 3&4) show summaries for the standard deviation surface and the prediction surface for a particular realisation. The realisations listed here are introduced one at a time as the chapter progresses.

Although the Z-test statistic is not entirely appropriate, due primarily to the lack of random sampling (the population is roughly normally distributed based on χ^2 at 15% significance), the very high value displayed (-18.33) indicates that the mean predictions of factor-of-safety in slide zones are significantly different from the population (i.e., the slope stability analysis is generally capable of distinguishing slide zones from non-slide zones). However, the probabilities associated with the Z statistic cannot be reliably estimated due to the violation of statistical assumptions. Further on in this chapter a sampling method will be discussed that bypasses some of these violations.

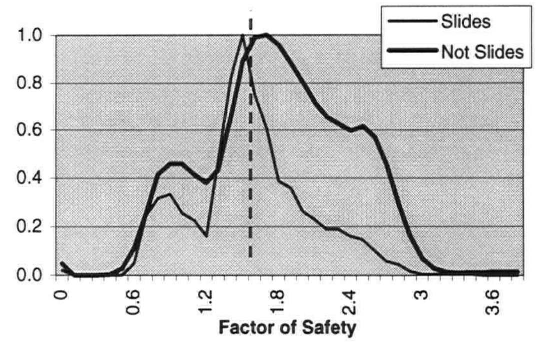


Figure 5.2. Relative frequency of slide zones and non-slide factor-of-safety values using an ML realisation.

These simple summaries do very little to explain the detailed information contained in the slope stability predictions. A first step in expanding the analysis is to look at the relative frequencies of the results. A pair of histograms illustrating the factor-of-safety (FS) predictions for slide zones and the non-slide zones are presented in Figure 5.2. The peak of the graph is chosen as a rough dividing line between 'slide' predictions and 'non-slide' predictions. This means that:

a) everything under the slide curve (light line) on the *left* side of the divide represents cells that were predicted as 'will slide' and did actually contain a slide (correct prediction);

b) everything under the slide curve on the *right* side of the divide represents areas that were designated 'safe', but actually contained a slide;

c) everything under the 'non slide' curve on the left represents areas that were predicted as slides that did *not* slide; and

d) Everything under the 'non slide' curve on the right represents areas predicted as 'safe' that *are* safe.

Keep in mind that this is a histogram, so the absolute values of the curves are not an issue (i.e., in an absolute graph the 'slide' curve would be tiny relative to the 'non slide' curve).

This highlights an important distinction in slope stability modelling: there are two types of wrong answer, comparable to Type I and Type II errors in hypothesis testing. Predicting that an area will fail when it does not (hereafter called Type A) creates only an economic problem (e.g., trees in the area are not harvested when they could have been). However, *not* predicting a slide that *does* occur (Type B) can have more than economic consequences. Damage to personnel, equipment and infrastructure can occur, in addition to environmental damage such as stream degradation and loss of soil. In this ML realisation, one-half (47%) of slide cells that were predicted as 'safe' (using the rough estimator discussed above) fall into the crucial Type B class.

In a typical Boolean analysis this would represent the final result. One-third of the cells would be poorly predicted, and we would go back and try and revise the slope stability model to increase predictive accuracy. However, the uncertainty model has retained a considerable amount of data regarding alternative realisations that can assist in reducing this Type B prediction error.

5.3.2. Alternative Realisations

As introduced and briefly demonstrated in Chapter Three, there is a wide range of possible realisations of slope stability uncertainty. The inclusion in this model of spatially variable certainty factor values allows more than the most probable value for each cell to be displayed. For example, the 'worst-case scenario' (WC) realisation utilises the (application-specific) lowest reasonable value for factor-of-safety rather than the most probable (where 'reasonable' is typically defined through an iterative process using a sliding scale). The WC version of the model allows error to be focused on the side of caution, by decreasing the possibility of Type B error at the expense of Type A. Note, however, that the acceptable balance between these two types of error must be determined by the application. As discussed in earlier chapters, this 'tolerance of risk' depends upon a number of external factors. The uncertainty model has retained sufficient information that the analyst can make this type of trade-off in the final stages of analysis (i.e., visualisation for decision support) rather than at the earliest data gathering stages (as with standard Boolean models).

A WC realisation was produced using the lowest factor-of-safety value predicted for each cell at a reasonable (i.e., > 0.55) certainty factor. Figure 5.3 shows the same values as Figure 5.2, but using this separate realisation. If we use the same dividing point as the previous figure (FS = 1.6 – the dotted line on the right) there is now only 20% of the slide area on the right side (predicted as 'safe', using area under the curve). Even if the dividing line is moved to the new graph peak (reflecting the lower mean FS that results from pessimistic predictions—see Table 5.1, rows 3 and 4), the right side still holds only 32%. Type B error has been reduced. However, this is evi-

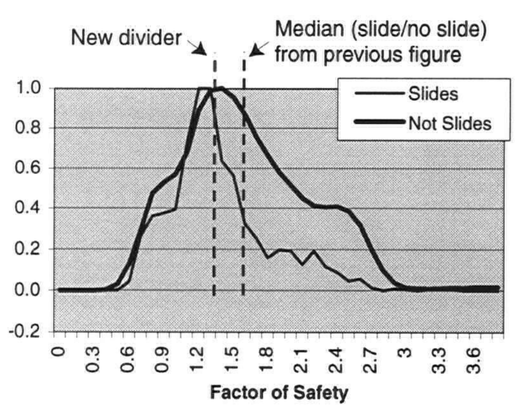


Figure 5.3. Relative frequency of slide zones and non-slide factor-of-safety values using a worst-case realisation.

dently at the expense of Type A error. The Type A curve (the 'non-slide' line on the left of the divide) has increased in relative area from Figure 5.2.

A variety of other realisations are possible, as the uncertainty model delays 'hardening' the data into a specific state for as long as possible in the analysis process. The purpose of the analysis will determine the appropriate realisation, as there is no 'best' way of looking at the data. For example, a road building project would have different requirements from a harvesting risk model, and a seismic crew utilising the data would have their own specifications for acceptable uncertainty.

5.3.3. Variance

While alternate realisations make use of some of the unique characteristics of an uncertainty model, the variance values have not, as yet, been utilised. The variance values represent, on a cell-by-cell basis, the spread of output generated by the Monte Carlo simulations in the uncertainty-based slope stability model (see §3.3.1). Once you decide on the realisation you will use, each cell *in* that realisation has a slope stability prediction (factor-of-safety number) and an associated standard deviation. The standard deviation in each cell will differ between realisations. This standard deviation value will help to determine with what certainty a particular prediction has been made, and how this uncertainty varies with both spatial and attribute variables.

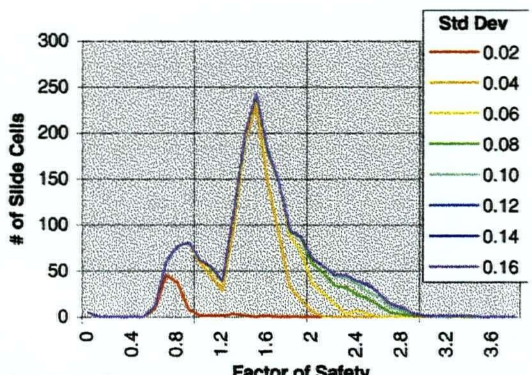


Figure 5.4. Factor-of-safety values for slide zones relative to number of cells. A series of standard deviation thresholds are used in separate curves.

In order to generate a summary of how standard deviation (SD) behaves relative to factor-of-safety, the SD values are thresholded at decreasing values and a histogram similar to Figure 5.2 generated (Figure 5.4). The largest SD is approximately 0.16, therefore, the 'SD=0.16' curve is virtually identical to the 'slide' curve in Figure 5.2. The 'SD=0.14' curve represents a histogram of all cells with a SD of 0.14 or less. This thresholding is continued down to the small SD=0.02 'bump' at the lower left (around FS=0.8).

Generally the graph shows that:

1. Some very dangerous areas are predicted with high prediction certainty, but there are not very many of these.
2. In the zone where most slides occur (FS 1.3-1.8; medium danger), they can only be predicted with intermediate rather than high certainty.
3. As predictions move towards the safer side of the series (> 1.6), the uncertainty in the prediction increases, demonstrated by the increasing spread on the graph.

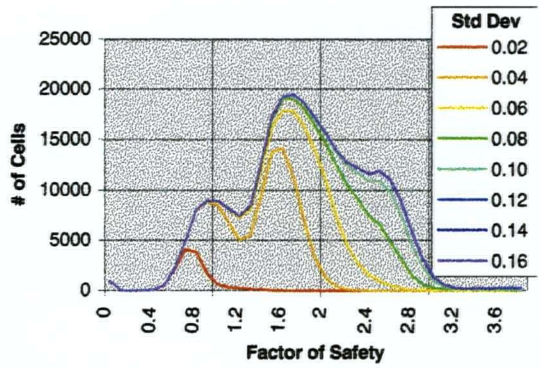


Figure 5.5. Factor-of-safety values for non-slide cells relative to number of cells. A series of standard deviation thresholds are used in separate curves.

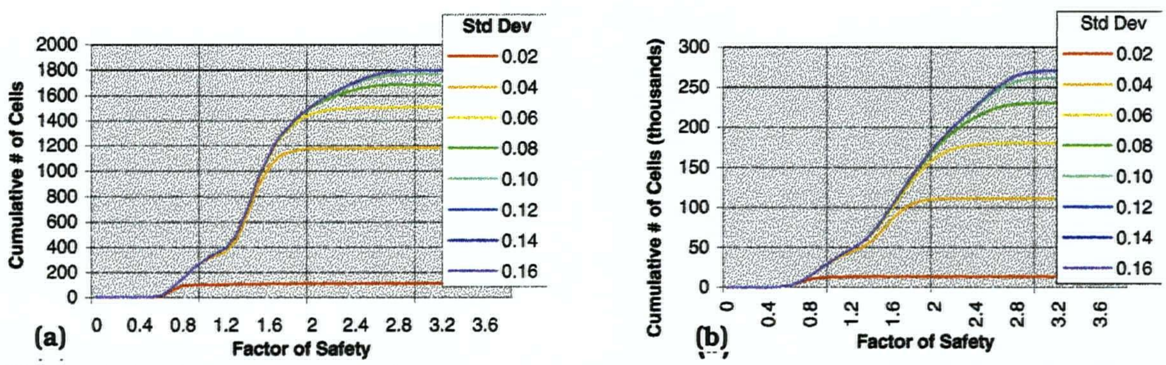


Figure 5.6. The previous two figures are here graphed using cumulative values. (a) details slide cells, while (b) shows the non-slide cells.

Viewing this same type of graph using the non-slide data (Figure 5.5), it is apparent that there is a different distribution of uncertainty in these data than exhibited by the slide data. The graphs are similar in dangerous areas; however, overall the non-slide demonstrates more uncertainty. Figure 5.6 highlights how the variance increases gradually for the non-slides towards the right of the curve (b), but quickly for the slide cells (a). These variations will be examined in further detail below.

5.3.4. Expected vs. Actual

The sections above have looked principally at descriptive statistical and graphic methods for examining differences between slide and non-slide areas. However, the evaluation of any type of model typically uses a comparison of expected vs. actual values. The problem is that, as introduced above, there is a *range* of expected values. High uncertainty and low predictive success must be considered as valuable (though of less practical use) than low uncertainty and high predictive success. For example, if one of the sources for the model introduced low resolution or low quality data, this would be represented in the results as high uncertainty and low predictive success. In this case the uncertainty model is highlighting a *data* problem rather than problems with the model itself. Even though the slide areas are poorly predicted, as an *uncertainty* model, the process is a success. Low predictive success with *low* uncertainty would indicate poor performance of the slope stability model (or some other factor such as calibration).

A second problem is that direct comparisons of slide areas and predictions are difficult, due to the necessity of juxtaposing binary (slide) data with cardinal predictions and their associated variability (enumerated in greater detail above). One possible way of addressing both of these problems is through determining what the expected distribution would be for both factor-of-safety and variance, and then graphing the expected zone. Success would be determined by graphing both the population and the slide cells and determining the percentage that fall into the 'expected' area.

This possible range of expected values is represented graphically as the area shaded darkly in Figure 5.7 using an FS variance vs. FS graph (the shading represents a general tendency rather than specific numeric values). However, a plot of the population (Figure 5.8a) shows that the population already exhibits this tendency (i.e., it falls roughly within the shaded area of Figure 5.7). Generally, high variance always is associated with safe areas, while low

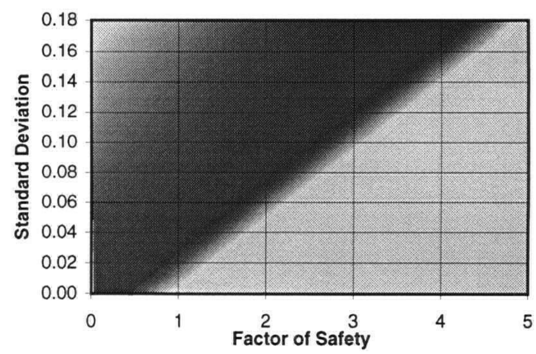


Figure 5.7. A scatter graph of variance vs. factor-of-safety for slide zones should fall within these general bounds if the predictive accuracy of the uncertainty model is good.

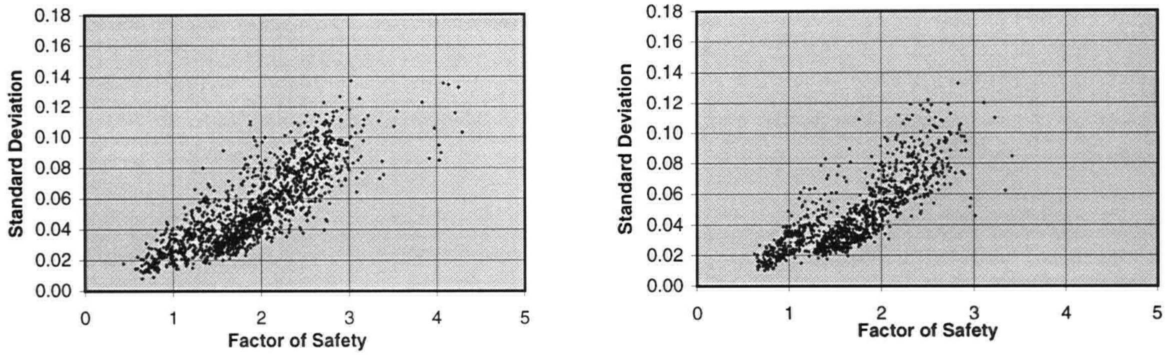


Figure 5.8 (a) Population (random subset) vs. standard deviation; (b) Slide cells (all) vs. standard deviation.

variance is associated with unsafe areas. Re-examining the standard deviation figure in Chapter Three (Figure 3.10), this association is graphically apparent. Highest variability exists on valley bottoms, most probably due to exaggerations of minor slope variations in the modelling process.

The fact that the population *already* exhibits the expected distribution makes an 'inside or outside the line' type of uncertainty model evaluation difficult. There are few cells that have a high FS *and* a low variance. This difficulty is compounded by a lack of specific numeric boundaries for the prediction region, making it impossible to compare slide predictions with the population on the basis of scatterplot shapes. (Note that the plot in Figure 5.8a displays a random subset of the population with a count equal to Figure 5.8b for the purpose of comparison and visual clarity)

Therefore, it is necessary to drop the stipulation noted above that 'high variability and high factor-of-safety equals predictive success.' It will instead be necessary to focus specifically on predictions of slides with low variability and low FS (i.e., on whether slide cells appear in the lower left side of the curve—the 'unsafe' area).

A comparison of the population and the slide cells (Figure 5.8) shows a slight shift in concentration to the left (decrease in mean FS in slide cells), and the disappearance of most safe outliers (far right). However, there is no substantial change in the concentration of cells. (The appearance of two apparent 'nuclei' or concentrations on the lower left are most probably remnants of the centroids in soil attribute space of the two most prevalent soil types). However, when the worst-case data are introduced (Figure 5.9), the degree of concentration in the lower left is further increased.

Predictive success has gone up (from the viewpoint of reducing variability). However, this is not the final comparison. There are still additional manipulations that are possible due to both the large amount of information in the uncertainty model and the spatially variable nature of that model.

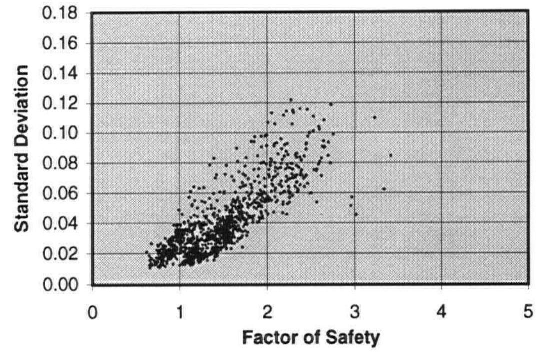


Figure 5.9. Worst case realisation factor-of-safety vs. standard deviation.

5.3.5. Zonal Spatial Limits

The WC realisation places constraints on the attributes in order to better focus on the overall goals of the model. *Spatial* constraints can also help this focus. Thus far in the analysis the comparisons have been between the predicted and actual values for the entire areas designated as mass wastage. However, within an individual slide there are different zones where different processes occur.

Of particular concern are the differences between the upper and lower areas of a slide. A certain percentage of the lower section of any slide can be considered the deposition zone. It consists of terrain features that are not conducive to continuing the slide, such as stable soils or a reduction in slope. This area is not technically part of the slide *process*, but certainly part of the disturbed region. The upper area of the slide is typically the initiation zone, and is more involved in the process of the landslide. Therefore, although the dividing line will vary, the upper slide areas should be predicted with greater accuracy than the lower areas, all else being equal.

As a generalisation, the lower 50% of each slide zone was removed from the 'slide' dataset. As a result, concentration in the lower left of the scatterplot substantially increases (Figure 5.10). This is reflected by a drop in the mean factor-of-safety value from 1.20 (WC, full slide) to 1.09 (WC, upper slide) in Table 5.1.

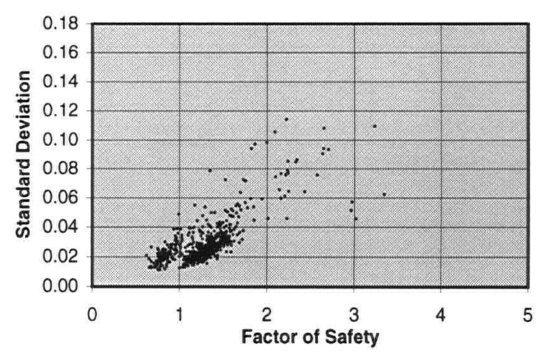


Figure 5.10. Worst case realisation using the upper 50% of slide zones.

With this latter refinement of predictions some of the minor variations become more visible. For example, in comparing the latter graph of upper slides with the maximum likelihood realisation, it is apparent that only one of the two main 'nuclei' shifts to the left. The WC realisation predicts the most failure-prone slide areas with equal probability to the ML realisation. Only the 'medium danger' predictions are shifted down (left).

The rough approximation of '50%' to divide slides is effective; however, there is likely a more effective division point that will further increase predictive accuracy. Therefore, all slide zone cells (observed) were coded based on their position in the local slide area. For example, a cell at elevation 100m on a slide ranging from 50m to 125m in elevation was assigned a 'position value' of 0.66 (66%). The slide cells were then divided into two groups based on the factor-of-safety predictions, making use of the graph peak (Figure 5.2, value of FS=1.6) as a dividing line between 'slide predicted' and 'no slide predicted' (note that, as mentioned earlier, due to the relative nature of the FS predictions this division is not necessarily ideal).

The 'position values' for the 'slide predicted' cells are graphed in Figure 5.11 (relative frequency). It is evident from the moving average overlay (thin line) that there is no clear single break-point on this graph to divide the slide zones by position. Of those that are evident, the small slope break at ~30% would remove 20% of the correctly classified cells from the set, while the break just above 50% would remove 38% of the slide predicted cells. From the data graphed here it is evident that any division in this general range will increase predictive accuracy, but at the expense of Type B errors (as discussed earlier).

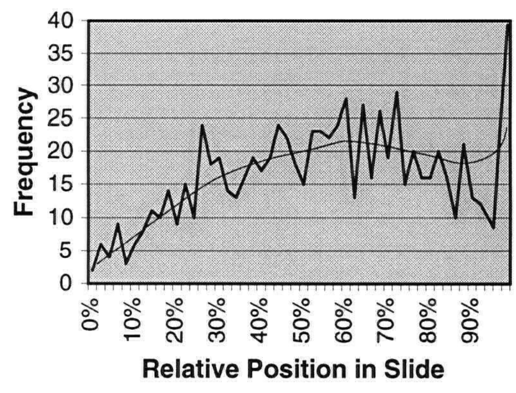


Figure 5.11. Relative frequency of the relative position in each slide for all correctly predicted cells (based on an FS=1.6 division). The thin line is a moving average of the raw data (thick line)

In order to further examine this issue a visual comparison of prediction accuracy and slide position is compiled (for three major slide areas) in Figure 5.12. In this figure it is evident that, generally, predictions fall in the upper sections of slides—particularly the large and/or long slides. However, medium and small slides would typically be either included

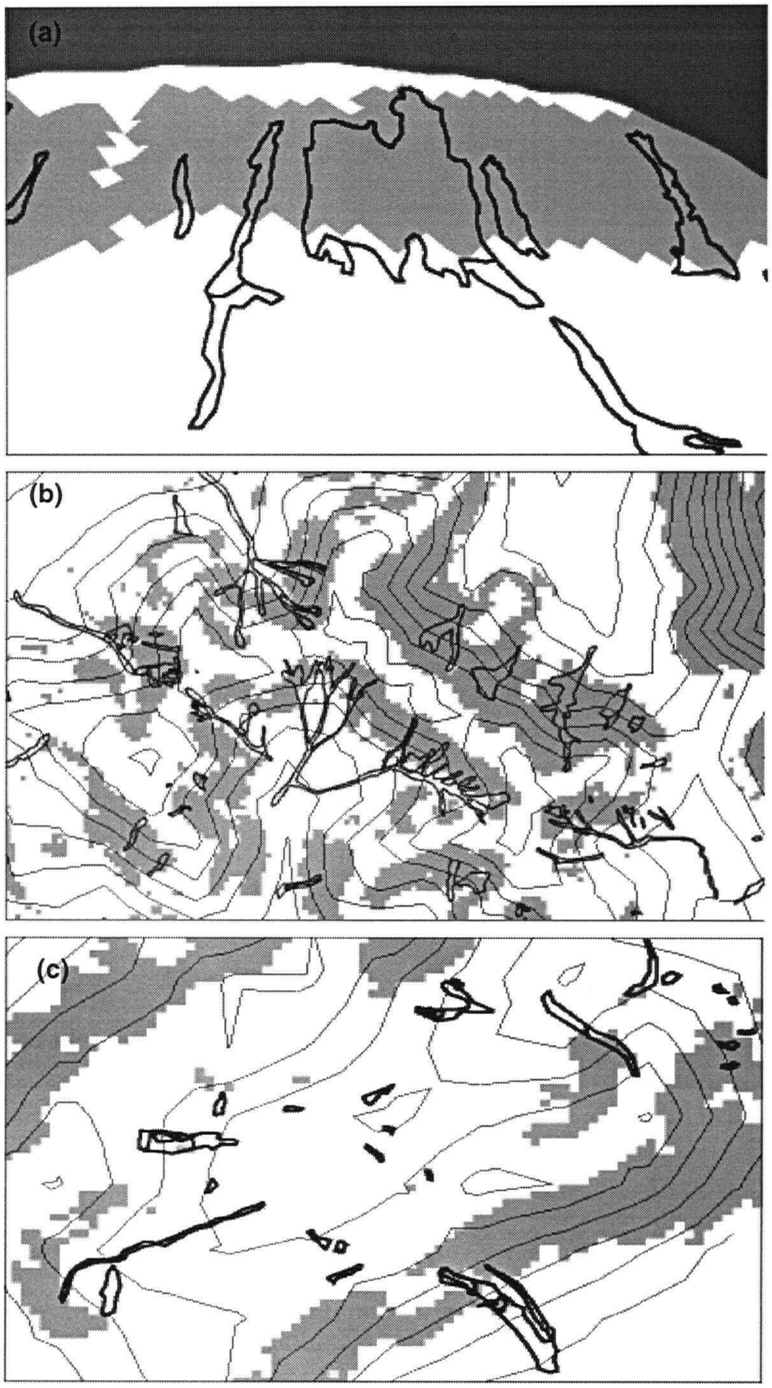


Figure 5.12. Position of low FS predicted areas (grey cells) relative to slides (dark lines) and their position and orientation on slopes (using 50m contours—thin lines). The 25m cell size provides a relative scale indicator. (a) is a draped perspective view of one side of the Gogit valley; (b) is a plan view of an intensive slide region in the centre of the island; (c) is a plan view of the Powrivco Valley (the initial test area).

or excluded from the predicted set. The strong influence of soil polygon divisions is also evident in this Boolean realisation of slide prediction, providing evidence indicating why the graph in Figure 5.11 shows no clear division. Note, however, that this binary division into 'predicted' and 'not predicted' is a somewhat arbitrary slice into a complex range of predictions. Extensions to this analysis might focus on how prediction behaves spatially as the division line slides up or down (i.e., the dotted line in Figure 5.2 moves left or right).

5.3.6. Spatial Constraints

Cell-by-cell predictions of slide zones violate the independent sample assumption required for most statistical significance tests. One method of overcoming this problem is subsampling the slide areas at regular intervals to reduce spatial dependence (based on a semivariogram sill). However, this dataset already is operating at the limit of its resolution. Many slides are composed of less than 4 pixels. Subsampling would effectively reduce the dataset, so that predictions would only focus on large slide areas—biasing the results.

Another option is to treat each individual slide as one event, and compile statistics based on this reduced summary dataset. Slide area FS values were reduced to one mean value per slide zone. For comparison, a series of areas of similar size to each slide were located randomly within the island's boundary. Statistics were compiled as with the slide zones. Although the Z test score was substantially reduced, even with a large decrease in N this still indicates a very high probability that the slides are not part of the background population (Table 5.1 above—last line).

5.3.7. Comparison: Old vs. New

A secondary purpose of this evaluation exercise is to determine if the changes made to the model in Chapter Four resulted in any increase in predictive accuracy. The following methodology is used:

1. Run the uncertainty model described in Chapter Three using the Lyell Island soil and forestry datasets (the soil dataset is described in Appendix E; the forestry dataset was generated using the 1990 orthophoto coverage).

2. Reset the uncertainty model parameters using the numbers determined by the allocation routines in Chapter Four. Generally, these routines determined that soil types one and three have a higher likelihood of misclassification with each other than the original parameters accounted for. Other minor changes were also made, but only in reference to the three most common soil types.
3. Repeat the model and compare results.

The graph in Figure 5.13 compares the two using a maximum likelihood realisation. It is apparent that changes in the model did not affect the slide zones to any significant degree. Exploratory spatial analysis indicates that the areas affected by the differences in the model (Figure 5.14 - shaded pixels) are not located in slide zones to any large degree. The shaded zones in this figure are located principally in bedrock zones (cross-hatched polygons), classified as Type six. This type is not present to any significant degree within mass wastage zones, where the dominant type(s) consist of sandy/silty colluvial or morainal blankets. Bedrock generates a very high factor-of-safety in the infinite slope stability model—so high that minor variations in model parameters can cause significant numerical differences between runs. However, these variations are all between very 'safe' values, and so are of little consequence.

Changes in the model (parameter and slope calibration) resulted in minor variations over the entire surface of the study area. Although these changes had little impact on the slide predictions, they may affect other applications that do not focus on soil cohesion or weight. Although exploration of the implications of this information would be best left to specialists in geomorphology or soil science, it is apparent that this *type* of comparison is potentially useful for exploring the details of environmental models. Here, although the model *may* now be more representative of reality, the effort expended to make it so did not translate into increased utility. This type of analysis is, in a sense, the spatial equivalent of a non-spatial sensitivity analysis (such as the type used by Hammond

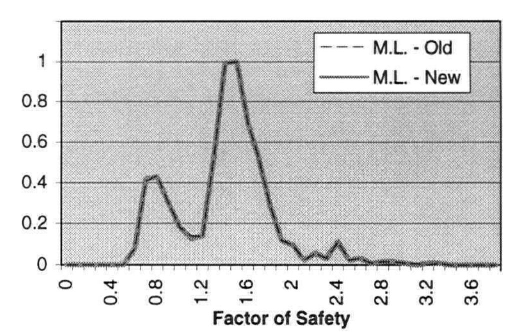


Figure 5.13. A comparison of predictive accuracy between the original and updated uncertainty models.

et al. 1992 to determine that slope and cohesion are the most sensitive components of the infinite slope stability model).

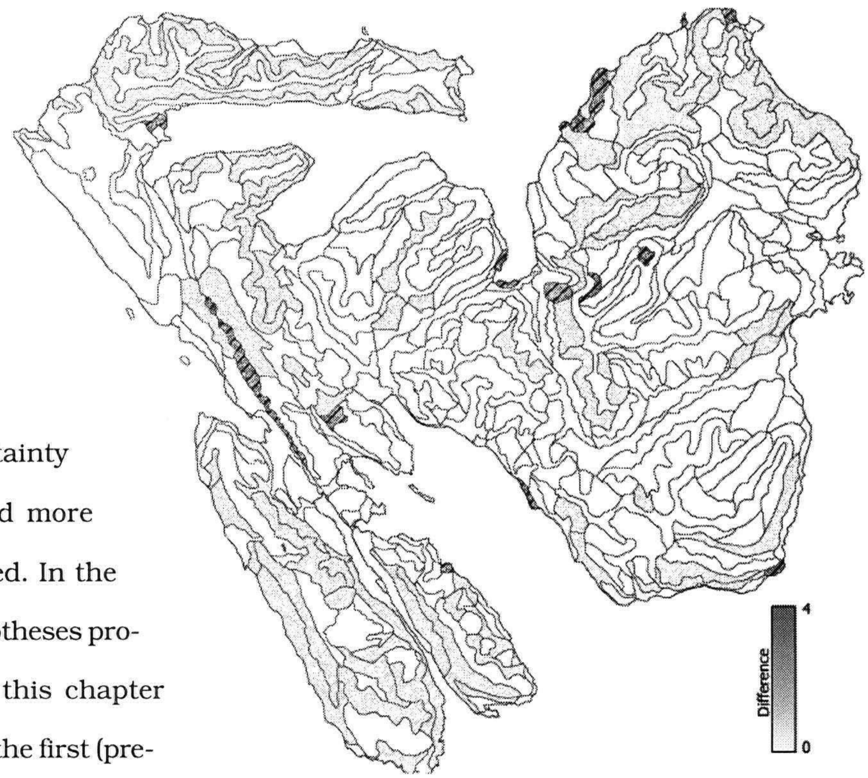


Figure 5.14. The location of differences between the two models relative to soil polygons. Significant differences are indicated by shade variations. Soil type six (bedrock) is highlighted with cross-hatching. Bedrock and significant variation tend to coincide.

5.4. Discussion

This evaluation of the uncertainty model has certainly generated more questions than it has answered. In the simplest case, the two null hypotheses proposed in the introduction to this chapter have been addressed. Namely, the first (predictions in slide zones are not significantly different than predictions for the population) was rejected, and the second (predictions made by the original version of the model are not significantly different than those made by updated version) was not rejected. The only general conclusions that can be drawn from these two items of information are that 1) the slope stability model works better than random assignments, and 2) calibration of the spatial constraints on the soil model input had no significant effect on predictions in this particular environment.

Delving deeper into the spatial and attribute structure of the model, it becomes apparent that realisations of the model other than maximum likelihood increase the prediction success. When 'success' is redefined using the Type A and B errors (as redefined above), realisations such as the 'worst-case scenario' can be used to concentrate the error within the less important of the two. However, it is difficult to compare realisations on other than a summary level, for each creates an entirely unique population.

Variance data allows some more specific conclusions to be drawn about prediction accuracy. For example, the most unsafe areas were predicted with highest certainty, while uncertainty in poor predictions (high FS in slide zones) increases with increasing FS. This generally confirms the expected performance of the model; however, the *numerical* significance of this correspondence cannot be directly ascertained.

A graphical analysis of actual vs. expected model results shows that the population data already exhibit the expected shape, eliminating some possible methods of analysis. However, the model, particularly in its spatially constrained realisations, has sufficient predictive success (slides vs. not) to allow evaluation to concentrate on this particular aspect, and ignore the possibility of low success at high levels of uncertainty (a secondary method of defining uncertainty model 'success'). The use of zonal spatial limits creates the greatest increase in slide prediction success, virtually eliminating high FS values from the slide cells.

One of the greatest difficulties with this type of analysis (as with any mass wastage modelling) is the problem of comparing the relative predictions of the model with the binary events of landslides. Most studies fall back on simple percentages and cut-off lines for evaluation methods (e.g., a cut-off FS of 1.7 classifies 70% of slides correctly). The methods used here offer more information that can be used to evaluate other models or alternative realisations of this model.

As for the new questions generated, there are numerous aspects of the data, the model and the underlying process(es) that can be illuminated through exploratory analysis. These might include:

- Working backwards by determining the realisation for each cell that gave best predictive accuracy, then determining why the certainty factor was not optimised. This would highlight problems with the soil database, the classification system or the model itself.
- The use of border spatial constraints, such as 'nibbling' the edges of the slide zones to reduce inaccuracies caused by data errors, such as vector-raster conversion, or due to physical process, such as the outside edges of slides being 'dragged along for the ride' by the main failure. Similarly, a higher resolution dataset could also reduce these problems, though at the cost of reducing the spatial extent of the area modelled, or increasing the processing requirements.

- Examining where the average deposition area begins on a slide, and increasing model accuracy in this way. By thresholding the model using different spatial constraints, variations in prediction accuracy could be used to determine the optimal way to represent a landslide in the database. Areas to investigate might include percentage area, slope effects, type of soil and nature of the deposition zone (slope effects have been noted by Fannin and Rollerson 1996).

5.5. Conclusions

This chapter has presented an evaluation of the slope stability uncertainty model developed and originally implemented on Louise Island. A high resolution database, detailing landslide timing and location on Lyell Island, was used to this end. For the most part, the evaluation focused on exploratory analysis of the data. Comparisons between various realisations of the model led to the conclusion that the worst-case-scenario version, coupled with spatial constraints limiting analysis to the upper 50% of slide zones, was the most effective for predicting slides with low variance. The second possible type of correct prediction—high FS with low variance—was not analysed due to the parameters of the population. A comparison of the model using its original parameters and a second run of the model with parameters updated through ground-truthing was performed. The changes did not appear to improve the model's predictive capability.

The development, testing and application of uncertainty modelling, as presented in the previous three chapters, is perhaps the simplest part of managing uncertainty in forest inventory. The crucial, most difficult step is widespread implementation of such tools in resource management. Certain aspects of management would require changes if uncertainty in data were to be recognised and addressed. The following chapter 'caps' much of the technical work by briefly pointing out the utility of this research in a management context.

Chapter Six

Discussion

6.1. Introduction

This dissertation has primarily been concerned with the development of techniques for storing, propagating, and in particular, verifying certain types of metadata. The work in this chapter focuses on the implications of these metadata, metadata manipulation techniques, and their verification in the realm of real-word natural resource management. The discussion will initially revisit the arguments used to justify this line of research, and then broaden out into the integration of uncertainty models into natural resource management, with a specific focus on forestry. The importance of uncertainty model verification will be highlighted throughout.

The broad discussion of uncertainty modelling in Chapter Two highlighted the fact that, in many natural resource applications, the traditional methods of analysis—based on a Boolean approach to data—are inadequate. A variety of uncertainty modelling methods are possible—one of which is demonstrated in Chapter Three. Much of the work in the remaining chapters has focused on verifying the inputs and outputs of this model, as well as the development of general techniques that allow this type of verification to occur in other uncertainty models. The final research question posed in the introduction is 'what are some of the implications [of these methods and techniques] for resource management decision making?' This is a rather open question, and has not been posed with the expectation of generating a complete answer. However, as one of the underlying themes of this research is the juxtaposition of uncertainty research with real-world data, the work would be incomplete without juxtaposing the general process of uncertainty modelling and verification with real-world decision making.

There is a need for the integration of uncertainty modelling, both spatial and non-spatial, with natural resource management decision making. The first part of this chapter will focus on justifying this statement. To begin, there are already several areas where basic types of uncertainty modelling exist in the decision making process. These include high level 'risk management' and lower level metadata. Considerable research has focused on the former, and the topic will only be addressed peripherally here. The latter is a rapidly growing area of resource data management (in fact, most spatial data management); however, this effort is primarily in the capture and storage stages, rather than in the 'how do we use this information?' stage. A discussion of this topic—the data management aspects of metadata—can be found in Chapter Two (§2.4). Here, the issue is the implications for management. The latter part of this chapter includes a discussion of other 'future research' directions highlighted or generated by the research in this document.

6.2. Resource Management

Although 'resource management' is a term that has been part of the discussion in most of the preceding chapters, here the term will be broken down into specific components. As yet, the term has not been explicitly defined. One simple, generic definition is "a series of deliberate interventions in system processes" (Iles 1994). Although this 'intervention' could include forcibly doing nothing (e.g., halting urban expansion to preserve certain habitat), more typically it involves a sequence of planned activities introduced at different points in space and time that focus on creating or maintaining a particular state of the resource (Iles 1994). The goal, or 'state of the resource', is usually defined externally (e.g., government policy, societal values). In any case, the 'activities'—the management strategy—are used to move towards this goal.

In delineating the types of problems that uncertainty management and verification might address, it will be necessary to partition the topic of management into three commonly accepted principal phases (or levels) often applied to both planning and management: strategic, tactical and operational. At each of these levels the problems, opportunities and constraints are somewhat different. The following sections will focus on the management of natural resources, with most examples drawn from the forestry sector.

6.2.1. Strategic Level

At the strategic level the concept of uncertainty management has received a considerable amount of attention. Most strategic management and planning occurs at a summary level, with little spatial specificity. At the strategic level, 'risk management' is quite a common component of decision support, usually in the form of identifying risks in various categories that are associated with various alternative decisions or plans (e.g. Marshall 1986, Pollard 1994). For example, the economic risk of a shift in harvesting strategies might be determined using models of uncertainty in the economy, uncertainties in public policy change, and various other inputs. Decisions might be made in order to minimise risk, or to minimise risk while maximising profitability or stability. In a more specific example, the expected value of forestry restoration programs is often calculated using the *expected value* approach to risk management, in which biomass additions and other values are calculated using the 'best available evidence' to assess relative probabilities of different decisions (Scarfe 1997). Conservative estimates are used to ensure that societal aversion to environmental risks are captured in the estimate. Rather than simply guessing at these conservative values (e.g. –3% per annum social discount rate; see Scarfe 1999), uncertainty analysis incorporated with risk management would provide this conservative approach with some actual variability data.

Uncertainty in the types of spatial models dealt with in this dissertation are rarely direct inputs to strategic level planning. Most planning at this level involves tabular data, such as summaries of inventory by region or by type (e.g., Traas 1994). However, uncertainty management at the spatially-specific (lower levels) is a crucial, and often ignored component of these summary data. Even when they have been estimated, it is often difficult to combine uncertainty in the various components into an overall statistic. A clear example is the calculations that lead to an allowable cut determination (AAC) for regions British Columbia. In these documents (e.g., BCMOF 1996), summary statistics are typically perturbed by a standard amount (± 10% in the example) and the calculations are repeated twice to set lower and upper bounds. Strategic decisions are then made based on the sensitivity of the model to each of the inputs *individually*. Uncertainty modelling and propagation have a clear role in a) identifying the actual variability in the figures, b) combining the various input variabilities through uncertainty propagation, and c) providing realistic estimates of the overall sensitivity of the models used.

Uncertainty models also have a place in long-term strategic forecasting. In attempts to determine the future state of a resource based on the present situation and a chosen strategic management strategy, uncertainty modelling can play a role (in addition to that discussed above) in deciding just how far to trust the model. One question is: when do the data become 'buried' in their own uncertainty?, and another is: what is the actual uncertainty in the long-term supplies of a resource? These are crucial strategic planning questions that properly verified uncertainty management models can help address.

6.2.2. Tactical Level

The tactical level of management or planning generally involves a short or medium term viewpoint, a greater detail of planning than the strategic level, usually more spatial specificity, and increased reliance on spatially-specific data. In the forest sector, tactical planning focuses on inventory and inventory updates, the allocation of resources, and various short-range (i.e., 2-5 year) management plans. Uncertainty management has the potential to be very useful at this level due to a) increased reliance on spatial data, b) increased variety of data required, and c) the emergence of multi-stakeholder decision-making systems.

As plans become more complex and begin to draw on spatially-specific information, spatial uncertainty management becomes increasingly important. For example, in forest inventory and inventory updates there is currently little information available regarding the spatial distribution of uncertainty. Updates are generally performed at fixed intervals and on large segments of the inventory (delineated by political or cartographic boundaries). The incorporation of a verified uncertainty model has the potential to allow identification of specific areas that have high uncertainty, leading to spatially-specific updates. For example, uncertainty might be relatively high in areas that were originally inventoried from smaller (than normal) scale photos, or in areas where growth is relatively fast. This type of update has the potential to decrease costs significantly (see Appendix D for further discussion of this topic).

At the tactical level a wider variety of data are required than at higher planning levels. For example, where strategic plans required a rough summary of the area covered by wetlands, tactical plans require knowledge of exact area values, and how these wetlands are distributed relative to other items. Similarly, tactical planning brings a number of different models together, such as growth and yield, slope stability, and regeneration models. Even if the uncertainty in each individual model is understood, what is generally lacking is the ability to bring these together at the landscape level. Verified uncertainty propagation techniques offer a way to perform this task.

Another relevant aspect of tactical level planning is the gradual shift from a relatively simple process with one agency, company or individual making decisions to complex multiple-stakeholder processes. Although this is not the case in all sectors of resource management, in those where multiple-stakeholder processes have been implemented the requirement to thorough justify planning decisions has increased substantially. These often-adversarial processes commonly involve different interest groups bringing *their* version of the resource data to the table, leading to arguments about data veracity (e.g., Basta 1990). The fact is, they may all actually be looking at the same data—with each group displaying a different tail of the variance curve. An understanding of data uncertainty, coupled with uncertainty models in which the actual distribution of that curve has been verified, could lead to a decrease in contention over this particular issue. Perhaps, in recognising the level of uncertainty in their data and models, such decision-makers (or committees) might implement greater levels of conservatism in their decisions.

This section has introduced several areas where uncertainty management might assist tactical level planning in resource management. There are also secondary issues—areas in which secondary effects of uncertainty management would widen the scope of planning. One example is the development of temporal models (discussed in greater detail in Appendix D). Inventories used on a tactical level time frame quite often are composed of different *versions* (e.g., 1995 version, 1999 version). It is difficult, often impossible, to perform efficient studies of broad scale change-over-time using such a system. This may be due to scale changes (greater/lesser detail), variation in interpretation (e.g., polygon boundaries redrawn) and other similar issues. Temporal-oriented data storage (such as the simple example in Appendix E, or more complex temporal models such as those discussed in Langran 1992) increases the potential for change-over-time modelling. However, uncertainty in these temporal objects is a crucial factor that needs to be addressed. For example, if 10m satellite data are added to a 1m airphoto-based temporal inventory, the level of uncertainty in the *new* information increases. If temporal databases do *not* track uncertainty using verified models, they stand to decrease their utility for analysis.

6.2.3. Operational Level

At the operational level of management and planning, spatially specific uncertainty management has a decided, though often different, role to play. Much of this role hinges on the fact that resource sectors such as forestry currently have smaller profit margins than they have had historically and, therefore, decreased room for error in operational planning. One example is the decreasing supply of old-growth timber, leading to operations in forests of marginal profitability. Uncertainty modelling can potentially support operational planning in several areas: basic data gathering, modelling at fine scales, and highly specific inventories.

Uncertainty modelling has the potential to assist operational level data gathering in a manner similar to that described in the previous section, although with a focus on large-scale data. By pointing out specific areas of high uncertainty (such as the polygon boundaries highlighted in the worst case scenario model in Chapter Three), operational planning can focus on gathering additional data only in areas of high uncertainty.

An example of the role that uncertainty modelling can play in fine scale operational modelling is found in Chapter Three. Knowledge of uncertainty in slope stability might lead to different decisions regarding road placement or the timing and location of harvesting.

Some forest resource management sectors are currently in the process of shifting to an operational planning level *lower* than the accustomed stand-level management. In the usual stand-level planning, it was assumed that errors tended to average out—offset each other when a number of stands were harvested. In the new, fine-scaled operational regime, operational plans that focus on selective removal will require a greater understanding of what actually constitutes a stand of trees. An understanding of uncertainty at the largest scale of inventory will be of increasing importance. This also applies to operational plan development. An increasing amount of planning and documentation is required in today's highly regulated resource sectors. It is crucial to the success of *future* plans that today's plans be correct in as many aspects as possible. If, for example, a visibility analysis indicates that the altered terrain will constitute 11% of the landscape, yet the actual visibility is 18%, future plans may be looked on with some scepticism. With terrain uncertainty and inventory uncertainty modelling, increasingly realistic operational planning estimates can be developed. Uncertainty management in general is of increasing importance in a leaner, information-rich operational management environment in all resource sectors.

6.3. Uncertainty Model Validation

Uncertainty modelling can potentially address many planning and management issues at strategic, tactical and operational levels. In addition, the validation of these models increases their importance in many ways.

At the strategic level, risk management is gradually becoming an indispensable part of decision-making. However, there is a clear need for validation of the risks associated with different actions. This will not be an easy task, but it remains a crucial one. At both the strategic and tactical levels there is a need to evaluate decisions made on the basis of uncertainty information; not simply risks, but specific plans (such as allowable cut). The principal impediment is the difficulty in evaluating decisions made in different information environments without both (or all) decisions being made in comparative (e.g., double-blind) experiments. Without such experiments it is difficult to tell what *might* have happened.

At operational levels validation is also important. Currently, the '15.72157 ha of mature timber' type of answers that result from GIS-based analysis and modelling are mistrusted—with good reason. GIS and associated spatial (and non-spatial) models must *prove* that their estimates are sound before they will be accepted at the operational level. For example, knowing that the level of harvestable timber is as graphed in Figure 6.1 might make for quite different operational planning than simply knowing the central number. The validity of those outside figures is as important as the central one.

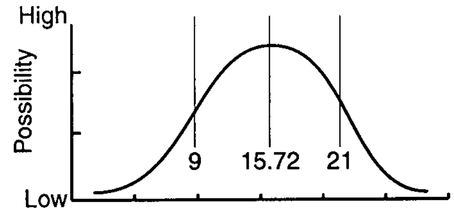


Figure 6.1. Typical graph of normally distributed uncertainty

Through tests of expert opinion estimates of risk or uncertainty, we end up with both better models and, through feedback, better understanding of uncertainty. This benefit cannot easily be measured. However, when planners, managers or scholars are forced to revise their estimates of the quality of the data they work with every day, they may also be forced to revise their methods, plans and research tactics. In the same way, *confirmation* of their quality estimates may also have positive benefits.

6.4. Further Research

The sections above have each included several recommendations for further research directions. Each of the chapters have also included specific suggestions for areas in which further work might enhance understanding of uncertainty, or lead off into separate research streams. This section summarises these points and indicates relative priorities of the suggestions, with the presentation progressing from general to specific items.

Research Programs. In reviewing the research literature in the general area of uncertainty management of spatial data (Chapter Two), it soon becomes apparent that this field particularly lacks research *programs*. Many laudable individual projects exist; however, the diverse range of applications (as indicated in documents such as the *Proceedings of the International Symposium on Spatial Accuracy of Natural Resource DataBases*, Congalton 1994) are rarely tied together in integrative programs. Verification of uncertainty metadata and models is a crucial element of such a program, and it is therefore hoped that the research presented in this document will assist in the development of such areas of study.

Integration into Management. As noted in sections above, there are a number of specific areas where properly verified uncertainty models and uncertainty propagation routines would be of use in real-world management. Potentially useful areas of research include the verification of the probabilities and possibilities associated with risk management scenarios through tests where various outcomes of a decision are followed through to their conclusion. At tactical or operational levels of management, uncertainty verification work is required on a more task-specific basis, with the goal of allowing a manager to *expect* to encounter a specific level of uncertainty, enabling her to put detailed contingency plans in place to deal with all likely eventualities.

Issues of uncertainty communication noted in Chapters Two and Three also have considerable impact on the integration of uncertainty modelling in natural resource management. When visualisation tools are developed (e.g., Fisher 1991b, Goodchild *et al.* 1994), this process usually ceases at the demonstration stage. Verification research is required to determine a) if what the visualisation routines/tools indicate is correct, and b) that the *impression* of uncertainty levels provided by these tools to non-technical users is in line with the actual level of uncertainty. Work such as this is underway (e.g., Antle, in prep.); further efforts are required in various resource sectors and tasks. A shift in 'spatial understanding' regarding uncertainty can only be judged through its effects on policy, resource decisions, scientific hypothesis generation or other bottom-line items.

The integration of verified uncertainty models into real-world management is undoubtedly the most crucial research issue to be discussed in this section. Using this general issue as a target, an efficient research program can thereby be planned and executed, tying in the various tasks of modelling, verifying and decision-support integration. Initial test cases such as the projects discussed in Davis (1994) and in Appendix E are needed to make decision-makers aware of the issues and importance of this work. However, concerted, application-specific research is required in order to develop applications for real-world, day-to-day management.

Uncertainty Model Input Verification. The research discussed in Chapter Four focused on verifying several of the inputs to a specific uncertainty model, based on a more general approach to verifying uncertainty models of classified and data. There are several possible research avenues that derive from this and the work on continuous data, including:

- Performing exhaustive sampling in limited areas in order to verify (and therefore better define) the behaviour of class uncertainty in attribute space.
- Performing comparative studies of slope representation in terrain models in order to determine the probable reasons for the noted variations from estimated uncertainty. It was noted in Chapter Four that, given the available data, it would be difficult to determine whether the extreme terrain, the elevation data contractors, the slope modelling routine assumptions, the GIS algorithm, or some other factor (or a combination) is to blame for the discrepancy. A comparative study might be undertaken using one or more of the following: increased intensity of slope sampling, a variety of comparative areas in different locations and with different types of terrain and ground cover, a comparison of photogrammetric techniques, and GIS algorithm evaluation in variable terrain. Although some studies exist that compare various types of elevation models (e.g., Sasowsky *et al.* 1992) or GIS techniques (e.g., Skidmore 1989), there is a lack of detailed ground evaluation studies, which has led to a potentially 'unhealthy' reliance on inventories such as TRIM.

Uncertainty Model Output Evaluation. In Chapter Five the uncertainty model was evaluated using the Lyell Island data. A number of new questions were generated that might be answered through exploratory analysis or with a more extensive dataset. These include:

- Working backwards by determining the realisation for each cell that gave best predictive accuracy, then determining why the certainty factor was not optimised. This would highlight problems with the soil database, the classification system and/or the model itself. This would require a more extensive dataset than is currently available for the island in order to bring secondary factors into play.

- The use of border spatial constraints, such as 'nibbling' the edges of the slide zones to reduce inaccuracies caused by vector-raster conversion. Similarly, a higher resolution dataset could also reduce these problems, though at the cost of reducing the spatial extent of the area modelled, or increasing the processing requirements.
- Examining where the average deposition area begins on a slide, and increasing model accuracy in this way. By thresholding the model using different spatial constraints, variations in prediction accuracy could be used to determine the optimal way to represent a landslide in the database. Areas to investigate might include percentage area, slope effects, type of soil and nature of the deposition zone. Once again, a higher resolution and more detailed dataset would be required for accurate determination of these factors. Some work on this topic has been conducted by Fannin and Rollerson (1996), who noted that deposition is generally triggered by a distinct change in slope gradient.

Other Areas. A number of the techniques developed during the mass wastage database production and testing (Appendices D and E) led to possibilities for further specific research. These include:

- The 'crossings index' was presented as a rough indicator of mis-registration under certain circumstances. In the text of Appendix D it was noted that a possible extension study would further explore the implications of line crossings relative to scale, digitising accuracy, and other factors. It was also noted that a more intensive study would be required to determine what the *absolute* implications of the crossings index number are (rather than the relative implications explored in the text).
- In Appendix D an attempt was made at automating the image registration procedure. It was noted that new types of search algorithms would be required to complete this procedure in reasonable time, quite possibly requiring some degree of understanding of the human vision

process. A logical extension of this work would be a study of the visual clues used by a manual system operator to register images in a small number of steps, and to translate this into a computer procedure—possibly using expert systems or other learning algorithms.

- A direct comparison between different terrain types focusing on similar targets would enable quantification of the apparent increase in accuracy of the ODFS registration in extreme terrain noted in Appendix D.

Communication of Uncertainty. Uncertainty models, such as the model presented and verified earlier, are typically complex assemblages of data. Communicating these data to both analysts and decision-makers presents a challenge in the area of data visualisation. No standard, proven methods for display of uncertainty data exist. A number of techniques have been discussed in the literature; however, few implementations exist, and the majority of these refer to artificial or sample datasets. Exploratory visualisation of practical datasets focusing on real-world problems is required to further develop this research field.

More specifically; an understanding of the implications of the results of an uncertainty modelling procedure cannot easily take place without visualising uncertainty measures in concert with the original data. For example, slope stability can be easily classified and displayed. However, there are many unknowns regarding how uncertainty in these data can be effectively communicated to a user accustomed to seeing crisp, Boolean-style data. Considerable research is needed in this area.

More specifically, the numerous dimensions involved in a multiple uncertainty representation place an increased burden on the spatial analyst. The database is far more flexible than in the Boolean case, but flexibility is coupled with complexity. Uncertainty and error can be combined at the summary/display stage, but only in an environment where the users' needs are thoroughly understood. The wide variety of possible representations allows the database to provide just about any answer desired. Therefore, considerable work is required to determine just what 'reasonable' uncertainty/error values are and how these translate into reality in the field. At this point the concepts of risk analysis and acceptable risk come into play. Once the display has been calibrated with field data (e.g., green: safe, yellow: a reasonable possibility of failure, and red: the near certainty of failure) using techniques such as those discussed in this document, it becomes possible for the user to set the desired risk level (e.g., 5% chance of being wrong) and proceed with an analysis. The term 'risk' implies that there are social, economic, or other factors interacting with the spatial uncertainty metadata in a decision-making context. This concept of 'acceptable risk level' may be easier for most users to interpret than the quantified uncertainty values used as internal representations in the database.

The added dimensions available through the use of dynamic visualisation tools also place an increased burden on the cartographer. The purpose here is not simply effective communication of a particular message. The 'message' imparted through visualising uncertainty information is far less tangible, and therefore far more difficult to evaluate. A shift in 'spatial understanding' regarding uncertainty can only be judged through its effects on policy, resource decisions, scientific hypothesis generation or other bottom-line items.

In addition to the issues discussed above, visualisation tool development would benefit from user evaluations—not simply regarding cartographic communication, but through a simulated decision-making scenario. The effectiveness of these tools can only be properly judged through a cross-comparison of decisions made based on different techniques, as well as comparisons with a control group using static, Boolean-based maps (for example see Antle, in preparation). It is likely that experience and innate understanding of uncertainty are already incorporated into many types of Boolean-based decisions. It will, however, be difficult to make predictions outside of particular application areas.

Comments on Further Research/Implementation. Although this work has focused specifically on a slope stability model implementation of uncertainty modelling, many of the techniques developed and implemented are also highly applicable to other types of natural resource modelling. First, the uncertainty model itself is generic in nature, in that it can encompass a wide variety of *types* of uncertainty, and acts as a shell around the process model. The basic requirements are: a good understanding of the nature of uncertainty in each of the model inputs, and the ability to interpret the extensive output of the procedure. In essence, the entire uncertainty modelling procedure forces the investigator to develop a thorough and complete understanding of the data she works with.

This research has indicated that the process of gathering expert opinion can be fraught with error, and that published error statistics are not necessarily trustworthy. The effort required to determine the actual values for uncertainty must be balanced with the depth of analysis required for the application at hand. There are many other possible ways of gathering this information; however, they too should be viewed with some suspicion if a precise accounting of uncertainty is necessary.

Overall, it is important to identify and then concentrate on the specific types of uncertainty that affect the resource model to the greatest extent. Some of this information might be gained from a sensitivity analysis, while others will simply be common sense. However, because of the potentially multiplicative nature of uncertainty, the other types should at least be estimated and included wherever possible.

6.5. Summary

The discussion in this chapter has touched on a number of resource management areas where uncertainty management, and particularly uncertainty model validation, may be of use in solving problems or increasing efficiency. The issues are different at different planning or management levels; therefore, the discussion was partitioned into the three main levels of planning: strategic, tactical and operational.

At each level, some of the more obvious implications of incorporating verified, spatially variable uncertainty models were discussed. As with the remainder of this document, forestry was utilised as the principal example of resource management. This sector's heavy reliance on data and models of a highly spatially variable resource with numerous associated uncertainties makes it a prime example. No doubt, in other sectors there are many other implications of data uncertainty and verification for resource management. It is hoped that these examples will bring some of the relevant issues to the forefront.

Chapter Seven

Conclusions

7.1. Summary of Study

The new research discussed in this document and its appendices took place over the span of four years. The major tasks can be broken down as follows:

Louise Island - Model Input Verification: Approximately six weeks were spent in the field gathering the data for this phase. Two of those weeks were used to gather existing data from various agencies in the Queen Charlotte Islands, two were spent on Louise Island and two on Lyell Island performing preliminary work for the next phase. In the lab, two months work went into developing the systems and performing the preliminary analysis; another four weeks were spent developing the visualisation routines for analysis and reporting. Development of the conceptual work took place over an extended period.

Lyell Island - Model Output Verification: Lyell Island is located approximately 100km from the nearest roads, and 200km from the nearest fuel supplies. Access required 2-10 hours of boat travel (weather dependent). A single water circumnavigation of the island required 3-4 hours in good weather. Therefore, much of the field effort involved in this research focused on logistics. During the second field season three weeks were spent on Lyell (author and assistant). All travel on the island was on decommissioned roads (by foot). Survey equipment was carried to fourteen of the major landslides, and the slides were physically surveyed. A typical survey of a single slide required 5-6 hours of hiking and 2-3 hours actually on the slide. Extremely remote slides were accessed via zodiac landings on highly exposed beaches.

The aerial survey work described in Appendices D and E required a single day of effort, and approximately one week of logistics and preparation. The system was also tested in a separate area (in work that is not described herein), with another two weeks of field time required.

Development and analysis of the Lyell data took place over two years. Approximately four months of full-time work (two months with an assistant) went into developing the orthophotos and the baseline database of mass wastage. Another three months were required for code development of the ODFS system described in Appendix D. Approximately two months were required to perform the rehabilitation case study described in Appendix E. As with the Louise analysis, the final analysis and conceptual work took place over an extended period.

7.2. Research Questions

The research presented in this dissertation has focused on the issue of uncertainty model verification. Specifically, the central research question was: **can a natural resource management uncertainty model be verified in order to evaluate its utility in real-world management?** In the introduction it was noted that there can be no simple yes or no solution, as there exists no simple statistic to determine if uncertainty as modelled equals uncertainty as sampled. As the research has shown, the issue of uncertainty model verification is a complex one; yet, through techniques such as exploratory data analysis, it has been possible to address the principal research question.

The question was addressed by breaking it down into a series of manageable questions, each of which focuses on one of the 'verification boxes' in Figure 1.1. The questions are as follows.

1. What are appropriate methods for modelling data uncertainty in natural resource management, making use of information typically available?

This question was addressed in previous research, summarised in Chapter Three, and used as a model base for the following chapters. It was noted that there are several types of uncertainty that must remain conceptually separate, but may be brought together in data summaries and queries of the resulting uncertainty model. A number of methods were discussed, and fuzzy sets were chosen for the test case (but only for particular types of uncertainty).

2. How appropriate are these methods, and how can this 'appropriateness' be determined? Specific questions include:

   2a. How effective is gathering metadata from expert opinion?

   2b. How effective is gathering metadata from published variability statistics?

The effectiveness of these two inputs to an uncertainty model was determined through ground verification of the modelled information. Methods were developed to allow comparison of sampled values with classification uncertainty levels, allowing the 'appropriateness' of the model to be determined. This development was the principal focus of Chapter Four, and was used in a test case to verify the model described in Chapter Three. Metadata gathered using expert opinion on soil uncertainty were found to underestimate uncertainty in all cases, with specific soil types exhibiting greater uncertainty than others. Therefore, it was concluded that expert opinion is not necessarily an ideal input to this particular model, and should be looked on with some trepidation in similar exercises. The results pointed to the apparent fact that soil scientists do not necessarily have a strong grasp of the overall level of uncertainty in the data they regularly employ. Uncertainty verification was shown to have considerable importance in 'tuning' this model input.

Tests on the effectiveness of gathering data from published variability statistics also showed that, in the most important input to the test case model (slope stability), the published values underestimated the variability found in reality. However, the tests also indicated that, in this case, corrections could be applied that would reduce this margin.

The question of 'appropriateness' is not one that can be answered with a direct yes or no. The methods of modelling uncertainty did not, in their initial 'laboratory value' state, effectively reflect uncertainty on the ground. However, even prior to verification, they still—by definition—reflected the ground condition *better* than a standard Boolean model. The focus of the research was on the latter half of Question Two: 'how can this appropriateness be determined?' The methods developed are applicable to a wide range of uncertainty models and data sources, and will allow the appropriateness of other types of data and models to be determined and compared.

3. What are appropriate methods for propagating these metadata through to information products (i.e., using a typical type of natural resource model)?
4. How appropriate are these methods, and how can this 'appropriateness' be determined?

The first question was addressed in Chapter Three, where a combination of techniques (fuzzy joint membership function and Monte Carlo simulation) were utilised to propagate dissimilar types of uncertainty through a typical model. The focus of this research was addressing the fourth question through the development of techniques to determine the appropriateness of this uncertainty propagation. Again, methods were proposed and developed that would be applicable in a variety of situations. Here, they were tested using the output of the slope stability uncertainty model. The principal issue was addressing the fact that the model predictions incorporate multiple realisations and variability data, while the verification data were back or white. The results showed that the uncertainty model and propagation methods were highly appropriate in developing information products for a typical natural resource model. The information retained in the uncertainty modelling system allowed the slope stability model predictions to be of greater use in predicting slope failure with high certainty than the typical Boolean model. The fact that the uncertainty level in the Boolean model (in a typical application) is an *unknown* value serves to highlight the utility of uncertainty modelling procedures in general.

In general, the methods used to answer Questions Two through Four also demonstrate appropriateness through greater flexibility. The uncertainty modelling and propagation techniques, when properly verified, serve to increase the number of questions that can be asked of the data—both the model source and results. This could lead to either increased utility for research, or wider practical applicability of data and models.

5. What are some of the implications of the methods outlined in the above questions for resource management decision making?

The final question is a more general one, and was answered in the previous chapter. The implications are many and varied, and were broken down into three levels of planning. The implications of uncertainty modelling and verification at the strategic level are primarily non-spatial, and revolve principally around the task of 'risk management'. Verified uncertainty models would allow the various possible outcomes of a decision model to be assigned metadata, increasing decision model utility. At the tactical and operational levels the spatial variability of uncertainty becomes more important. Here, one of the crucial implications uncertainty modelling is in understanding the data used for planning, possibly allowing planners to understand that, in some cases, many of the presented scenarios fall within the range of possibility. Verified uncertainty models may also allow planning for leeway in strategic and operational plans.

Finally, returning to the overall question, it is apparent from the above research that a natural resource uncertainty model ***can*** be verified in order to determine its utility in real world management and, furthermore, one of the principal utilities of such a model is to allow greater flexibility and understanding of datasets and resource models by real world manager. However, this increased flexibility is combined with increased complexity. Communication becomes a crucial tool in bringing uncertainty models to the desktop of real world managers.

The research presented in this document represents a major step in an overall research program intended to integrate uncertainty management into natural resource decision making. Uncertainty models, uncertainty model verification, and resource model uncertainty verification feed directly into management integration. With another piece added to the puzzle, it is hoped that research into the process of integration will continue to grow.

Bibliography

ABEL, D.J., YAP, S.K., WALKER, G., CAMERON, M.A., AND ACKLAND, R.G. 1992. Support in spatial information systems for unstructured problem-solving. In *Proceedings of the Fifth International Symposium on Spatial Data Handling.*

ADKINS, K.F., 1994. Accuracy assessment of elevation data sets using GPS. *Photogrammetric Engineering and Remote Sensing,* 60(2):195-202.

ALONSO, W., 1968. Predicting best with imperfect data. *Journal of the American Institute of Planners,* July:248-55.

ALTMAN, D., 1994. Fuzzy set theoretic approaches for handling imprecision in spatial analysis. *International Journal of Geographical Information Systems,* 8(3):271-289.

ANTLE, A., in preparation, Interactive visualisation tools for data and metadata, Ph.D. Dissertation, University of British Columbia Department of Geography.

ARONOFF, S., 1982. Classification accuracy: a user approach. *Photogrammetric Engineering & Remote Sensing,* 48(8):1299-1307.

AUNG, H., 1992. Probabilistic analysis of earthen slope stability (slope failure). Ph.D. Dissertation, Marquette University. 159pp.

BANAI, R., 1993. Fuzziness in geographic information systems: contributions from the analytic hierarchy process. *International Journal of Geographical Information Systems,* 7(4):315-329.

BASTA, D., 1990. Information technology for multiple use decision making. In *Coastal Ocean Space Utilization,* edited by R.B. Abel and S.D. Halsey (Elsevier Science Publishing Co.), pp. 309-313.

BC MINISTRY OF ENVIRONMENT, 1988. Terrain Classification System for British Columbia (revised edition). (Victoria, BC: Gov't of BC).

BC MINISTRY OF FORESTS, 1996. Queen Charlotte Islands Timber Supply Area Rationale for (AAC) Determination. (Victoria, BC: Queen's Printer), 44 p.

BEARD, M.K. AND BUTTENFIELD, B.P. (eds), 1991. NCGIA Research Initiative 7: Visualization of spatial data quality. NCGIA Technical Paper 91-26, (NCGIA) .

BERTIN, J., 1985. *Graphical Semiology* (Madison, Wisconsin: University of Wisconsin Press).

BEZDEK, J.C., 1974. Numerical taxonomy with fuzzy sets. *Journal of Mathematical Biology,* 34:1-15.

BEZDEK, J.C., 1992. Fuzzy models: what are they, and why? *IEEE Transactions on Fuzzy Systems,* 1(1):1-5.

BEZDEK, J.C., EHRLICK, R., AND FULL, W., 1984. FCM: the fuzzy c-means clustering algorithm. *Computers in Geoscience,* 10:191-203.

BILODEAU, J.-M., BEDARD, Y., AND LOWELL, K.E., 1993. Using a Geographic Information System to plan a forest inventory that respects the spatial distribution of forest strata. *Northern Journal of Applied Forestry,* 10(4):161.

BLAKEMORE, M., 1983. Generalization and error in spatial databases. *Cartographica*, 21(2/3):131-139.

BOUILLE, F., 1992. Fuzzy neural processing by an object-oriented expert system: application to geographic information systems. In *International Fuzzy Engineering Symposium*, Fuzzy Engineering Toward Human Friendly Systems. Yokohama, Japan, Nov 1991.(Tokyo: IOS Press), pp. 574-585.

BRANDT, J.W., 1994. Convergence and continuity criteria for discrete approximations of the continuous planar skeleton. *Computer Vision,Graphics and Image Processing*, 59(1): 116-125.

BRIMICOMBE, A.J., 1993. Combining positional and attribute uncertainty using fuzzy expectation in a GIS. In *Proceedings, GIS/LIS Annual Conference.* Minneapolis, Minnesota, November 2-4.(Bethesda, Maryland: American Society for Photogrammetry and Remote Sensing), pp. 73-81.

BROWN, D.G. AND BARA, T.J., 1994. Recognition and reduction of systematic error in elevation and derivative surfaces from 7.5 minute DEMs. *Photogrammetric Engineering and Remote Sensing*, 60(2):189-194.

BRUNEAU, P. AND GASCUEL-ODOUX, C., 1995. Sensitivity to space and time resolution of a hydrological model using digital elevation data. *Hydrological Processes*, 9:69-81.

BUCKLES, B. P. AND PETRY, F. E., 1985. Generalized database and information systems. Unpublished Manuscript.

BUITEN, H.J. AND VAN PUTTEN, B., 1997. Quality assessment of remote sensing image registration - analysis and testing of control point residuals. *ISPRS Journal Of Photogrammetry And Remote Sensing*, 52(2):57-73.

BURROUGH, P.A., 1986a. Five reasons why geographical information systems are not being used efficiently for land resource assessment. In *Auto-Carto London*, edited by M. Blakemore. London, 14-19 Sept.(AutoCarto London), pp. 139-148.

BURROUGH, P.A., 1986b. *Principles Of Geographical Information Systems For Land Resources Assessment* (Oxford: Clarendon).

BURROUGH, P.A., 1989. Fuzzy mathematical methods for soil survey and land evaluation. *Journal of Soil Science*, 40:477-492.

BURROUGH, P.A., 1991. Soil Information Systems. In *Geographic Information Systems: Principles and Applications*, *2*, edited by D.J. Maguire, M.F. Goodchild, and D.W. Rhind (Harlow, Essex, England: Longman Scientific), pp. 153-168.

BURROUGH, P.A., 1993. Spatial data quality and error analysis issues: GIS functions and environmental modeling. In *GIS and Environmental Modelling Conference.* Colorado. Manuscript.

BURROUGH, P.A., MACMILLAN, R.A., AND VANDEURSEN, W., 1992. Fuzzy classification methods for determining land suitability from soil profile observations and topography. *Journal of Soil Science*, 43:193-210.

BUTTENFIELD, B.P., 1991. Visualizing cartographic metadata. In *NCGIA Research Initiative 7: Visualization of Spatial Data Quality*, edited by M.K. Beard and B.P. Buttenfield (NCGIA), p. C17-C24.

BUTTENFIELD, B. AND BEARD, M.K., 1994. Graphical and geographical components of data quality. In *Visualization in Geographical Information Systems*, edited by H.M. Hearnshaw and D.J. Unwin (New York: John Wiley and Sons), pp. 150-157.

BUZUG, T.M., WEESE, J., FASSNACHT, C., AND LORENZ, C., 1997. Image registration: convex weighting functions for histogram-based similarity measures. *Lecture Notes in Computer Science,* 1205:203-225.

CALIFORNIA DEPT. OF FORESTRY. 1992. Implications of sample size and sampling methodology for forestland map accuracy assessments. Manuscript.

CAPRA, A., NICOSIA, O., AND SCICOLONE, B., 1995. Application of fuzzy set theory to drainage and salinity problems. *Rivista Di Ingegneria Agraria,* 26(4):200-207.

CHEN, Z.Z. AND FINN, J.T., 1994. The estimation of digitizing error and its propagation with possible application to habitat mapping. In *International Symposium on Spatial Accuracy of Natural Resource DataBases: Unlocking the Puzzle,* edited by R.G. Congalton. Williamsburg, Virginia, May 16 - 20 1994.(Bethesda, Maryland: American Society for Photogrammetry and Remote Sensing), pp. 57-66.

CHORAS, R.S., 1993. Image codong by morphological skeleton transformation. *Lecture Notes in Computer Science,* 71: 216-222.

CHRISMAN, N.R., 1982. A theory of cartographic error and its measurement in digital databases. In *Proceedings, Fifth International Symposium on Computer-Assisted Cartography. AUTO-CARTO 8.* Falls Church, Virginia(ASPRS & ACSM), pp. 159-168.

CHRISMAN, N.R., 1989. Modelling error in overlaid categorical maps. In *The Accuracy Of Spatial Databases,* edited by M. Goodchild and S. Gopal (Bristol, PA: Taylor & Francis), pp. 21-34.

CHRISMAN, N.R., 1991. The error component of spatial data. In *Geographical Information Systems: Principles And Applications, 1,* edited by D.J. Maguire, M.F. Goodchild, and D.W. Rhind (Harlow, U.K.: Longman Scientific And Technical), pp. 165-174.

CHRISMAN, N.R., 1997. *Exploring Geographic Information Systems,* (New York: John Wiley and Sons), 298 pp.

CHRISTIAN, J.T., LADD, C.C., AND BAECHER, G.B., 1994. Reliability applied to slope stability analysis. *Journal of Geotechnical Engineering - ACSE,* 120(12):2180-2207.

CLAGUE, J.J., 1989. Quaternary geology of the Queen Charlotte Islands. In *The Outer Shores,* edited by G.G.E. Scudder and N. Gessler (Skidegate, Queen Charlotte: Queen Charlotte Islands Museum), pp. 65-74.

CLEAVES, D.A., 1994. Assessing uncertainty in expert judgements about natural resources. USDA Forest Service General Technical Report 50-110, (USDA Forest Service) .

CONGALTON, R.G., 1994. (ed.), *International Symposium on Spatial Accuracy of Natural Resource DataBases: Unlocking the Puzzle.* Williamsburg, Virginia, May 16 1994-May 20 1994.(Bethesda, Maryland: American Society for Photogrammetry and Remote Sensing),

CONGALTON, R.G. AND MACLEOD, R.D., 1994. Change detection accuracy assessment on the NOAA Chesapeake Bay pilot study. In *International Symposium on Spatial Accuracy of Natural Resource DataBases: Unlocking the Puzzle,* edited by R.G. Congalton. Williamsburg, Virginia, May 16 1994-May 20 1994.(Bethesda, Maryland: American Society for Photogrammetry and Remote Sensing), pp. 78-87.

CROWLEY, J.L. AND DEMAZEAU, Y., 1993. Principles and techniques for sensor data fusion. *Signal Processing,* 32(1):5-27.

CSAPLOVICS, E., 1992. Analysis of colour infrared aerial photography and SPOT satellite data for monitoring land cover change of a heathland region of the Causse du Larzac (Massif Central, France). *International Journal of Remote Sensing,* 13(3):441-449.

CUNIA, T. AND WHARTON, E.H., 1986. *Estimating Tree Biomass Regressions and their Error*, Proceedings of the Workshop on Tree Biomass Regression Functions and Their Contribution to the Error of Forest Inventory Estimates. Syracuse NY, May 26-30.(USDA Forest Service). 450pp.

DAVIS, J.C., 1986. *Statistics and Data Analysis in Geology* (New York: John Wiley and Sons).

DAVIS, T. J., 1994. Modeling and Visualizing Spatial Uncertainty using Fuzzy Logic and Monte Carlo Simulation. M.Sc. Thesis: University of Victoria. 131pp.

DAVIS, T.J. AND KELLER, C.P., 1997a. Modelling uncertainty in natural resource analysis using fuzzy sets and monte carlo simulation: slope stability prediction. *International Journal Of Geographical Information Science*, 11(5):409-434.

DAVIS, T.J. AND KELLER, C.P., 1997b. "Modelling and Visualising Multiple Spatial Uncertainties", *Computers and Geosciences*, 23(4), p.397-408.

DAVIS, T.J., KELLER, C.P., AND KLINKENBERG, B. 1998. A preliminary evaluation of the Lyell Island Rehabilitation Project. Report for The Archipelago Management Board, Gwaii Haanas. 72pp.

DE ROO, A.J.P., HAZELHOFF, L., AND BURROUGH, P.A., 1989. Soil erosion modelling using 'answers' and geographic information systems. *Earth Surface Processes and Landforms*, 14:517-532.

DELAVAR, MAHMOUD REZA, 1997. Development of probability maps to assess the accuracy and reliability of information in the output of a GIS system. PhD Dissertation: University Of New South Wales (Australia).

DIBIASE, D., MACEACHREN, A.M., KRYGIER, J.B., AND REEVES, C., 1992. Animation and the role of map design in scientific visualization. *Cartography And Geographic Information Systems*, 19(4):201-221.

DJAMDJI, J.-P., 1993. Geometrical registration of images: the multiresolution approach. *Photogrammetric Engineering And Remote Sensing*, 59(5):645.

DOMBURG, P., DEGRUIJTER, J.J., AND BRUS, D.J., 1994. A structured approach to designing soil survey schemes with prediction of sampling error from variograms. *Geoderma*, 62:151-164.

DRUMMOND, J., 1988. Fuzzy sub-set theory applied to environmental planning in GIS. In *EuroCarto 7*, edited by J.C. Muller. Enschede, Netherlands, Sep 20- 22 1988.(ITC). pp 121-134.

DU, L. AND LEE, J.S., 1996. Fuzzy classification of earth terrain covers using complex polarimetric SAR data. *International Journal Of Remote Sensing*, 17(4):809-826.

DUNN, J.C., 1974. A fuzzy relative of the isodata process and its use in detecting compact, well-separated clusters. *Journal of Cybernetics*, 3:22-57.

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*, 4(4):385-398.

DUTTON, G., 1992. Handling positional uncertainty in spatial databases. In *Proceedings of the Fifth International Symposium on Spatial Data Handling*, edited by P. Bresnahan, E. Corwin, and D. Cowen. Charleston, South Carolina, Aug. 3-7.(University of South Carolina), pp. 460-468.

EBERBACH, E., 1993. Representing spatial and temporal uncertainty. *Lecture Notes In Computer Science*, 682:129.

ECOSAT GEOBOTANICAL SURVEYS, 1989. Forest Harvesting Activities and LANDSAT Thematic Mapper Analysis of Lyell Island. Report submitted to Parks Canada, Queen Charlotte, BC.

EDWARDS, G. AND LOWELL, K.E., 1996. Modeling uncertainty in photointerpreted boundaries. *Photogrammetric Engineering And Remote Sensing*, 62(4):377-391.

EHLSCHLAEGER, C.R. AND GOODCHILD, M.F., 1994. Uncertainty in spatial data: defining, visualizing and managing data errors. In *Proceedings: GIS/LIS*. Phoenix, Arizona, Oct. 25-27.(Bethesda, Maryland: ASPRS/ACSM). pp 240-247.

EMMI, P.C. AND HORTON, C.A., 1995. A Monte Carlo simulation of error propagation in a GIS-based assessment of seismic risk. *International Journal Of Geographical Information Systems*, 9(4):447-461.

ESTEP, K.W., MACINTYRE, F., AND NOJI, T.T., 1994. Seal sizes and habitat conditions assessed from aerial photography and video analysis. *ICES Journal of Marine Science*, 51:253-261.

ESTES, J.E., STOW, D., AND JENSEN, J.R., 1982. Monitoring land use and land cover changes. In *Remote Sensing for Resource Management*, edited by C. Johannsen and J. Sanders (Ankeny, Iowa: Soil Conservation Society of America), pp. 100-110.

EVANS, B., 1997. Dynamic display of spatial data-reliability: does it benefit the map user? *Computers and Geosciences*, 23(4):254-270.

EVANS, D.L., 1994. Some considerations for AVHRR forest classification accuracy assessments. In *International Symposium on Spatial Accuracy of Natural Resource DataBases: Unlocking the Puzzle*, edited by R.G. Congalton. Williamsburg, Virginia, May 16-20 1994.(Bethesda, Maryland: American Society for Photogrammetry and Remote Sensing), pp. 161-167.

EVERITT, J.H., ESCOBAR, D.E., AND JUDD, F.W., 1991. Evaluation of airborne video imagery for distinguishing black mangrove (Avicennia germinans) on the Lower Texas Gulf Coast. *Journal of Coastal Research*, 7(4):1169-1173.

EVERITT, J.H., ESCOBAR, D.E., AND VILLARREAL, R., 1993. Integration of airborne video, global positioning system and geographic information system technologies for detecting and mapping two woody legumes on rangelands. *Weed Technology*, 7:981-987.

FANNIN, R.J. AND ROLLERSON, T.P., 1996, Assessing debris flow behavior in coastal British Columbia: Runout behavior. FERIC Special Report, SR-116.

FELICISIMO, A.M., 1994. Parametric statistical method for error detection in digital elevation models. *ISPRS Journal Of Photogrammetry And Remote Sensing*, 49(4):29-33.

FERGUSON, R.L., 1977. Linear regression in geography. Concepts and Techniques in Modern Geography #15. (Norwich, England: Geo-Books). 44p.

FERGUSON, R.L., WOOD, L.L., AND GRAHAM, D.B., 1993. Monitoring spatial change in seagrass habitat with aerial photography. *Photogrammetric Engineering And Remote Sensing*, 59(6):1033-1038.

FIGHT, R.D. AND BELL, E.F., 1994. Coping with uncertainty: a conceptual approach for timber management planning. USDA Forest Service General Technical Report PNW-5920pp.

FISHER, P.F., 1991a. First experiments in viewshed uncertainty: the accuracy of the viewshed area. *Photogrammetric Engineering and Remote Sensing*, 57(10):1321-1327.

FISHER, P.F., 1991b. Modeling and visualizing uncertainty in GIS. In *NCGIA Research Initiative 7: Visualization of Spatial Data Quality*, edited by M.K. Beard and B.P. Buttenfield (NCGIA), p. C63-C70.

FISHER, P.F., 1991c. Modelling soil map-unit inclusions by Monte Carlo simulation. *International Journal of Geographical Information Systems*, 5(2):193-208.

FISHER, P.F., 1992. First experiments in viewshed uncertainty: simulating fuzzy viewsheds. *Photogrammetric Engineering and Remote Sensing*, 58(3):345-352.

FISHER, P.F., 1993. Visualizing uncertainty in soil maps by animation. *Cartographica*, 30(2-3):20-29.

FISHER, P.F., 1994. Visualization of the reliability in classified remotely sensed images. *Photogrammetric Engineering and Remote Sensing*, 60(7):905-910.

FISHER, P.F., 1996. Extending the applicability of viewsheds in landscape planning. *Photogrammetric Engineering And Remote Sensing*, 62(11):1297-1302.

FLAMM, R.O. AND TURNER, M.G., 1994. Alternative model formulations for a stochastic simulation of landscape change. *Landscape Ecology*, 9(1):37-46.

FONSECA, L.M. AND MANJUNATH, B.S., 1996. Registration techniques for multisensor remotely sensed imagery. *Photogrammetric Engineering And Remote Sensing*, 62(9):1049-1056.

FOODY, G.M., 1996. Approaches for the production and evaluation of fuzzy land cover classifications from remotely-sensed data. *International Journal Of Remote Sensing*, 17(7):1317-1340.

FRIDLAND, V.M., 1974. Structure of the soil mantle. *Geoderma*, 12:35-41.

GARCIA, M.A., 1994. Terrain modeling with uncertainty for geographic information systems. *Proceedings- SPIE- The International Society For Optical Engineering*, 2357//PT1:273.

GIMBARZEVSKY, P., 1988. Mass wasting on the Queen Charlotte Islands: A regional inventory. (Victoria, BC: BC Ministry of Forests and Lands) ; Land Management Report #29.

GOLD, C.M., 1991. Applications of dynamic voronoi data structures. In *Second European Conference on GIS*. Brussels, April. p. 1090-1098.

GOODCHILD, M.F., 1980. Algorithm 9: simulation of autocorrelation for aggregate data. *Environment and Planning A*, 12:1073-1081.

GOODCHILD, M.F., 1989. Modeling error in objects and fields. In *The Accuracy Of Spatial Databases*, edited by M. Goodchild and S. Gopal (Bristol, PA: Taylor & Francis), pp. 107-114.

GOODCHILD, M.F., 1991. Issues of quality and uncertainty. In *Advances In Cartography*, edited by J.C. Muller (New York: Elsevier Applied Science Series). pp 140-172.

GOODCHILD, M.F., 1994. GIS error models and visualization techniques for spatial variability of soils. In *15th World congress of soil science*. Acapulco; Mexico, Jul 1994.(Mexican Society of Soil Science; ISSS), pp. 683-698.

GOODCHILD, M.F., BUTTENFIELD, B., AND WOOD, J., 1994. Introduction to visualizing data quality. In *Visualization in Geographical Information Systems*, edited by H.M. Hearnshaw and D.J. Unwin (New York: John Wiley and Sons), pp. 141-149.

GOODCHILD, M.F., CHIH-CHANG, L., AND LEUNG, Y., 1994. Visualizing fuzzy maps. In *Visualization in Geographical Information Systems*, edited by H.M. Hearnshaw and D.J. Unwin (New York: John Wiley & Sons, Inc.), pp. 158-167.

GORDON, J. AND SHORTLIFFE, E.H. 1992. The Dempster-Shafer theory of evidence: rule-based expert systems. In *The MYCIN Experience of the Stanford Heuristic Programming Project*, edited by B.G. Buchanan and E.H. Shortliffe (Reading, MA: Addison-Wesley)

GOTTESFELD-BROWN, L., 1992. A Survey of Image Registration Techniques. *ACM Computing Surveys*, 24(4):325-375.

GOVERNMENT OF .B.C., 1994. Forest Practices Code of British Columbia Act. S.B.C. 1994, c. 41.

GRAHAM, L.A., 1993. Airborne video for near-real-time vegetation mapping. *Journal of Forestry*, (Aug):28-34.

GUPTILL, S.C. AND MORRISON, J.L., 1995. *Elements of Spatial Data Quality* (New York: Elsevier).

HAINING, R.P., GRIFFITH, D.A., AND BENNETT, R.J., 1983. Simulating two-dimensional autocorrelated surfaces. *Geographical Analysis*, 15:247-253.

HALL, G.B. AND WANG, F., 1992. Comparison of Boolean and fuzzy classification methods in land suitability analysis by using geographical information systems. *Environment & Planning A*, 24(4):497-516.

HAMMOND, C., HALL, D., MILLER, S., AND SWETIK, P., 1992. Level I Stability Analysis (LISA) Documentation. Intermountain Research Station, General Technical Report INT-285, (Ogden, Utah: United States Forest Service, Department of Agriculture) .

HANNAH, M.J., 1981. Error detection and correction in digital terrain models. *Photogrammetric Engineering and Remote Sensing*, 47(1):63-69.

HARRIS, P.M., 1995. The integration of geographic data with remotely sensed imagery to improve classification in an urban area. *Photogrammetric Engineering And Remote Sensing*, 61(8):993.

HARTIGAN, J.A., 1975. *Clustering Algorithms* (New York: Wiley).

HEARNSHAW, H.M. AND UNWIN, A.R., 1994. *Visualization in Geographical Information Systems* (New York: John Wiley & Sons).

HERRING, J.R., 1991. Using spline functions to represent distributed attributes. In *Proceedings: Auto-Carto 10* . Baltimore, Maryland, March., pp. 46-57.

HEUVELINK, G.B.M., 1995. Identification of an error model for quantitative spatial attributes under different models of spatial variation. In *International symposium : Spatial accuracy of natural resource data bases: unlocking the puzzle*, edited by R. Congalton. Williamsburg, VA, May 1994.(Bethesda, MD: ASPRS), p. 267.

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*, 3(4):303-322.

HEUVELINK, G.B. AND BURROUGH, P.A., 1993. Error propagation in cartographic modelling using Boolean logic and continuous classification. *International Journal of Geographical Information Systems*, 7(3):231-246.

HOPE, A.C.A., 1968. A simplified Monte Carlo significance test procedure. *Journal of the Royal Statistical Society*, B30:583-598.

HORD, R.M. AND BROONER, W., 1976. Land-use map accuracy criteria. *Photogrammetric Engineering & Remote Sensing*, 42(5):671-677.

HOUGARD, T.R. AND VALDEZ, R.A., 1994. A multi-temporal, multi-accuracy fisheries GIS database for the Colorado River in Grand Canyon, Arizona. In *International Symposium on Spatial Accuracy of Natural Resource DataBases: Unlocking the Puzzle*, edited by R.G. Congalton. Williamsburg, Virginia, May 16-20 1994.(Bethesda, Maryland: American Society for Photogrammetry and Remote Sensing), pp. 67-77.

HOWES, D., HARPER, J., AND OWENS, E., 1994. British Columbia physical shore-zone mapping system. (Victoria: Resources Inventory Committee) 71p.

HUDSON, W.D. AND RAMM, C.W., 1987. Correct formulation of the kappa coefficient of agreement. *Photogrammetric Engineering & Remote Sensing,* 53(4):421-422.

HUNTER, G.J., 1995. A new model for handling vector data uncertainty in geographic information systems. In *33nd Annual conference — Information technology linking the Americas: your network to an expanded world,* edited by M. Salling. San Antonio; TX, Jul 1995., pp. 410-419.

HUNTER, G.J., ROBEY, B.H., AND GOODCHILD, M.F., 1996. Experimental development of a model of vector uncertainty. In *Spatial Accuracy Assessment in Natural Resources and Environmental Sciences,* edited by H.T. Mowrer, R.L. Czaplewski, and R.H. Hamre. Fort Collins, Colorado, May 21-23.(Fort Collins, Colorado: Rocky Mountain Forest and Range Experimental Station), pp. 217-224.

ILES, K., 1994. Directions in Forest Inventory. *Journal of Forestry,* 92(12):12.

JOURNAL, A.G., 1996. Modelling uncertainty and spatial dependence: stochastic imaging. *International Journal of Geographical Information Systems,* 10(5):517-522.

JOY, M.W., KLINKENBERG, B., AND CUMMING, S., 1994. Handling uncertainty in a spatial forest model integrated with GIS. In *GIS '94.* Vancouver, B.C., February. (Ministry of Supply and Services, Canada), pp. 360-365.

KANAZAWA, K., 1994. GOO: a database for temporal uncertainty management. In *Proceedings of the 10th Biennial conference of the Canadian Society for Computational Studies of Intelligence.* Banff, Alberta, May., pp. 197-204.

KELLER, C.P., 1994. Exploratory spatial data analysis (ESDA) - The next revolution in GIS. In *GIS '94.* Vancouver, B.C., February. (Ministry of Supply and Services, Canada), pp. 298-302.

KELLER, J.M., 1997. Fuzzy set theory in computer vision: A prospectus. *Fuzzy Sets and Systems,* 90(2):177-185.

KING, D. AND VLCEK, J., 1990. Development of a multispectral video system and its application in forestry. *Canadian Journal of Remote Sensing,* 16(1):15-22.

KOLLIAS, V.J. AND VOLIOTIS, A., 1991. Fuzzy reasoning in the development of geographical information systems. FRSIS: A prototype soil information system with fuzzy retrieval capabilities. *International Journal of Geographical Information Systems,* 5(2):209-223.

KONTOES, C., WILKINSON, G.G., BURRILL, A., GOFFREDO, S., AND MEGIER, J., 1993. An experimental system for the integration of GIS data in knowledge-based image analysis for remote sensing of agriculture. *International Journal Of Geographical Information Systems,* 7(3):247.

KOSKO, B. AND ISAKA, S., 1993. Fuzzy Logic. *Scientific American,* July:76-81.

KUNKEL, R. AND WENDLAND, F., 1997. WEKU - a GIS-supported stochastic model of groundwater residence times in upper aquifers for the supraregional groundwater management. *Environmental Geology,* 30(1-2):1-9.

LANGRAN, G., 1992. *Time in Geographic Information Systems* (Philadelphia: Taylor & Francis).

LANTER, D.P. AND VEREGIN, H., 1992. A research paradigm for propagating error in layer-based GIS. *Photogrammetric Engineering & Remote Sensing,* 58(6):825-833.

LEE, J., SNYDER, P.K., AND FISHER, P.F., 1992. Modeling the effect of data errors on feature extraction from digital elevation models. *Photogrammetric Engineering and Remote Sensing,* 58(10):1461-1467.

LEESE, M.N. AND MAIN, P.L., 1994. The efficient computation of Mahalanobis Distances and their interpretation in archaeometry. *Archaeometry,* 36(2):307.

LIU, R. AND HERRINGTON, L.P., 1996. The expected cost of uncertainty in geographic data. *Journal of Forestry,* 94(12):27-31.

LIU, R., 1995. The effects of spatial data errors on the GIS-based forest management decisions. PhD Dissertation: State University Of New York Col. of Environmental Science & Forestry. 224p.

LIVINGSTONE, D. AND RAPER, J., 1994. Modelling environmental systems with GIS: theoretical barriers to progress. In *Innovations in GIS 1*, edited by M.F. Worboys (Bristol, PA: Taylor & Francis), pp. 229-240.

LODIN, M. AND SKEA, D., 1996. A new method for evaluating positional map accuracy. In *Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, edited by H.T. Mowrer, R.L. Czaplewski, and R.H. Hamre. Fort Collins, Colorado, May 21-23.(Fort Collins, Colorado: Rocky Mountain Forest and Range Experimental Station).

LOWELL, K.E., 1993a. Initial studies on the quantification and representation of uncertainty in forestry data. In *GIS '93 Symposium*. Vancouver, BC, Feb 1993., pp. 791-796.

LOWELL, K.E., 1993b. Predictive model development and evaluation with unknown spatial units. *Photogrammetric Engineering and Remote Sensing,* 59(10):1509-1515.

LOWELL, K.E., 1997. An empirical evaluation of spatially based forest inventory samples. *Canadian Journal of Forest Research,* 27(3):352.

LOWELL, K.E., EDWARDS, G., AND LANGRAN-KUCERA, G., 1996. Modeling heterogeneity and change in natural forests. *Geomatica,* 50(4):425-440.

MACDOUGALL, E.B., 1975. The accuracy of map overlays. *Landscape Planning,* 2:23-30.

MACEACHREN, A.M., 1992. Visualizing uncertain information. *Cartographic Perspectives,* 13:11-17.

MACEACHREN, A.M., 1994. Time as a cartographic variable. In *Visualization in Geographical Information Systems*, edited by H.M. Hearnshaw and D.J. Unwin (New York: John Wiley & Sons, Inc.), pp. 115-130.

MACLEAN, A.L., D'AVELLO, T.P., AND SHETRON, S.G., 1993. The use of variability diagrams to improve the interpretation of digital soil maps in a GIS. *Photogrammetric Engineering and Remote Sensing,* 59(2):223-228.

MAFFINI, G., ARNO, M., AND BITTERLICH, W., 1989. Observations and comments on the generation and treatment of error in digital GIS data. In *The Accuracy Of Spatial Databases*, edited by M. Goodchild and S. Gopal (Bristol, PA: Taylor & Francis), pp. 55-68.

MAGNUSSEN, S., 1996. A co-ordinate-free area variance estimator for forest stands with a fuzzy outline. *Forest Science,* 42(1):76-85.

MARK, D.M. AND CSILLAG, F., 1989. The nature of boundaries on 'area-class' maps. *Cartographica,* 26(1):65-77.

MARSHALL, P.L., 1986. Risk and uncertainty in the analysis of timber supply. Report #85-12. Dept. of Forest Resource Management, Faculty of Forestry, UBC. 12 p.

MARSMAN, B. AND DE GRUIJTER, J.J., 1984. Dutch soil survey goes into quality control. In *Soil Information Systems Technology*, edited by P.A. Burrough and S.W. Bie (Pudoc, Wageningen, The Netherlands), pp. 127-134.

MARSMAN, B. AND DE GRUIJTER, J.J., 1986. Quality of soil maps: a comparison of survey methods. Soil Survey Papers No. 15. (Netherlands: Netherlands Soil Survey) .

MASCARENHAS, N.D., BANON, G.J., AND CANDEIAS, A.L., 1996. Multispectral image data fusion under a Bayesian approach. *International Journal Of Remote Sensing,* 17(8):1457-1471.

MATHER, P.M., 1995. Map-image registration accuracy using least-squares polynomials. *International Journal Of Geographical Information Systems*, 9(5):543.

MATHIEU, S., LEYMARIE, P., AND BERTHOD, M., 1994. Determination of proportions of land use blend in pixels of a multispectral satellite image. *IEEE Transactions*, 1154-1156.

MCBRATNEY, A.B., 1994. Allocation of new individuals to continuous soil classes. *Australian Journal of Soil Research*, 32:623-633.

MCBRATNEY, A.B. AND DEGRUIJTER, J.J., 1992. A continuum approach to soil classification by modified fuzzy k-means with extragrades. *Journal of Soil Sciences*, 43:159-175.

MCBRATNEY, A.B. AND MOORE, A.W., 1985. Application of fuzzy sets to climatic classification. *Agricultural and Forest Meteorology*, 35:165-185.

MENDOZA, G.A. AND SPROUSE, W., 1989. Forest planning and decision making under fuzzy environments: an overview and illustration. *Forest Science*, 35(2):481-502.

MINISTRY OF FORESTS, BC., 1995. Computer assisted methods to assess positional accuracy of forest cover map updates. Manuscript. 42 p.

MONCKTON, C., 1993. An investigation into the spatial structure of error in digital elevation data. In *First National conference on GIS research.* (Bristol, PA; Taylor and Francis) pp. 201-214.

MORRISON, P. H., 1975. Ecological and geomorphological consequences of mass movements in the Alder Creek watershed and implications for forest land management. B.Sc. Thesis: Eugene, Oregon: University of Oregon.

MOWRER, H.T., 1995. Monte Carlo techniques for propagating uncertainty through simulation models and raster-based GIS. In *International symposium : Spatial accuracy of natural resource data bases: unlocking the puzzle*, edited by R. Congalton. Williamsburg, VA, May 1994.(Bethesda, MD: ASPRS), pp. 179-188.

MOWRER, H.T., 1997. Propagating uncertainty through spatial estimation processes for old-growth subalpine forests using sequential Gaussian simulation in GIS. *Ecological Modelling*, 98(1):73-86.

MUCHONEY, D.M. AND HAACK, B., 1994. Change detection for monitoring forest defoliation. *Photogrammetric Engineering And Remote Sensing*, 60(10):1243-1251.

MULDER, J.A. AND CORNS, I.G.W., 1995. NAIA: a decision support system for predicting ecosystems from existing land resource data. *AI Applications*, 9(1):49-60.

NAESSET, E., 1996. Use of the weighted kappa coefficient in classification error assessment of thematic maps. *International Journal Of Geographical Information Systems*, 10(5):591-603.

NCDCDS, NATIONAL COMMITTEE FOR DIGITAL CARTOGRAPHIC DATA STANDARDS., 1988. The proposed standard for digital cartographic data. *The American Cartographer*, 15(1):9-140.

NEWCOMER, J.A. AND SZAJGIN, J., 1984. Accumulation of thematic map errors in digital overlay analysis. *The American Cartographer*, 11(1):58-62.

ODEH, I.O., MCBRATNEY, I.B., AND CHITTLEBOROUGH, D.J., 1992. Soil pattern recognition with fuzzy c-means: Applications to classification and soil-landform interrelationships. *Soil Science Society of America Journal*, 56:505-516.

OPENSHAW, S., 1989. Learning to live with errors in spatial databases. In *The Accuracy Of Spatial Databases*, edited by M. Goodchild and S. Gopal (Bristol, PA: Taylor & Francis), pp. 263-276.

OPENSHAW, S., CHARLTON, M., AND CARVER, S., 1990. Error propagation: a Monte Carlo simulation. In *Handling Geographical Information*, edited by I. Masser and M. Blakemore (UK: Longman Scientific), pp. 78-95.

OPENSHAW, S., WAUGH, D., AND CROSS, A., 1994. Some ideas about the use of map animation as a spatial analysis tool. In *Visualization in Geographical Information Systems*, edited by H.M. Hearnshaw and D.J. Unwin (New York: John Wiley and Sons). pp 131-139.

OR, D. AND HANKS, R.J., 1992. Spatial and temporal soil water estimation considering soil variability and evapotranspiration uncertainty. *Water Resources Research*, 28(3):803-814.

ORLAND, B., 1994. Visualization techniques for incorporation in forest planning geographic information system. *Landscape and Urban Planning*, 30:83-97.

OWENS, T. AND MCCONVILLE, D., 1996. A method for measuring the spatial accuracy of coodinates collected using the global positioning system. In *Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, Second International Symposium, edited by H.T. Mowrer, R.L. Czaplewski, and R.H. Hamre. Fort Collins, Colorado, May 21-23.(Fort Collins, Colorado: USDA Forest Service), pp. 351-358.

PALUBINSKAS, G., 1994. Comparison of texture-based and fuzzy classification approaches for regenerating tropical forest mapping using Landsat TM. In *Proceedings: The International Society for Optical Engineering*, Spatial Information From Digital Photogrammetry and Computer Vision, edited by H. Ebner and C. Heipke. Munich; Germany, September.(Society for Photogrammetry and Remote Sensing.), pp. 640-647.

PEIPE, J., 1995. High resolution still video camera for industrial photogrammetry. *Photogrammetric Record*, 85:135.

PERKAL, J., 1966. An attempt at objective generalization. *Geodezia Es Kartografia*, VII(2):130-142.

PEUCKER, T.K., 1975. A theory of the cartographic line. In *Proceedings International Symposium on Computer-assisted Cartography. AUTO CARTO II*, September 21-25.(US Department of Commerce and the ACSM), pp. 508-518.

PHINN, S.R., STOW, D.A., AND ZEDLER, J.B., 1996. Monitoring wetland habitat restoration in Southern California using airborne multispectral video data. *Restoration Ecology*, 4(4):412.

PIETRZYK, U., HERHOLZ, K., AND HEISS, W.-D., 1995. An interactive technique for three-dimensional image registration: validation for PET, SPECT, MRI and CT brain studies. *The Journal of Nuclear Medicine*, 35(12):2011.

POLLARD, S.J.T., 1994. Embracing uncertainty: developing risk assessment methodologies for environmental management in Europe. In *2nd International Symposium on Environmental Contamination in Central and Eastern Europe*. Budapest, Sep 1994., pp. 574-576.

PREPARATA, F.P. AND SHAMOS, M.I., 1985. *Computational Geometry* (Berlin: Springer-Verlag).

PRICE, K.P., 1992. Shrub dieback in a semiarid ecosystem: the integration of remote sensing and geographic information systems for detecting vegetation change. *Photogrammetric Engineering And Remote Sensing*, 58(4):455.

RAMLAL, B., 1996. A data quality model for the representation of mixed spatial variation. PhD Dissertation: University Of Maine. 266p.

REED, W.A., 1991. Planting decisions in the face of uncertainty. FEPA Research Unit Working Paper #156, University of BC., (Vancouver, BC: UBC) 30p.

RIDD, M.K. AND LIU, J., 1998. A Comparison of Four Algorithms for Change Detection in an Urban Environment. *Remote Sensing of Environment*, 63(2):95-102.

RINGROSE, S. AND MATHESON, W., 1992. The use of Landsat MSS imagery to determine the aerial extent of woody vegetation cover change in the west-central Sahel. *Global Ecology and Biogeography Letters,* 2(1):16-25.

ROBBINS, B.D., 1997. Quantifying temporal change in seagrass areal coverage: the use of GIS and low resolution aerial photography. *Aquatic Botany,* 58(3):259-265.

ROBINOVE, C.J., CHAVEZ, P.S., GEHRING, D., AND HOLMGREN, R., 1981. Arid land monitoring using LANDSAT albedo images. *Remote Sensing of Environment,* 11:133-156.

ROBINSON, V.B., 1988. Some implications of fuzzy set theory applied to geographic databases. *Computers, Environment and Urban Systems,* (12):89-98.

ROGOWSKI, A.S., 1996. Incorporating soil variability into a spatially distributed model of percolate accounting. In *Spatial Accuracy Assessment in Natural Resources and Environmental Sciences,* edited by H.T. Mowrer, R.L. Czaplewski, and R.H. Hamre. Fort Collins, Colorado, May 21-23.(Fort Collins, Colorado: Rocky Mountain Forest and Range Experimental Station). pp 178-185.

ROGOWSKI, A.S., 1996. Quantifying soil variability in GIS applications: Spatial distribution of soil properties. *International Journal of Geographical Information Systems,* 10(4):455-475.

ROHRER, W.H. AND SPARKS, D.L., 1993. Express saccades - the effects of spatial and temporal uncertainty. *Vision Research,* 33(17):2447-2460.

ROOD, K.M., 1984. An aerial photograph inventory of the frequency and yield of mass wasting on the Queen Charlotte Islands, B.C. (Victoria, BC: BC Ministry of Forests and Lands Management Report #34).

RUBENSTEIN, R.Y., 1981. *Simulation and the Monte Carlo Method* (New York: John Wiley & Sons).

RUIZ, M. H., 1995. A model of error propagation from digital elevation models to viewsheds: PhD Dissertation: University of Florida. 175p.

RUSINEK, H., TSUI, W.-H., AND LEVY, A.V., 1993. Principle axes and surface fitting methods for three-dimensional image registration. *The Journal of Nuclear Medicine,* 34(11):2019-2026.

RUSPINI, E.H., 1969. A new approach to clustering. *Information and Control,* 15:22-32.

SASOWSKY, K.C., PETERSEN, G.W., AND EVANS, B.M., 1992. Accuracy of SPOT digital elevation model and derivatives: utility for Alaska's North Slope. *Photogrammetric Engineering and Remote Sensing,* 58(6):815-824.

SCARFE. B.L., 1997, Screening criteria for restoration projects. In P.A. Slaney and D. Zaldokas (eds.), Fish Habitat Rehabilitation Procedures, Watershed Restoration Technical Circular No. 9, B.C. Watershed Restoration Program, Vancouver. pp. 4.1-4.11.

SCARFE. B.L., 1999, Evaluating resource benefits from hillslope and stream restoration programs. *Streamline,* 3(3): 5-12.

SHAFER, G., 1976. *A Mathematical Theory of Evidence* (Princeton, NJ: Princeton University Press).

SKIDMORE, A.K., 1989. A comparison of techniques for calculating gradient and aspect from a gridded digital elevation model. *International Journal of Geographical Information Systems,* 3(4):323-334.

SKIDMORE, A.K., TURNER, B.J., BRINKHOF, W., AND KNOWLES, E., 1997. Performance of a neural network: mapping forests using GIS and remotely sensed data. *Photogrammetric Engineering And Remote Sensing,* 63(5):501-514.

SRMB - SURVEY AND RESOURCE MAPPING BRANCH, 1990. British Columbia Specifications and Guidelines for Geomatics. Content Series, Volume 3, (Victoria, B.C.: Survey and Resource Mapping Branch) .

STEHMAN, S.V., 1996. Estimating the kappa coefficient and its variance under stratified random sampling. *Photogrammetric Engineering And Remote Sensing,* 62(4):401-407.

STEIN, A., 1994. Regionalized variable theory and geostatistics for modeling and representing soil spatial variability in GIS. In *Transactions: 15th World Congress of Soil Science.* Acapulco, Mexico, July 1994.(Mexico: International Society of Soil Science), pp. 699-717.

STOMS, D., 1987. Reasoning with uncertainty in intelligent geographic information systems. In *GIS '87, Second Annual International Conference.* San Francisco, Oct 26-30 1987.(Falls Church, VA: ASPRS), pp. 693-700.

SURYANA, N., 1993. The application of evidence and fuzzy sub-set theory in handling imprecision and uncertain data handled by geographic information systems. In *7th Annual Symposium on Geographic Information Systems in Forestry, Environment and Natural Resources Management - GIS 93.* Vancouver; Canada, Feb., pp. 871-882.

SURYANA, N., 1993. The application of evidence and fuzzy subset theory using certainty factor in handling imprecision and uncertain data handled by geographical information systems (GIS). In *GIS '93.* Vancouver, BC, Feb 1993., pp. 871-882.

TAYLOR, J.R., 1982. *An Introduction to Error Analysis* (Oxford: University Press).

THOM, C. AND JURVILLIER, I., 1997. An aerial digital frame camera for photogrammetry: Radiometric comparison with scanned imagery. *Geomatica,* 51(2):133-142.

THOMPSON, W.A. AND VERTINSKY, I., 1991. Evaluation of silvicultural investments under uncertainty using simulation. FEPA Research Unit Working Paper #161, University of BC., (Vancouver, BC: UBC) 18pp.

TRAAS, T.P., 1994. Uncertainty analysis and risk assessment combined: application to a bioaccumulation model. In *Proceedings of The International Conference on Predictability and Nonlinear Modelling in Natural Sciences and Economics.* Wageningen; The Netherlands, Apr 1993, pp. 516-526.

UENOHARA, M. AND KANADE, T., 1995. Vision-based object registration for real-time image overlay. *Computers in Biology and Medicine,* 25(2):249.

VAN DER WEL, F.J.M., HOOTSMAN, R.M., AND ORMELING, F.J., 1994. Visualization of data quality. In *Visualization in Modern Cartography,* edited by A.M. MacEachren and D.R.F. Taylor (Amsterdam: Elsevier Scientific Publishing Co.), pp. 67-92.

VANKOOTEN, G.C., BULTE, E., AND ELLS, A., 1996. Uncertainty and forest land use allocation in BC: Fuzzy decisions and imprecise coefficients. Pacific Forestry Centre Working Paper 96.11, (Victoria, BC: Canadian Forest Service) 45p.

VANKOOTEN, G.C., VANKOOTEN, R.E., AND BROWN, G., 1990. Modelling the effect of uncertainty on harvest age for the boreal forest of N.W. Alberta. FEPA Research Unit Working Paper #139, University of BC., (Vancouver, BC: UBC) 15p.

VEREGIN, H., 1989. A Taxonomy of error in spatial databases. Technical Paper 89-12, National Centre for Geographic Information and Analysis, (Santa Barbara, California).

VEREGIN, H., 1996. Error propagation through the buffer operation for probability surfaces. *Photogrammetric Engineering And Remote Sensing,* 62(4):419-428.

VOGELMANN, J.E. AND ROCK, B.N., 1988. Assessing forest damage in high-elevation coniferous forests in Vermont and New Hampshire using Thematic Mapper data. *Remote Sensing of Environment,* 24:227-246.

WANG, F., HALL, G.B., AND SUBARYONO, 1990. Fuzzy information representation and processing in conventional GIS software: database design and application. *International Journal of Geographical Information Systems*, 4(3):261-283.

WEBSTER, R., 1985. Quantitative spatial analysis of soil in the field. *Advances in Soil Science*, 3:1-70.

WEETMAN, G.F., 1990. Biological uncertainty in growth and yield benefits of second-growth silviculture actions and old-growth protection and liquidation strategies in British Columbia. Working Paper #148, FEPA Research Unit, UBC, (Vancouver, BC: FEPA Research Unit, UBC) 23p.

WEISMILLER, R.A., KRISTOF, S.J., SCHOLZ, D.K., ANUTA, P.E., AND MOMIN, S.N. 1990. Evaluation of change detection techniques for monitoring coastal zone environments. In *Symposium on Remote Sensing of Environment, 2* (Ann Arbor, Michigan: Environmental Research Institute of Michigan), pp. 1229-1235.

WESTLAND RESOURCES GROUP, 1994. Gwaii Haanas Ecological Land Classification. Report and Databases. 227 p.

WICKWARE, G.M. AND HOWARTH, P.J., 1981. Change detection in the Peace-Athabasca Delta using digital LANDSAT data. *Remote Sensing of Environment*, 11:9-25.

WILKINSON, G.G., 1996. A review of current issues in the integration of GIS and remote sensing data. *International Journal Of Geographical Information Systems*, 10(1):85-101.

WINSTON, P.H., 1984. *Artificial Intelligence* (Reading, MA: Addison-Wesley).

WOHLERS, T., 1997. The technology behind 3D digitizing. *Computer Graphics World*, 20(3):47.

WOOD, J., 1994. Visualizing contour interpolation accuracy in digital elevation models. In *Visualization in Geographical Information Systems*, edited by H.M. Hearnshaw and D.J. Unwin (New York: John Wiley and Sons), pp. 168-180.

XIAO, Y., 1996. Topographic characterisation for DEM error modelling: University of Britih Columbia447pp.

YOSHINARI, Y., PEDRYCZ, W., AND HIROTA, K., 1993. Construction of fuzzy models through clustering techniques. *Fuzzy Sets and Systems*, 54( ):157-165.

YOUCAI, H. AND WENBAO, L., 1997. Building the estimation model of digitizing error. *Photogrammetric Engineering & Remote Sensing*, 63(10):1203.

ZADEH, L.A., 1970. Quantitative Fuzzy Semantics. *Information Science*, 3:1-17.

ZHANG, Z., SHIMODA, H., FUKUE, K., AND SAKATA, T., 1988. A landcover classification using fuzzy inference. Proceedings of ACRS, Bangkok, Thailand (manuscript).

ZIMMERMAN, H., 1990. Problems and tools to model uncertainty in expert and decision support systems. *Mathematical Computer Modelling*, 14:8-20.

Appendix A

The Hard k-Means and Fuzzy c-Means Algorithms

Both of these algorithms are designed to divide a number of individual samples (individuals) into logical classes. The fuzzy c-means version is a generalisation of the hard k-means version, in that it allows membership in more than one class. The hard k-means algorithm is structured as follows.

Given a set of n individuals (samples) divided into k discontinuous classes, each individual is a member of exactly one class. This can be represented by a n x k matrix of memberships $\mathbf{M} = (m_{ic})$, where $m_{ic}=1$ if individual I belongs to class c and m=0 otherwise. The following conditions apply to assure that classes are mutually exclusive:

$$\sum_{c=1}^{k} m_{ic} = 1, \qquad i = 1,\ldots,n \qquad (1)$$

$$\sum_{n=1}^{n} m_{ic} > 0, \qquad c = 1,\ldots,k \qquad (2)$$

$$m_{ic} \in \{0,1\}, \qquad i = 1,\ldots,n;\ c = 1,\ldots,k. \qquad (3)$$

In other words, the sum of all class memberships of a sample is 1, while the number of classes is greater than zero, and each sample can only have a membership value of 1 (belongs) or zero (does not belong). The theory of fuzzy sets relaxes condition (3), allowing memberships to be partial. Thus condition (3) is replaced by:

$$m_{ic} \in [0,1], \qquad i = 1,\ldots,n;\ c = 1,\ldots,k. \qquad (3a)$$

Now the membership values of each sample no longer have to be only 1 or 0, but can fall in between, as long as they sum to 1 for all classes.

The hard-k-means algorithm minimises the within-class sum-of-squares errors function $J(\mathbf{M},\mathbf{C})$ under conditions (1), (2), and (3):

$$J(\mathbf{M},\mathbf{C}) = \sum_{i=1}^{n} \sum_{c=1}^{k} m_{ic} d^2(\mathbf{x}_i,\mathbf{c}_c), \qquad (4)$$

where $\mathbf{C} = (c_{cv})$ is a k x p matrix of class centres, c_{cv} denoting the value of the centre of class c for variable v.

$\mathbf{x}_i = (x_{i1},\ldots,x_{ip})$ is the vector representing individual i.

$\mathbf{c}_c = (c_{c1},\ldots,c_{cp})$ is the vector representing the centre of class c, and $d^2(\mathbf{x}_i,\mathbf{c}_c)$ is the square distance between $\mathbf{x}_i$ and $\mathbf{c}_c$ according to a chosen definition of distance, further denoted as d^2_{ic}.

This function measures the distance from the sample to the centre of each class along each attribute axis, and calculates the sum of the distances squared. If the purpose of the procedure is to allocate a new sample to known classes, then the smallest of these distances determines which class it is assigned to. If the purpose is to generate the classes themselves, then the procedure is used iteratively, looking for a set of classes with the smallest within-class total error.

The fuzzy generalisation employed in fuzzy-c-means provides the memberships with an exponent φ which determines the fuzziness of the solution. A value of 1 represents a hard partition, while values above this increase the fuzziness of the memberships. In other words, as values of this exponent increase, the 'fuzzy boundary' around a class grows, and samples that are distant from the class (in attribute space) are given higher membership values.

If a value of 1 is used for φ, the solution is usually solved by iterative relocation of individuals to the classes as noted above. If $\varphi > 1$, then the solution can be minimised by iteration of the following equations:

$$m_{ic} = \frac{d_{ic}^{-2/(\varphi-1)}}{\sum_{j=1}^{k} d_{ij}^{-2/(\varphi-1)}} \qquad I = 1,\ldots n;\ c = 1,\ldots,k \qquad (4)$$

$$\mathbf{c}_c = \frac{\sum_{i=1}^{n} m_{ic}^{\varphi}\mathbf{x}_i}{\sum_{i=1}^{n} m_{ic}^{\varphi}} \qquad c = 1,\ldots,k \qquad (5)$$

The main algorithm is as follows:

---

1) choose the number of classes k;
2) choose a value for the fuzziness exponent φ, with $\varphi > 1$;
3) choose a definition of distance in the attribute space;
4) choose a value for the stopping criterion ε ($\varepsilon = 0.001$ gives reasonable convergence);
5) initialise $\mathbf{M} = \mathbf{M}_0$, using either random memberships or memberships from the hard-k-means partition;

6) at iteration $l = 1,2,3,\ldots$ (re-)calculate $\mathbf{C} = \mathbf{C}_l$ using equation (5) and $\mathbf{M}_{l-1}$;

7) re-calculate $\mathbf{M} = \mathbf{M}_l$, using equation (4) and $\mathbf{C}_l$;

8) compare $\mathbf{M}_l$ to $\mathbf{M}_{l-1}$; if $<= \varepsilon$, then stop; otherwise return to step (6).

By relaxing the 'hard' boundaries around classes, and allowing partial memberships in each class, the number of possible solutions becomes effectively infinite. The purpose of this algorithm is to look for a reasonable solution, rather than the best solution.

Appendix B

GPS Accuracy Statistics

The field sites that were surveyed with GPS were also tested for GPS positional accuracy. The first site (Louise Islands) was surveyed with a Trimble Scoutmaster GPS in non-differential mode. The second site, the Lyell Island aerial and ground truthing work, was surveyed with the same unit, also in non-differential mode. Differential correction was unavailable (post-processing was also not possible due to the remote location of the sites). However, differential tests *were* part of related research not reported in this document. This other research made use of a a Magellan Mark 10 GPS, and was corrected to a station approximately 60km distant. For comparison purposes accuracy tests were performed for both differential and non-differential co-ordinates. In both cases the methodology utilised was as follows:

1. Tie in to the local survey grid via a control point.
2. Take multiple readings with the GPS. For each site 2500 readings were taken at four different times during the day.
3. Readings are processed as offsets from the known position.
4. Statistics are calculated as below.

Standard Error is calculated as: $$\sigma_Z = \left(\sum_{i=1}^{n} \frac{\left(x_i - \bar{x}\right)^2}{n-1} \right)^{0.5} \tag{1}$$

Mean Square Error = $$\sigma_{xy} = \left(\sigma_x^2 + \sigma_y^2\right)^{0.5} \tag{2}$$

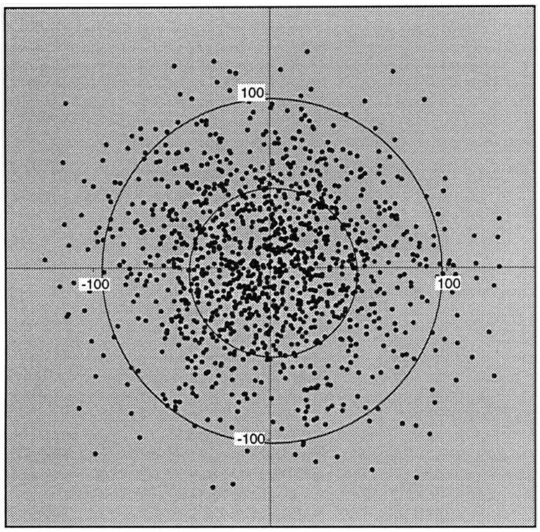
(at 90% probability multiply by 1.520)

Circular Standard Error is: $$\sigma_c = 0.7071\left(\sigma_x^2 + \sigma_y^2\right)^{0.5} \tag{3}$$

at 90% probability this is the Circular Map Accuracy Standard (CMAS); multiply by 2.146.

The following graphs show a random subset of the 2500 readings. They have a tested normal distribution.

Lyell Island Test, CMAS = 110m



Differential Test, CMAS = 2.8m

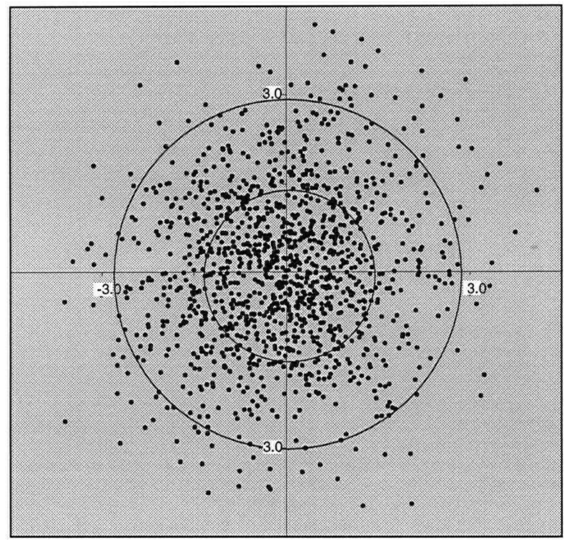


Figure B.1. Random subset of the 2500 test points gathered for non-differential and differential tests.

Appendix C

Cross Correlograms and Significance Tests for Sample Transects

The tests summarised below (Figure C.1) are a t-test of the significance of the correspondence between the sample transects and the model. A 'y' indicates a significant correspondence, while a 'n' indicates lack of significance. For example, the graph below (Figure C.2) shows the actual curves (the model and sample) that are compared in generating the cross-correlogram (next page) for type 2, transect 2, which generated a significant t-test result of 5.45 when a lag of 1 and an Avg(9) smoothing was applied.

Original Data									Optimised Data						
Type	Transect	X-Corr	Lag	Avg(x)	n	DF	t test	Signif?	X-Corr	Lag	Avg(x)	n	DF	t test	Signif?
1	1	0.888	2	9	30	28	**10.22**	y	0.638	-4	5	30	28	**4.384**	y
2	1	0.67	1	9	30	28	**4.776**	y	0.535	-4	5	30	28	**3.351**	y
3	1	0.468	-4	7	32	30	**2.901**	y	0.801	3	9	32	30	**7.328**	y
1	2	0.542	-2	7	48	46	**4.374**	y	0.24	-3	5	48	46	**1.677**	y
2	2	0.635	1	9	46	44	**5.452**	y	0.565	1	9	46	44	**4.542**	y
3	2	-0.0114	3	Raw	50	48	**-0.079**	n	-0.14	1	Raw	50	48	**-0.98**	n
1	3	0.093	-4	5	41	39	**0.583**	n	0.449	-1	Raw	41	39	**3.138**	y
2	3	0.21	-4	9	39	37	**1.307**	n	0.35	0	7	37	35	**2.21**	y
3	3	0.445	1	9	39	37	**3.023**	y	0.217	-2	Raw	39	37	**1.352**	n
1	4	0.204	0	9	22	20	**0.932**	n	0.392	4	9	26	24	**2.087**	y
2	4	0.7	-1	9	22	20	**4.384**	y	0.32	1	5	28	26	**1.722**	y
3	4	-0.06	4	0	26	24	**-0.294**	n	0.559	2	9	26	24	**3.303**	y

Figure C.1. t-test results for the cross-correlogram and lag generating maximum correspondence between sample and modelled transects (5% signif.)

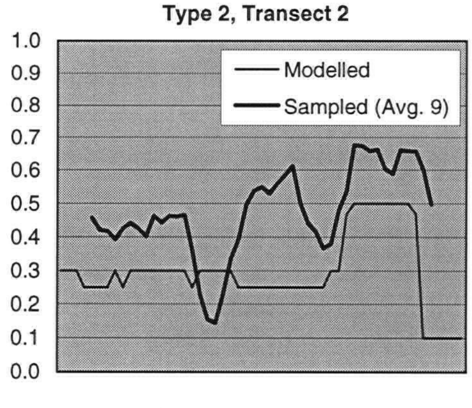


Figure C.2. Sampled and modelled transect example.

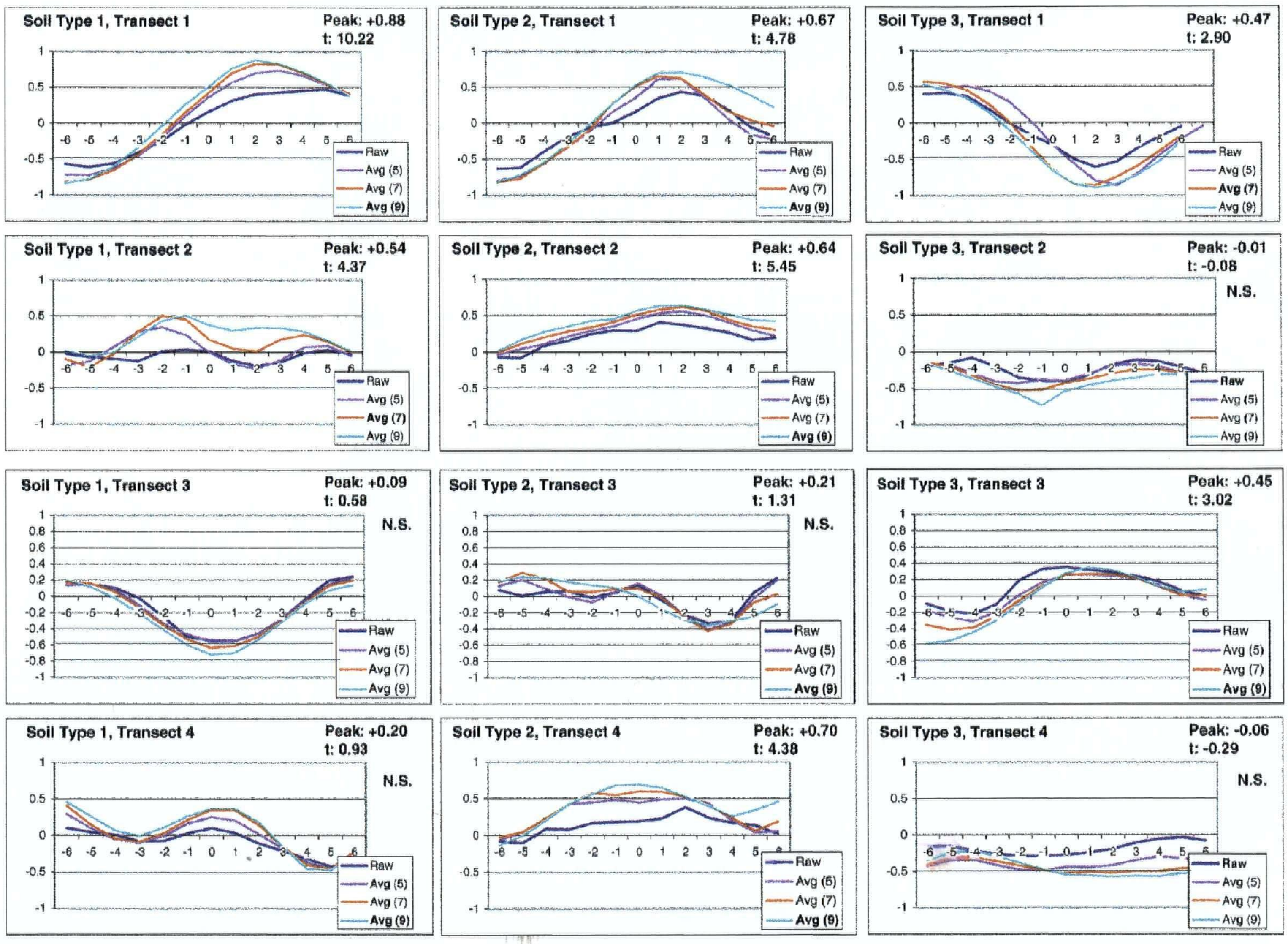


Figure C.3 Cross-correlograms for the original model. t-test results and significance are noted on each graph

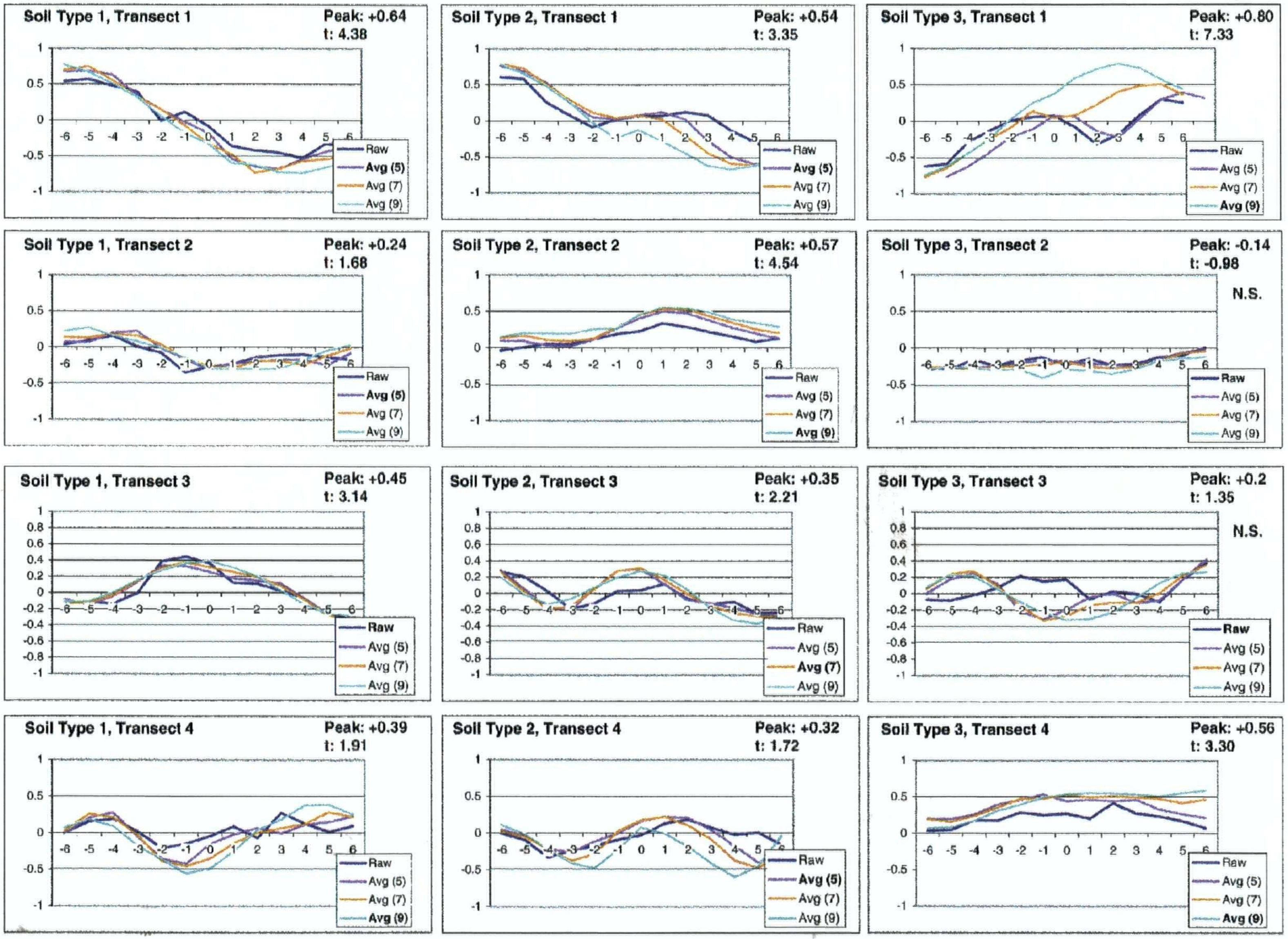


Figure C.4 Cross-correlograms for the updated model. t-test results and significance are noted on each graph

Appendix D

Database Development

Introduction

The research presented in this appendix focuses on inventory update in large-scale (i.e., limited area) projects. It involves the merging of several research streams: a subset of the uncertainty models discussed in some of the chapters, image fusion of dissimilar images in a GIS, and change detection in forest inventory. The questions addressed in this research are:

- How can areas identified as 'uncertain' (e.g., out of date or lacking data relative to a particular application) in a spatial inventory model be efficiently updated?
- How can data uncertainty be tracked through this process?
- What other types of inventoried objects can be updated (i.e., other than polygons), and with what level of accuracy?

This work has been placed in an Appendix to this dissertation because it represents a tangent to the overall theme of uncertainty model verification. The tools developed in this work are used to develop the landslide database used in Chapter Six to verify uncertainty model output. The data update tool development project stands on its own as a 'nested' piece of research. This appendix presents it as such.

Updating Areas of High Uncertainty

As noted in Chapters Two and Three, one major advantage of uncertainty tracking databases and modelling procedures is their ability to highlight areas of unreasonably high uncertainty (where 'unreasonable' is defined through risk assessment). For example, data combinations such as those displayed in Figure 3.11 can highlight key areas where uncertainty is high and information is crucial. Given the spatial specificity of this information, it would be highly inefficient to attempt data updates through a complete re-inventory. A more practical solution would be to gather data only in the areas where uncertainty is high. However, this presents a number of practical problems, chief of which is the issue of knitting these small inventories together into the bigger picture. Depending on the technique(s) utilised, the information might arrive with variations in scale, temporal variations, different levels of skewing error, and other types of variability. An update system will have to address the relevant types of variability in a more explicit manner than a standard inventory system.

For example, a particular stand of trees may have had its boundaries delineated in 1985 using a source with a spatial accuracy of five metres. In 1995, an update is produced using a system accurate to one metre. In the new data the stand boundary appears to have shifted four metres to the west. Unless there is spatially specific metadata available to indicate the 1985 spatial uncertainty level, this four metre shift will be considered to be an actual change in the dataset, when in fact it may be an artefact of the process.

Chapters Three and Four focused on addressing uncertainty in basic terrain inventory data and in modelling routines; however, an update system has the potential to address a wider variety of uncertainties. For example, certain stands of trees grow faster than others due to species composition, site situation, or treatments. In a growth model the uncertainty in the volumes of these stands increases over time at a higher rate than other stands, making them candidates for site-specific inventory update. Similarly, high elevation areas or sections containing slow growing species could be updated on a much longer cycle.

Unfortunately, although the technology exists, the potential for local inventory update has largely remained unrealised. This is due to both inertia in traditional inventory systems, as well as to a variety of technical factors. For instance, there is no quick way of registering a new air photo into an old air photo series and transferring observed changes into the inventory. Similarly, when new data are entered, the old and new lines will undoubtedly mismatch to some degree, generating slivers during comparison operations.

Shifting the traditional re–inventory system to an inventory update system could produce considerable advantages in the area of forestry planning. First, an update system could drastically reduce the costs of inventory or, if desired, make inventory more timely by allowing site-specific updates between major inventories. Second, by maintaining a standard baseline that is gradually altered over time, an update system would make a fully temporal database possible. Detailed analysis of change over time would be just one benefit of such a system. This baseline makes possible a third advantage: properly integrated temporal models of growth, terrain processes, and many others. Decision support systems would also benefit from this update system. Currency in information is perhaps the single greatest defect in decision modelling in this industry (Liu and Herrington 1996).

The tracking of uncertainty through the updating process allows the decision-maker to make a quantified risk assessment of factors such as 'cost of update vs. cost of being wrong'. It also integrates the update process into the overall process of inventory management under uncertainty.

Data Sources for Performing Updates

One area where inventory update work is well advanced is remote sensing satellite image analysis. In this field, 'change detection' is the term used to refer to algorithms that allow spatial and attribute comparisons between images whose acquisition dates vary (e.g., Congalton and MacLeod 1994). Although originally limited to the detection of small-scale changes due to sensor resolution, recent advances in have led to some large-scale applications in a wide variety of areas, including forestry. Change detection algorithms are particularly suited to determining overall percentage changes, as in 'a 10% decrease in wetlands in Chesapeake Bay'; however, they are generally less suited to repeated measurements of small spatial changes as required for large-scale forestry inventory work.

Change detection using a spatial inventory as the base requires a close integration between remote sensing analysis software and GIS. The GIS maintains the spatial data, performs analysis, integrates with models and decision support tools and provides output capability. The remote sensing software focuses on satellite image correction, enhancement and spatial control. Accurate, high-resolution data, applied to a closely integrated remote sensing and GIS package, can make change detection possible at reasonably large scales.

Yet satellites are not the only source of remotely sensed data. Aircraft mounted sensors can generate extremely high-resolution datasets that are also useful for change detection. Aerial photos, digital sensors, imaging spectrometers, and digital video are some of the sources utilised. Aircraft-based detection has several advantages over satellites. It can be of higher resolution, can be flown when required (rather than relying on satellite schedules), and sensors can be tuned for a particular task. Yet aircraft-based detection is expensive. Specialised aircraft or instrument mounts are required, and in remote areas the aircraft may have significant transit times. This type of detection is not generally suitable for smaller projects.

However, one relatively inexpensive source of data is found in oblique aerial photography or aerial video. Oblique images have long been part of aerial photography; they are invaluable in helping to identify structures and objects on the ground. However, they have been of little use in constructing or maintaining spatial databases. With existing tools only a vertical photo can be registered to a planimetric database. An oblique image would require a much more complex system of registration that would have to address image depth, variability in skewing, and numerous other factors. However, oblique images have the advantages of being inexpensive, quick, and useful under cloud cover and in extreme terrain. In this appendix techniques are developed to make use of oblique images as a data source for inventory update.

Tracking Uncertainty

There are several reasons why measures of uncertainty must be an explicit part of a spatial database that will be used for inventory update. For one, it is necessary to specify the areas where uncertainty is high. Second, uncertainty must be tracked through updates in order to determine whether *visible* change is *actual* change. Third (and specific to the oblique data source), the variability created by skewing an oblique image must be merged with existing metadata and newly added data in order to determine the uncertainty in the final product.

The project presented in this appendix addresses the questions of inventory update and related uncertainty tracking through the development of an 'image fusion' system that is suitable for inventory update and change detection. The system developed herein is not designed to replace inventory or most re-inventory systems. Its utility is primarily in areas where site-specific updates are required between larger inventories, where regular updates are required on small areas, or where terrain and remote locations make other options too expensive. Uncertainty management is an integral part of the system, making possible both registration under spatial uncertainty, uncertainty tracking, and change detection decisions regarding artefactual vs. real change.

Updating Other Types of Objects

The system developed in this research focuses on the update of the boundaries of polygonal or linear structures. However, this image fusion system is also capable of addressing other types of measurements. For example, slope distance, ground distance, and the height of objects in oblique frames can be measured directly on the image. This makes possible other types of forest inventory with the system. A number of forest mensuration parameters typically gathered through ground survey can be measured. Other research projects will examine these applications separately.

Background

This section reviews three areas relevant to data update through the integration of oblique data with GIS. The first is 'change detection' as implemented in remote sensing applications and research. Change detection (as applied in remote sensing) is the analysis of several vertical images or photographs of a particular area taken at different time periods in order to quantify change. The second area is 'image fusion'—which refers to the merging of images of the same scene, but gathered with different sensors or at different times. In the application introduced herein, the two 'sensors' will be an oblique image of a scene and a GIS-generated perspective view. The third area is 3-D digitising, which will be required to extract information from the 'fused' images.

Change Detection

Remote sensing change detection uses two or more satellite or aircraft-based images (photographic or digital), spread over time, to quantify spatial or land-use changes on the surface of the earth. Change detection work was originally based on aerial photography and photo interpretation. Even the basic update of a forest inventory could be considered as change detection. However, the term is more commonly used in reference to digital images and algorithmic methods of automatically detecting change.

Change detection in remote sensing is not a straightforward process of image comparison. Two images of the same location at different times will have numerous subtle differences in addition to any *real* change. Differences in illumination and slight differences in angle will create variations in reflectance. Therefore, a processing step is required before the images can be compared. For example, principle components analysis can be used to maximise data variance along principle axes. This process can separate out minor reflectance variation from forest cover variation, and then separate those factors from urban development. Other algorithms for change detection include image differencing (using a variety of bands—e.g., Estes *et al.* 1982), spectral-temporal change classification (Weismiller *et al. 1990),* and post-classification change detection differencing (e.g., Wickware and Howarth 1981).

Change detection routines have been applied in many areas. The principle ones are the evaluation of land cover changes associated with urbanisation (Wickware and Howarth 1981; Ridd and Liu 1998), desertification ((Robinove *et al.* 1981), and coastal zone monitoring (Congalton and MacLeod 1994). In forestry remote sensing, change detection has been used for specific types of defoliation studies (e.g., Vogelmann and Rock 1988; Muchoney and Haack 1994), as well as more general woody land cover change (Ringrose and Matheson 1992).

Despite the numerous advances in image analysis, change detection with aerial photography is still a useful technique. In some tasks photography is better suited to the target being modelled. For example, change detection of coastal and shoreline features may require special films and very specific timing to match tidal fluctuations (e.g., Ferguson *et al.* 1993 or Robbins 1997). In other cases historical data exists only as photographic images. And in some instances the high-resolution available from photos is important (e.g., Csaplovics 1992).

This latter example (Csaplovics 1992) discusses the integration of aerial photography with a satellite sensor in order to increase available information. This type of combination, generally termed 'data fusion', is becoming a common remote sensing tool and analytical data source.

Data Fusion under Uncertainty

The term 'data fusion' has typically been used in the design of advanced avionics control systems, or for the combination of several sensors in the field of image processing for computer vision (e.g., Crowley and Demazeau 1993; Mascarenhas *et al.* 1996). Typically, two or more representations of the same scene are registered to each other and the data about the scene are combined—leading to stereo computer vision, increased detail in radar displays, or simply more precise images. In most cases the two sensors being 'fused' represent slightly different versions of the same scene. For example, two radar antennas may operate at different frequencies, or two cameras may be slightly offset. In contrast, the research discussed herein focuses on the fusion of two very different representations of a scene: the archived, planimetric, stylised vectors of a GIS database fused with images obtained from oblique aerial video imagery.

A number of researchers have explored the integration of remote sensing data with GIS data (e.g., Price 1992; Kontoes *et al.* 1993; Harris 1995; Wilkinson 1996). Such work brings to light a number of issues, the main ones being positional ground control, grid rectification, and variations in cell shape and size. Minimising the loss of data through these procedures is a central goal. Such integration has generally been very successful, particularly in the case of pre-processed satellite data. In fact, the increasing availability of satellite data has made such integration crucial for the success of many GIS.

Airborne sensors, operating much closer to the earth, do not generate data that are quite as easily integrated into the fixed grids used in GIS raster structures. For example, data derived from imaging spectrographic scanners may require considerable correction of aircraft induced registration problems (roll, pitch...) before even beginning the process of GIS integration. As sensors move closer to the earth, orthographic correction also becomes an important element of data pre-processing.

Video

Recently, airborne video cameras have been utilised as data gathering tools for a variety of purposes. Video frames are georeferenced with a GPS system, allowing data gathering without the use of precise flightlines. Used as multispectral scanners, they can provide relatively inexpensive data for many types of studies (e.g., King and Vlcek 1990). Airborne video has also been used in place of standard aerial photography in several situations: 1) where large areas need to be covered quickly, but the detail of aerial photography is not required; 2) where repeated images are required, and the cost of aerial photography is prohibitive; and 3) when quick production is required. Examples include Everitt *et al.* (1993) identifying shrubs on rangeland, Graham (1993) mapping forest vegetation, Estep *et al.* (1994) determining seal size and location on ice floes, and a variety of others covering many of the topics formerly reserved for expensive sensors.

Most video systems do not offer the resolution available from aerial photographs (although high-resolution video photogrammetric systems are under development, e.g., Peipe 1995; Thom and Jurvillier 1997). Nevertheless, video has carved out a large niche as a source of spatial data. For many projects, price and speed are of greater importance than resolution. Video systems are also commonly used to gather descriptive (i.e., non-georeferenced) information. Oblique video—shot out of helicopter doors or from remote camera pods—provides reference information for a variety of tasks. Foresters use video to examine cutblocks and monitor code adherence. Large sections of the province's coastline have been captured on oblique video, allowing rough estimates of shoreline type to be developed (Howes *et al.* 1994). Oblique video has a number of advantages over vertical video:

1. oblique images can cover a wider area than vertical images;
2. oblique images can be taken below cloud cover, when vertical capture would be impossible;
3. the view obtained is more natural—allowing easier interpretation of landforms, vegetation, and human-made objects;
4. specialised camera mounts are not required, allowing almost any aircraft to be used, and allowing video gathering to be incorporated into other aerial work; and
5. targets located on steep terrain can be relatively 'invisible' to a vertical image. An oblique image can capture a better representation of area and shape.

The principle reason that oblique video is relegated to gathering merely descriptive data is that georeferencing and registration to spatial databases are not currently possible. For example, those using oblique imagery (photography) to detect change in cutblocks commonly utilise a series of pictures taken at fixed ground photo-stations. Such images can be compared over time, but actual area estimates are very difficult to obtain from these oblique images. Oblique aerial video is also often used as a preliminary survey technique; a quick over-flight and review of video leads to more detailed ground surveys.

There is, however, sufficient information available to allow oblique imagery to be georeferenced. There are numerous video systems in use that store GPS data on video frames. Vertical video systems use these data to locate an image's nadir point, allowing registration to a spatial database. For oblique imagery, the aircraft location data, combined with camera angle and lens information, can be used to register an oblique image.

Image Registration for Fusion

Registration is a fundamental task in image processing. It is used to match two or more images taken, for example, at different times, from different sensors, or from different viewpoints. Most registration tasks in geo-information processing involve tying an image into an existing planimetric database (remote sensing/GIS integration tools), or registering two images taken at different times. However, registration is also used for many other purposes: military target acquisition systems, robotic stereo vision for autonomous navigation, and aligning images from different medical sensors for diagnosis.

A broad range of registration techniques have been developed for these various types of data and problems. These techniques have been, for the most part, developed independently in different fields. However, using a taxonomy developed by Gottesfeld-Brown (1992), three major groupings can be distinguished. The first group are techniques that address variations due to differences in acquisition. These differences (e.g., camera movement) cause images to be misaligned. They can be addressed by global or local translation techniques. The optimum translation technique for a particular application is determined by knowledge about the nature of the variation. For example, aircraft roll during sensor operation would be corrected using a series of local translations, typically with specialised software.

The second group of registration techniques also focus on image variations caused by differences in acquisition, but here, the differences cannot be so easily modelled. These include lighting differences, atmospheric distortion or perspective distortion. The third group deals with variations due to differences of interest, such as change, growth or movement. Overall, the set of registration techniques chosen for a particular application must address the first type of variation while overcoming the second (which make an exact match impossible), and also avoiding removal of the third—the information content.

In this project the registration problem falls under both the first and second of these major groups. The purpose is to register an oblique image of a scene to an oblique representation generated from a planimetric database. Differences between the two scenes will be caused primarily by differences in viewpoint position, angle, and frame size. Both perspective distortion and differences in acquisition (i.e., angles and viewpoints) must be accounted for. Other sources of registration error, such as DEM and other data uncertainty, must also be accounted for in the process.

Both registration and data entry using an oblique perspective view generated by a GIS require that the image be accessible for co-ordinate input. This technology is not current part of any GIS. However, such systems are in place in other disciplines.

3-D Digitising

Almost all terrain-oriented GIS are capable of generating perspective views of a surface. However, none incorporate the ability to use this representation of data as an input device to enable the registration and digitising of oblique images. These oblique images, including still photos and video frames, therefore represent an untapped source of spatial information.

There are several reasons why oblique digitising presents a problem for typical GIS: 1) continuous variations in scale from the foreground to the background will create lines with varying levels of uncertainty; 2) hidden areas (terrain shadowing) must be explicitly dealt with; 3) registration of the image requires visual, rather than simply numeric alignment; and, on a more basic level, 4) with few exceptions GIS are not three-dimensional—the Z value is simply an attribute. All three-dimensional procedures require translation from a 3-D conceptual space to the '2.5-D' spatial attribute space.

The concept of three dimensional digitising is not new—it has gradually become an established technology in the area of computer aided design (CAD), biology (Uenohara and Kanade 1995), and nuclear medicine (Pietrzyk *et al.* 1995; Rusinek *et al.* 1993). Nevertheless, it remains both an art and a science (Wohlers 1997). Typically, specialised workstations are used to perform on-screen manipulation of images. Structures are stored as full 3-D representations in memory. Digital manipulation of the object requires the computer to capture actions not only in the 2-D plane of the standard screen, but also using a pseudo depth. The user's action (e.g., clicking a point on screen) is translated into a vector that is extended into the 'depth' dimension until it intersects an object. A variety of translation functions are required for even the simplest manipulations. There are separate co-ordinate systems for each object set, the object environment, display, and the workstation itself. Hardware and software designers must also utilise the 'art' side of the system to make the 3-D environment comfortable and intuitive.

Many of these problems are simply technical in nature. However, the variations in uncertainty represent a more fundamental problem. There are uncertainties in the database used as reference, and uncertainties in the data to be digitised from. Rendering these data in an oblique display will create continuous variations in uncertainty due to variations in visual depth. For example, although uncertainty may be constant throughout the database, lines further away in the visual field will have narrower uncertainty bands around them. Secondly, the accuracy of 3-D digitising decreases as field depth increases, creating a complex visual environment in which uncertainties vary continuously, as does the 'product' of these uncertainties: the digitised line.

This complex environment will require unambiguous uncertainty visualisation tools to allow the operator to 1) decide what is a datum and what is an artefact of uncertainty; 2) decide what is an appropriate scale for visualisation to perform the most effective digitising, and 3) to decide how far 'into' the image a digitising session should go (i.e., what is the limit of acceptable target distance).

The combination of these three factors—an aerial video source, image fusion and uncertainty visualisation—make possible the effective combination of oblique and planimetric data in a GIS. Note that this term 'effective' refers to how useful the resulting data are for analytical purposes. For purposes of *visual* comparison over time, it may be of greater use to store the information in a multi-media database.

Purpose

The system developed herein has two principle purposes: 1) in the context of the work presented in this dissertation, the practical purpose of the system is to update a landslide database in areas of high uncertainty, leading to a highly accurate, current database of mass wastage and feeding into the work described in Chapter Five; and 2) at a general application level, to demonstrate a practical application of uncertainty visualisation and management—the development of a tool that makes use of a data source for GIS that would, otherwise, be merely descriptive.

An inventory update system capable of both tracking and capturing uncertainty is developed herein. The system is termed the Oblique Data Fusion System (ODFS). Its practical purpose is to enable planimetric inventory data updates from low-resolution oblique data sources, while tracking uncertainty through the process. The following sections of this appendix focus on development steps, problems encountered, tests applied to validate information, and tests of system application to various types of objects. The results of applying the tool in a terrain inventory update will be presented and discussed in Appendix E; however, as development occurred in concert with this application, much of following sections will employ examples from that study. The development study area, Lyell Island, British Columbia, is introduced in detail in Appendix E. Here it is deemed the 'test area'. The landslide areas used for the test are here deemed 'targets', as they represent any visible object with distinct boundaries.

Methods

After performing initial tests to determine the viability (i.e., the level of accuracy that could be obtained) of registering oblique images and GIS-generated perspective scenes, the following methodology is utilised:

1. gather and choose appropriate frames from video source;
2. capture, translate and view one frame;
3. load nadir position from GPS data file;

4. generate a perspective view of the planimetric database using default nadir parameters and overlay the wire-frame perspective view on the video frame;
5. interactively manipulate the digital viewpoint until visual registration is achieved;
6. visualise vector location uncertainty;
7. decide on validity of registration; if invalid, either repeat from (4) or choose a new frame and repeat from (1);
8. if new data are present (i.e., outside of uncertainty bounds) digitise new linework;
9. gather localised information on registration uncertainty, apply registration uncertainty information to new linework, and store in vector database;
10. merge new linework into original data and update topology as appropriate for the data type;
11. if appropriate, verify linework and uncertainty by repeating from (2) using a different image of the same feature;
12. continue from (2) with a new feature.

Each of these steps is detailed in the following sections.

Initial Tests

The development of the Oblique Data Fusion System (ODFS) followed a series of preliminary field tests to determine the potential accuracy of the tool. A series of mass wastage zones, visible from aerial photography, were surveyed in the test area. Video frames were acquired from ground stations at various angles and distances from the targets (Figure D.1).

Aerial photos from several time periods (1977, 1980 and 1990) were converted into orthophotos using TRIM planimetry (See Chapter Three for a discussion of these data source) and survey control points as the base. Mass wastage zones were digitised directly from these orthophotos. Digitising uncertainty was estimated at $\epsilon = 1$m. When an orthophoto is skewed to a reference database there are areas of high and low correspondence. Although it would be ideal to capture this information and carry it through into later processes (so that it is possible to know the skewing uncertainty at any given point in the image), proprietary software routines make this task difficult. Lacking this information, a conservative estimate of the overall skewing uncertainty can be taken from the maximum error value. This number is deemed 'conservative' due to the fact that logged areas are typically rich in accurate control points, allowing maximal skewing accuracy in such zones. The overall epsilon distances (one SD for a normal distribution) for the images are: 1977: 3.1m; 1980: 2.9m; 1990: 1.9m. The production of the images is described in detail in Appendix E.

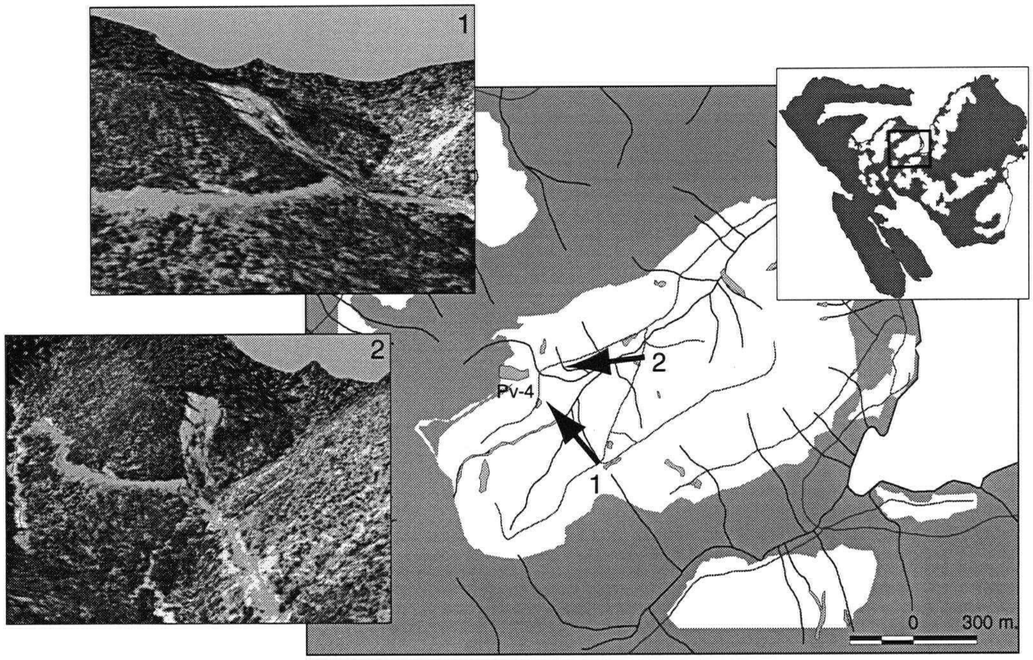


Figure D.1. The initial test site, located near Powrivco Bay, Lyell Island. Slide Pv-4, one of the slides tested, is shown from two of the base locations used to compare image geometry. Shaded areas in the planimetric views represent forested areas.

Using a GIS-derived perspective view based on the location data from the ground stations, the angle, zoom and other display factors (for the perspective view) were adjusted to bring the video image and the display into maximal correspondence (Figure D.2). Deviation values from key points were estimated on-screen (Figure D.3). The procedure was repeated with a second viewpoint (i.e., a new ground station and a new GIS-derived perspective view) and results compared. Over 95% of the lines (by length) fell within the epsilon distance of the most detailed image (1990). Additional statistics of correspondence are not available due to the non-Gaussian skewing uncertainty and the small number of samples available. Estimates of polygon area from the two viewpoints were within 7%. In comparison, digitising error could vary the area by as much as 5%, and slope area corrections would add 14% to the polygon area. Standard photogrammetric analysis (forest openings at 1:40,000) involves area uncertainty of approximately 6-15% (Ministry of Forests 1995). Two other test sites provided comparable values. Therefore, uncertainty in the 3-D digitising and oblique registration procedures for the test sites were judged to be well within tolerable limits set by data manipulation.

Video Frame Acquisition

Although any type of oblique image would be useful for data input, video frames were chosen for several reasons:

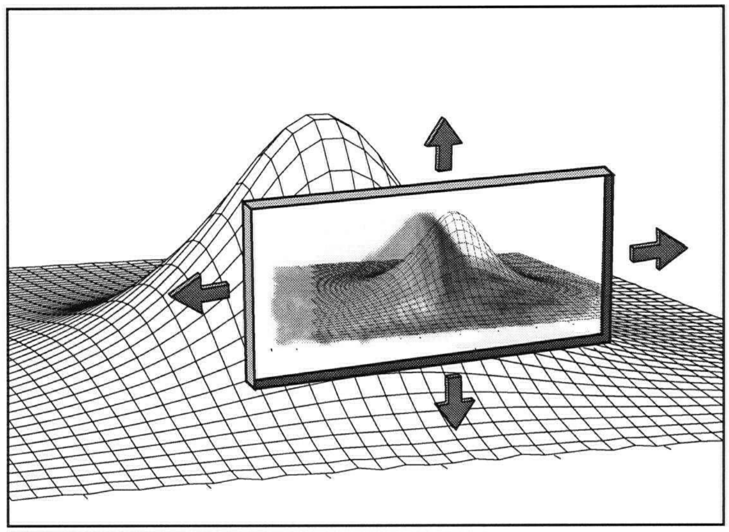


Figure D.2. Matching a video frame (inset) with the digital representation of the terrain. Changing the perspective view is equivalent to moving, rotating and warping the image window.

1. location information can be stored with the video data, or easily added during post-processing;
2. a continuous series of frames are available, so that any vibration-damaged images have many replicates;
3. image stabilisation technology is available at much lower cost for video cameras than for still-image cameras;
4. video frames can be directly translated into computer image formats.

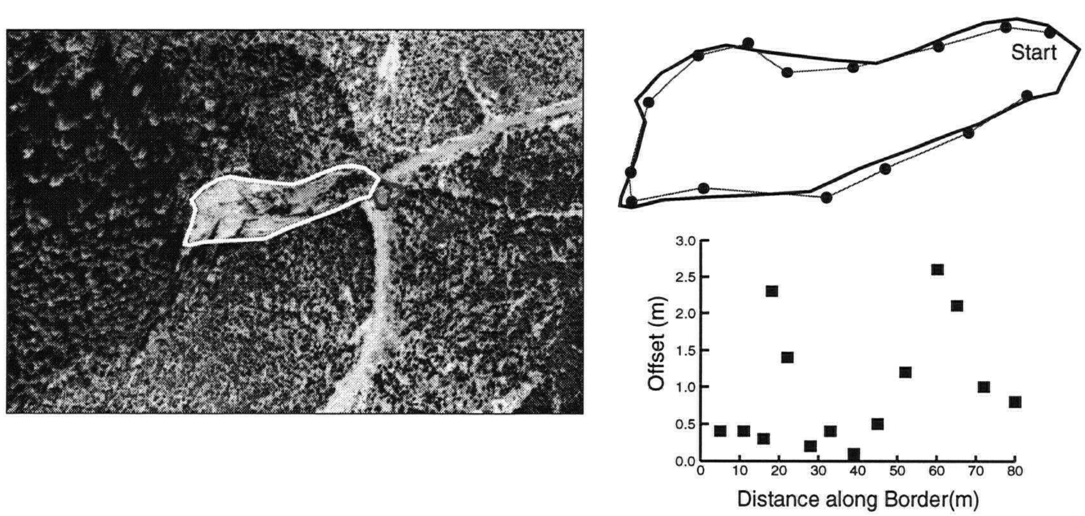


Figure D.3. Linework captured from video is compared with field sample points for slide Pv-4.

The primary disadvantage is the (relatively) low resolution of the images. An 800x600 pixel image was the maximum available after translation to computer image format. This disadvantage is offset by optimising distant-to-target as much as possible. It is also offset by (2) above.

The video data were gathered using a Super-8mm one-CCD camera, with a time-code link to a recording GPS. Both differential (Magellan M10) and standard (Trimble Scout) GPS were utilised in the trials. The GPS data logging was performed on both the GPS and a laptop computer. A position was stored every second, using output sentences that included position, heading and elevation (Figure D.4).

Helicopters utilised in the trials included both aluminium and carbon fibre bodies. The latter was found to be far more conducive to GPS data reception, given that aircraft regulations require internal antennae be used for ancillary equipment. However, in this project the GPS reception was found to be adequate (99% data recovery) for the aluminium frame as well. In other trials, using less sensitive GPS antennae, the aluminium airframe was found to block signals as much as 30% of the time.

Although a single pass over each zone being imaged would typically suffice, in these trials a number of passes were taken at various elevations and target distances. Later evaluation showed that the camera/aircraft configuration that leads to maximum registration accuracy is (Figure D.5):

- A maximum aircraft speed of 25-30 knots (slow enough to reduce buffeting).
- The aircraft positioned at a point where a maximum number of reference points are in view, balanced with moving as close to the target as possible. These reference points include (in order of utility) roads (particularly in the foreground), ridgelines (above the target and in the background), rivers, and other planimetric data (e.g., cutblock boundaries). Coastlines were found to be less useful due to tidal variations. Running at just above the local maximum height

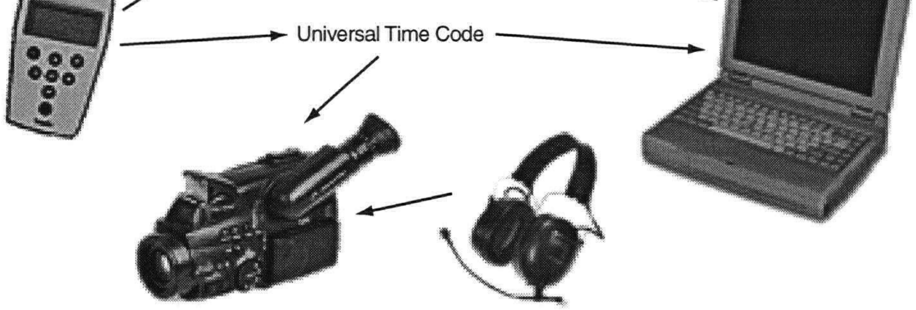


Figure D.4. Schematic of the data gathering system for aerial video data.

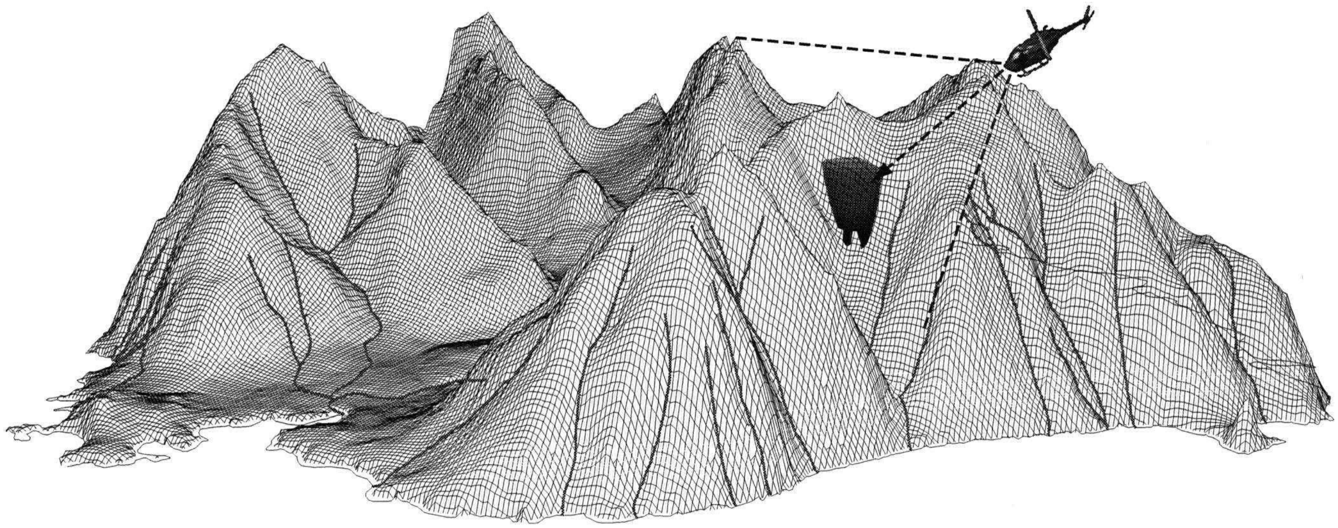


Figure D.5. Maximum registration accuracy is achieved using a balance between minimising target distance and maximising visible reference points.

of land and approximately 800-1000m away from the target proved to be the ideal zone given the conditions in the test area.

- A camera technique that involved both wide shots to establish reference points, coupled with zooms to capture detail. Although zooming is useful, if the aircraft is too far away from the target the image can easily be degraded due to sun glare, air disturbance or clouds and mist. Zooms of over 10X also reduce the usability of images due to camera vibration.

Translate and View Frame

Frames were captured from videotape and loaded as computer images, referenced by Greenwich time. GPS data were downloaded, differentially corrected if required, and stored in a tabular format. Frames were chosen so that every target could be viewed from at least two positions if possible. Frames were down-sampled to 640x480 without appreciable loss of quality (image blurring presented more of a problem than frame size).

Although there was some supposition that the small lens of the video camera would introduce distortion into the image, tests (using various lens settings and a controlled grid target) showed that minimal distortion occurred at common lens settings. At long zoom settings the distortion in the image periphery was noticeable, however such lens settings were not used in the fieldwork due to vibration blur. In other applications using longer lens settings a focal length-based correction might be required.

The geographic information system utilised for development of the ODFS was ARC/INFO. Display and user-interface oriented procedures were developed in the Arc Macro Language (AML), while other procedures were developed using C language subroutines. All relevant code is available on the disk. The system was developed over a six-month period, with approximately four months related to development and two to testing and data input. The orthophotos and databases discussed in this appendix and Appendix E took approximately six months of additional effort to develop. The only other researcher directly involved was a digitising assistant who independently duplicated the slide zone identification and digitising to assure accuracy in both crucial tasks.

Load Nadir Position

The nadir position of each frame is available through cross-reference to the GPS data file. These values are loaded automatically the first time each frame is loaded into the system. Subsequently, preference values detailing refined location, viewing angles and other information are saved and loaded in place of the nadir position.

The original information loaded from the GPS file also contains aircraft heading information. These values are converted to camera heading (normally -90°). The downward inclination of the camera was separately gathered from the soundtrack of the tape (estimated using an inclination device attached to the camera), which had been patched into the internal aircraft communication system. A second value for camera heading was also taken from the soundtracks. However, the GPS information and an estimated -10° camera angle were found to provide sufficient information to initialise the registration procedure. Nevertheless, the audio system was found to be useful for tracking other ancillary information.

Generate Perspective View

For each image a wire-frame perspective view is then generated using the values loaded above. Initially, a large default area of the DEM is chosen for viewing (a square with a width five times the distance to target). This area can then be widened or narrowed based on target parameters. The perspective view is generated on top of the video frame. The ability to toggle on/off each layer of the database is provided, as are controls for grid density. As each target is in a unique situation, each perspective view is configured differently. Tests showed no ideal initial configuration (Figure D.6).

Registration of Image

The image is registered to the database through interactive manipulation of view parameters in order to achieve a visual match between the two. Two methods were attempted: manual interaction through graphical user interface (GUI) controls (buttons, sliders, etc.), and automatic registration through iterative procedures. Manual controls were developed first, and then automated procedures were investigated.

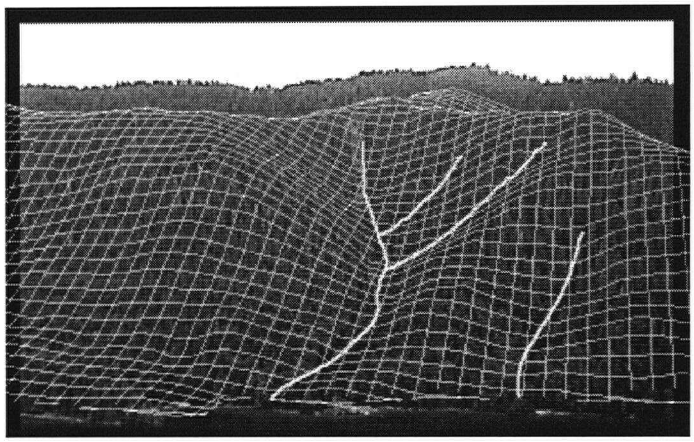


Figure D.6. The initial configuration of a frame. The unique situation of each target allows no standardised configuration to start the registration. Some frames are close to their final registration, while others require considerable adjustment.

Manual controls were developed which directly accessed view parameters and allowed for fast, interactive changes. The controls include: zooming, panning, tilting, movement in 3-D space, target definition, as well as changes in elevation model limits and resolution. These latter items affect the speed of display as well as the availability of background information to assist in registration. For example, a typical registration session is presented in Figure D.7.

Ideally, this type of registration procedure would be automated. However, standard registration methods do not offer ready-made solutions to this problem. The primary area of application where this type of problem arises is in computer vision, where it is termed 'viewpoint registration' (Keller 1997). In these robotic applications the general problem is the registration of images taken from different viewpoints, with the application being depth or shape reconstruction of objects. Here, the methods primarily focus on feature correspondence. When a plane is the target, the problem is somewhat simplified. However, with 3-D targets, perspective and occlusion become problems (Figure D.8).

Neither global nor local methods of transformation are applicable to this type of registration (Gottesfeld-Brown 1992). The changes in perspective that occur when the viewpoint is shifted are not amenable to direct calculation (in feasible time). Iterative procedures are called for. The target of an iterative procedure would be to bring a series of point-pairs (defined by the user to represent 'from' and 'to' locations in the database and image respectively) into maximal correspondence through manipulation of viewpoint parameters. However, though the goal is easily defined, the route to the goal is not necessarily linear. Both perspective and terrain distortions make the goal a complex one, particularly in extreme terrain.

There are a variety of search algorithms that have been developed or adapted for image registration (summarised in Table D.1). The implementation of one of these strategies to this unique situation is a non-trivial problem, but an obvious goal if the system developed in this project is to be applied to a large number of images.

In this initial implementation an iterative procedure was implemented that focuses on exploiting the spatial relations between features in a manner similar to the 'relaxation' method, though with

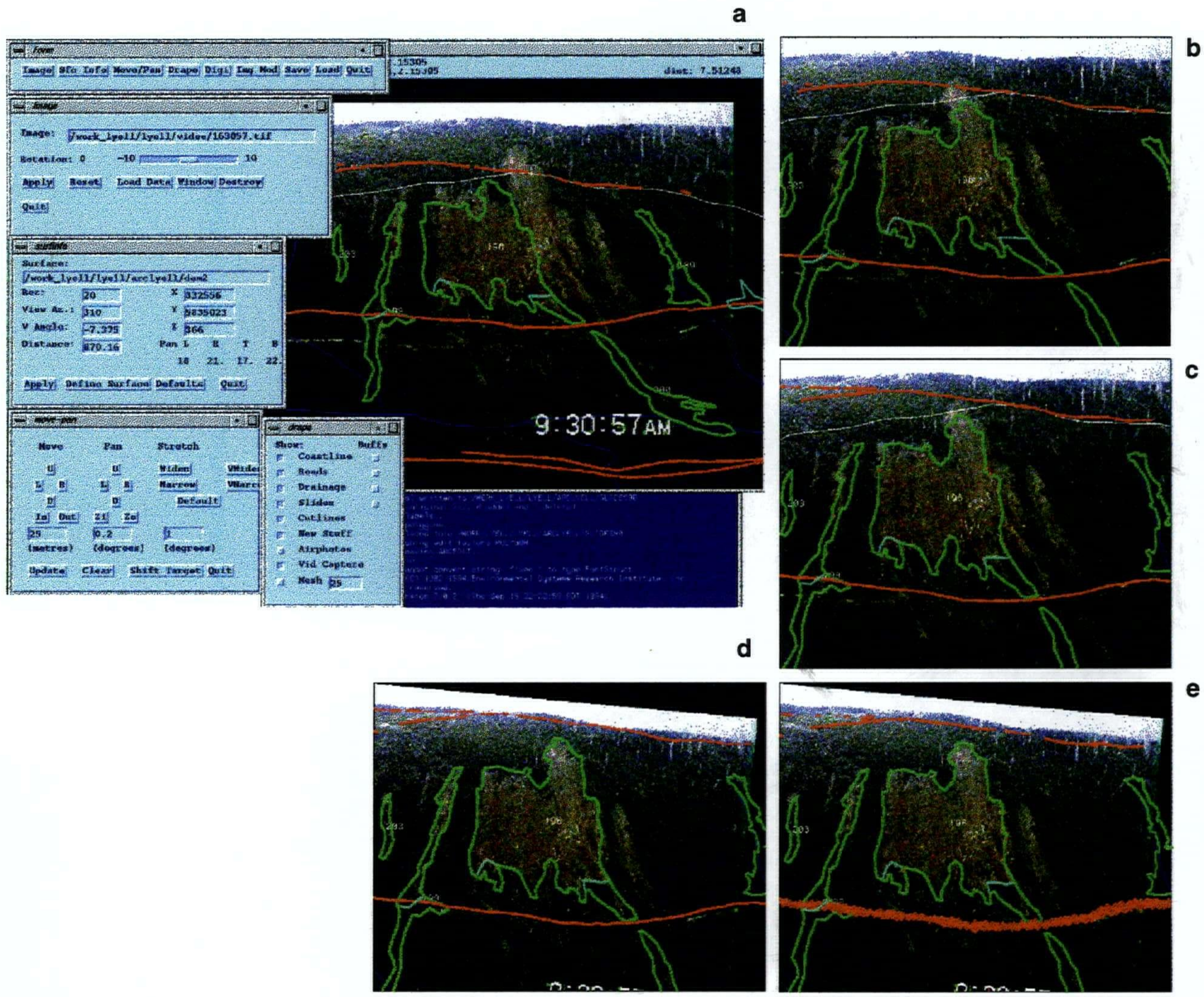


Figure D.7. An example of image registration (simplified). Red lines indicate roads, green - slide boundaries and white - cutblock boundaries (turned off in d and e). From the initial position (a), the operator first moves the viewpoint to the right (b), then gains altitude (c). The image is then rotated (d) to complete the registration. In (e) the uncertainty in road vector location is visualised, indicating that the registration is generally within specified tolerances. The visualisation in (e) provides feedback as to where the session should end (i.e., return for effort expended begins to decrease). Lines are exaggerated for clarity.

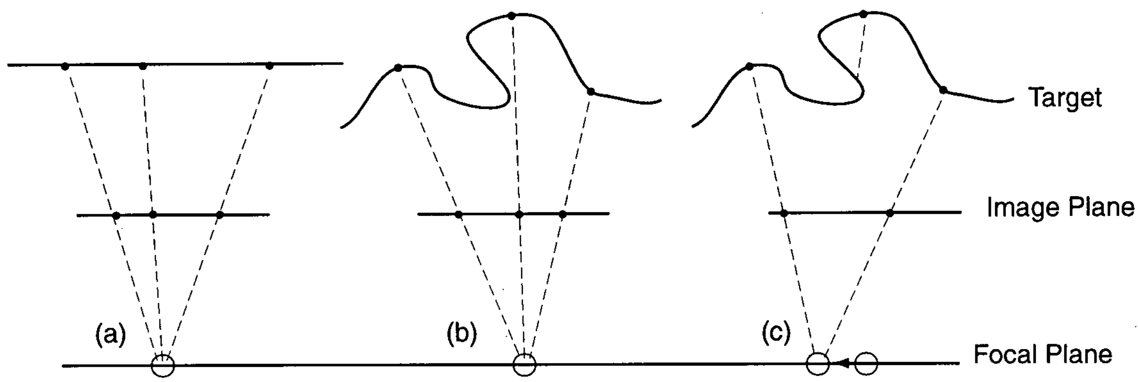


Figure D.8. Although the arrangement of points on a 3-D surface will not change when the viewpoint shifts, their *relative* spacing will vary in a non-linear fashion. Points will also appear and disappear due to occlusion by intervening terrain.

Search Strategy	Advantages and Applications
Decision Sequencing	Improved efficiency for optimising rigid transformations
Relaxation	A practical approach for finding global transformations when local distortions are present by exploiting spatial relations between features
Dynamic Programming	Good efficiency for finding local transformations when intrinsic order for matching is apparent
Hough Transform	For shape matching of rigidly displaced contours through the use of 'dual-parameter space'
Linear Programming	For solving systems of linear inequalities, used for rigid transformations in point matching
Hierarchical Techniques	Applicaable to speed up many other approaches by guiding a search through progressively finer resolutions
Tree and Graph Matching	Good for search minimisation using inexact matching of higher-level structures

Table D.1 – Search algorithms for image registration (adapted from Gottesfield-Brown 1992)

a local transformation focus. The procedure accepts as input a series of point pairs from the image/perspective view that define from-to targets. It then uses a search order defined through observation of an operator using manual methods to produce target correspondence. It is purposefully designed as a 'forgiving' procedure (i.e., 'relaxed'), in that small deviations from the goal do not immediately force the abandonment of the current configuration. The general algorithm is as follows (Figure D.9):

1) Define target point pairs and step sizes, STEP;

2) Calculate ground distance between pairs and produce an overall RMS statistic: TOTAL_ERR;

3) Move the viewpoint STEP in direction –Z, and recalculate TOTAL_ERR;

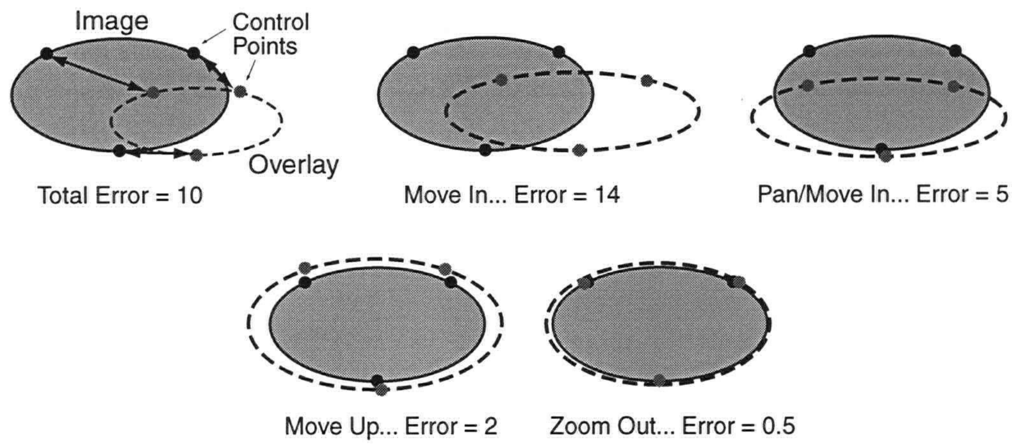


Figure D.9. A simplified example of the iterative search algorithm implemented for automated registration.

4) Repeat (3) ten times. If TOTAL_ERR continues to generally decrease, repeat (3) and (4). If TOTAL_ERR increases or stops decreasing, stop repetition.
5) Repeat (3) and (4) in all 26 (3-D) directions. Retain the position in which minimum TOTAL_ERR is calculated;
6) Pan the image STEP in direction LEFT, and recalculate TOTAL_ERR;
7) Repeat (6) ten times. If TOTAL_ERR continues to generally decrease, repeat (6) and (7). If TOTAL_ERR increases or stops decreasing, stop repetition.
8) Repeat (6) and (7) in all other 7 (panning) directions. Retain minimum TOTAL_ERR;
9) Repeat the procedure using 'zoom' (2 directions), 'tilt' (2 directions), and stretch (2 directions).
10) Repeat entire procedure from (3) until TOTAL_ERR is minimised.

The implementation of this algorithm made possible automated registration of the oblique image. However, the time taken to complete the registration procedure was far in excess of manual methods (on the order of 10 to 30 times). This difference is primarily attributed to the human capacity for visual processing. A human operator of the system can intuitively grasp the spatial relationship between the two scenes superimposed on the screen. The operator can then move directly to the registration goal in a small series of steps. The operator is also not deterred by an increase in overall error at intermediate stages, because they are aware of the path to the goal. The computer, however, lacks this intuitive sense, and so must explore a large number of spatial relationships while gradually approaching the goal.

For example, in Figure D.10 a typical situation the system operator might face is demonstrated. Figure D.10b represents a target video image, while (a) is the current GIS-derived wire-frame per-

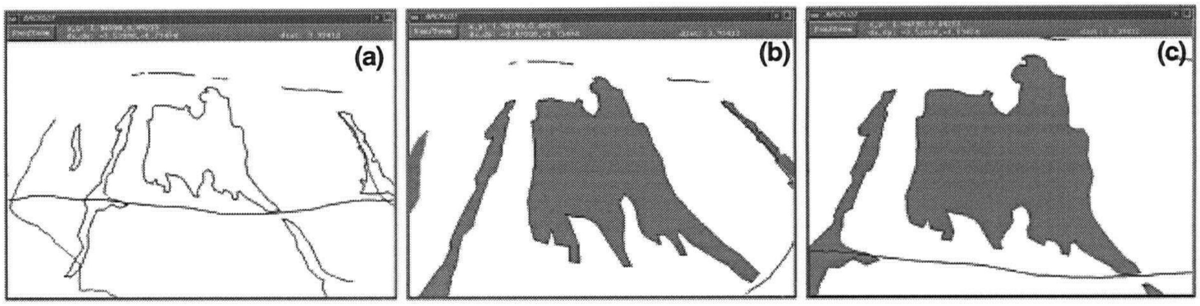


Figure D.10. Zooming and moving generate different image geometry. (a) Current viewpoint for reference; (b) Target image where a **zoom** (from (a) to this view) is appropriate (exaggerated for effect); (c) Target image where a **move** is appropriate.

spective view. The purpose of the procedure is to manipulate the viewpoint of (a) until it matches the geometry of (b). Visually comparing the two, the operator would note the differences in overall geometry, where the lower section of (b) is elongated relative to its' upper section (when compared with (a)). This simple clue is all the operator requires to ascertain that the target is located on a slope, and that the lower section is closer to the viewpoint than the upper section. 'Zooming in' the viewpoint in (a) will create the distortion observed in (b). However, if the target is instead (c), the lack of relative distortion indicates that a simple 'move in' would suffice to change the viewpoint of (a) to match (c). This entire process is an elementary function (i.e., based on intuitive processing) for a human operator, but would require numerous iterations of the algorithm presented above to produce the same results.

Intelligent search procedures and the introduction of expert systems and other learning algorithms would no doubt increase the efficiency of the automated procedure (e.g., Kontoes *et al.* 1993). However, such implementation is beyond the scope of this project. The automated procedure is currently part of the registration software, and is useful for 'fine-tuning' the registration of an image once the operator has manually performed a coarse registration.

In order to speed up the registration system, parallel processing was implemented for visualisation. The procedure utilised is as follows:

---

1) Make all data available to other network computers;
2) Poll the network for available GIS sessions;
3) If *X* sessions are available, start remote processes that generate the next *X* likely viewpoints (e.g., the most likely is a repeat of the previous step);
4) If the next step chosen by the user has been generated, load it directly;
5) If the next step chosen by the user is currently being generated, wait for processing to complete, then load;
6) Otherwise generate the next view locally, and repeat from (3).

The uncertainty in registration is quantified in a separate procedure discussed below.

Uncertainty Visualisation

The purpose of displaying uncertainty in this procedure is twofold: first, viewing uncertainty in the planimetric data used for registration will assist in determining if registration correspondence (i.e., RMS error) around the feature of interest has been maximised (RMS error minimised). If a section of a road in the image falls within the epsilon distance of its planimetric counterpart, further registration effort (for features in its vicinity) may be futile. Secondly, visualising uncertainty in the linework of the target (in this example the mass wastage boundaries) will enable the operator to determine if changes visible in the image are in fact new data or are simply artefacts of the database.

The data for uncertainty visualisation were produced prior to this procedure. Data were gathered from global values published as metadata for the source data layers, as well as from skewing error associated with images used to update these layers. The data represent the epsilon distance (one standard deviation for Gaussian distributed uncertainty), and are stored in an attribute field. The metadata and the collection procedure are discussed in more detail in Appendix E.

Several types of uncertainty visualisation are implemented in the system (Figure D.11). The simplest provides a crosshatched overlay onto draped linework with a width epsilon. In the example, the slide boundaries have a crosshatched overlay (green lines) that visually demonstrate the uncertainty in their position. This type of overlay enables image data beneath the crosshatch to be visible to the operator. A second method utilises dot density to indicate the probability of line location (red areas); providing more information about the structure of the uncertainty, but also increasing visual clutter. A third method using colour saturation bands was also implemented. Here, a user-defined number of bands are draw around the object, with the saturation of the colour indicating probability of line location (yellow). This third method produces a similar amount of information to the second method, while allowing some visibility of features occluded by the second. However, it substantially increases display processing time.

Each of these methods was found to be useful in certain situations. For registration, the first method—crosshatching—is most useful due to its display speed. The second and third are both useful in the 'target growth vs. artefact' situation discussed above. The second is most useful for optimising speed of display, while the third is best for complex uncertainty or where precision is crucial.

3-D Digitising

A 3-D digitising system was developed to enable direct input of new information from the registered image to the planimetric database. The system functions as follows:

A grid of x-y co-ordinates is drawn over the perspective view (drawing pen off) and the co-ordinates of each intersection, as well as the corresponding screen co-ordinates, are stored in a table. The density of the grid represents the resolution of the digitising session, and is user-defined. When the user digitises a line, each point picked on the screen is looked-up in the table, and the closest planimetric co-ordinate is entered into the (new) layer.

The system is based on the concept of ray-tracing, but rather than tracing each digitised co-ordinate individually, all calculations are performed on the grid before the actual digitising begins. Although more calculations are required than for individual tracing, this method is the only one compatible with the GIS access routines.

Figure D.12 shows the digitising of a new landslide into an existing database (1:1200 scale, 0.5m resolution). All lines between points digitised follow the ground contours.

All input is buffered, so the user can step back (i.e., undo) through the points already entered. Due to the availability of extensive editing facilities in the GIS, no on-screen editing has been implemented. Once input is complete, the user can edit co-ordinates in a separate window using a planimetric view.

Quantify Registration Uncertainty

One of the principle problems with 3-D digitising is the variability of uncertainty across the visible image. For example, a digitised line running from the foreground (target distance = 400m) to the background (800m) would see digitising uncertainty (epsilon) effectively double along its length. There is also the problem of registration uncertainty which, as discussed above, will vary across the modelled surface. There is no global or local spatial transformation function to draw these uncertainty values from, therefore a separate function is required to determine the spatial distribution of this uncertainty. In addition, the routine described below also effectively captures the variations in the epsilon distance function along the digitised line. It is illustrated in Figure D.13.

1. The user enters a series of points to define the general outline of a working area in the model (to minimise processing time).
2. Although the user has registered the wire-frame perspective view to the image as closely as possible, there will quite often be discrepancies between the two. Such discrepancies will commonly vary in magnitude and direction, depending upon the complexity of the topography, the resolution of the data, and other factors. There are certain loca-

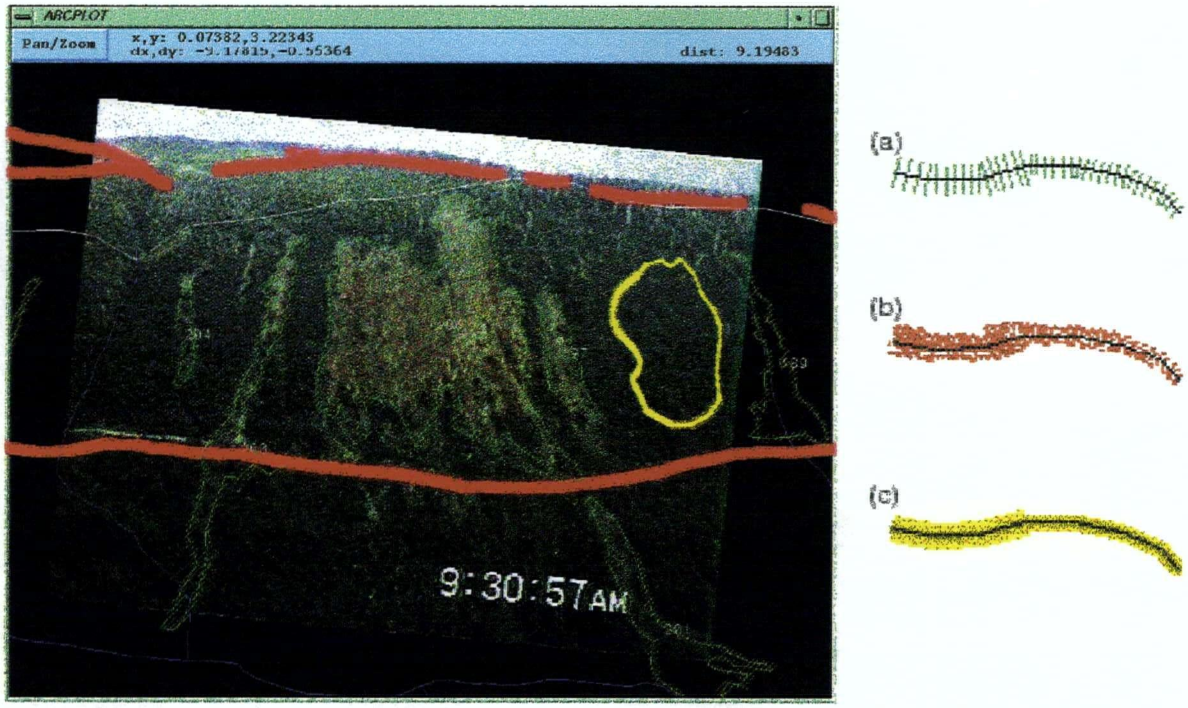


Figure D.11. Various methods of displaying vector uncertainty. (a) crosshatching; (b) dot-density; (c) saturation bands. Note: this 6-bit depth printed image considerably compresses the colour range of the original 24-bit screen image. The width of each band or crosshatch is also dependent on vector segment's distance from the vantage point.

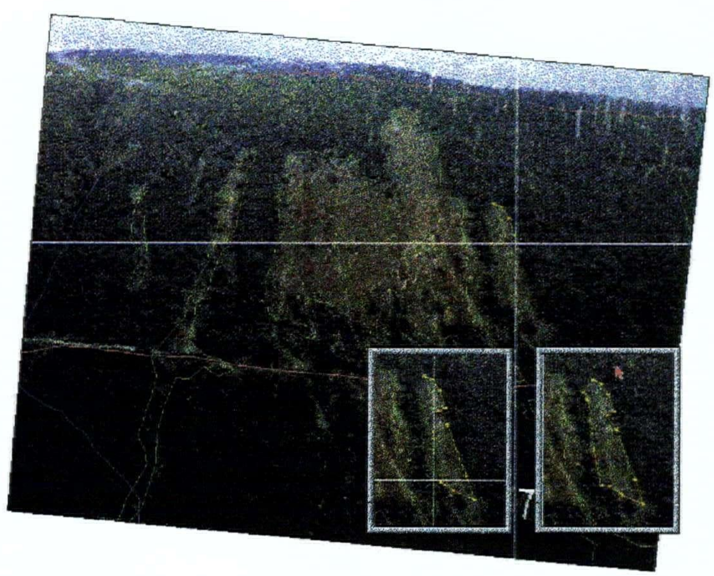


Figure D.12. Digitising a new landslide into an existing database. Left inset - in progress; Right inset - completed polygon.

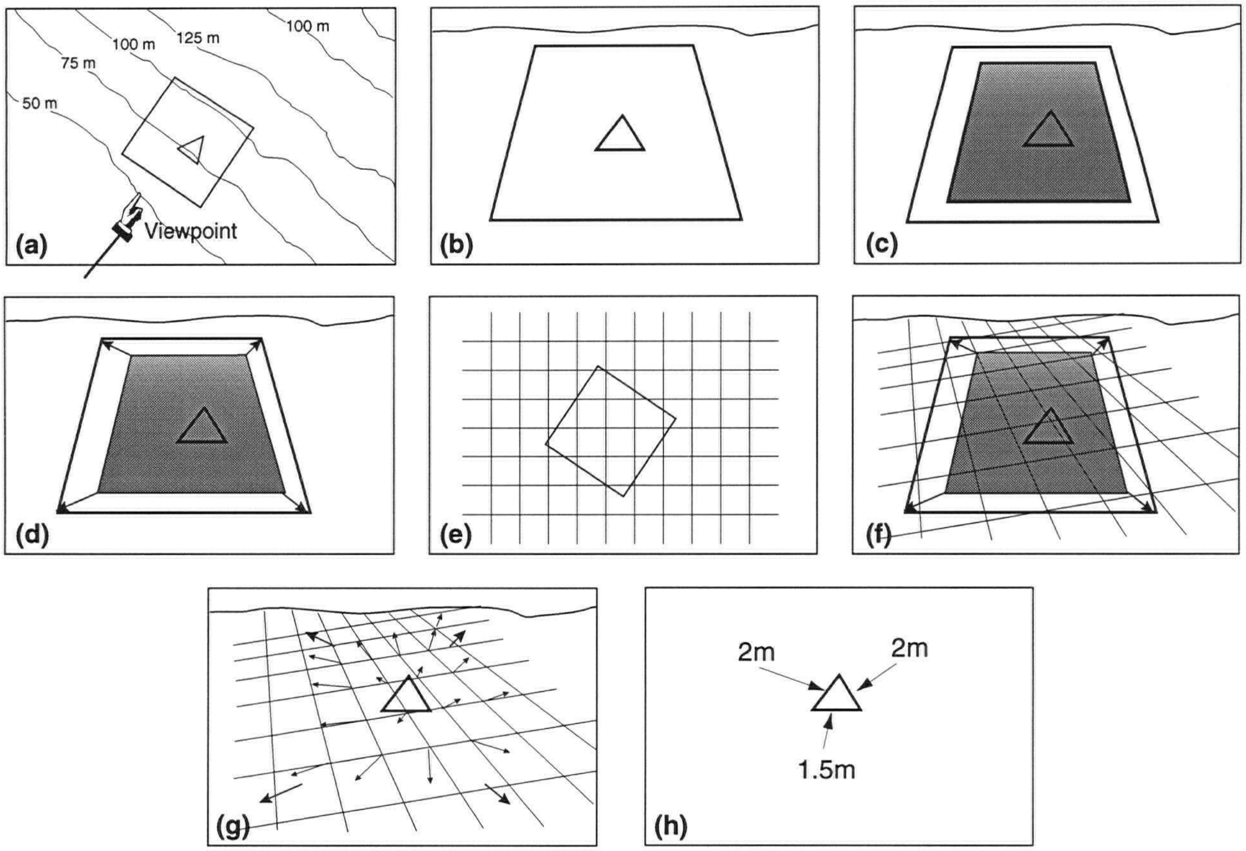


Figure D.13. Skewing procedure to quantify the registration uncertainty. (a) plan view of a well-defined existing object (square) and a new object (triangle); (b) perspective view of the objects; (c) the image of the square object is added (the wireframe is drawn over the image—the image is not draped); (d) point pairs are chosen between visible corresponding points; (e) a grid is generated; (f) the grid is draped (for illustrative purposes only); (g) each point in the grid is skewed using a distance-weighted function; and (h) numerical skew values are applied from the nearest grid point to the new object.

tions in the image/wire-frame overlay where the discrepancies are obvious, such as river confluences, road intersections, or the corners of polygons. In order to capture these discrepancies, the user now enters a series of point-pairs, in which the first point refers to the image, while the second refers to the corresponding position on the wire-frame (Figure D.13c).

3. The two points in each pair are converted from screen co-ordinates to map co-ordinates through reversal of the projection process (similar to 3-D digitising).

4. A north-south, east-west grid is generated (within the working area at a user-specified resolution; Fig. D.13e-f). Each point-pair's 'from' point is matched to its nearest grid intersection. The vector (i.e., the distance and direction) of the pair is then used as one input to a skew function. This is repeated for each point-pair, and then the skew func-

tion is applied to the entire grid. This 'skew' consists of all grid intersection points being shifted based on a distance weighted vector sum of all point-pair vectors (Fig. D.13g).

5. This new 'skew grid' is then compared with the original grid; the offset at each grid point represents an estimated epsilon variance for that area of the current image/wire-frame overlay.
6. These values are then applied to each line segment of the newly digitised line (Fig. D.13h)

Ideally, each cell that the line segment passes through would be averaged (weighted by sub-segment length) to derive this overall uncertainty value. However, to simplify the procedure (and to speed it up to an acceptable rate) the grid value at the midpoint of each line segment is taken as 'typical' for the entire segment. Given the digitising point density used in this application, the use of a midpoint is roughly comparable to the ideal value (as judged through trials). In applications where line segments are typically longer it would be necessary to either break the segments up (e.g., using a 'densify' type of procedure) or alter this current procedure as discussed above.

The effectiveness of this uncertainty assignment procedure depends upon both a careful choice and an adequate amount of point pairs. They must be defined in all working areas of the image, in areas of both high and low registration error. Users will typically focus on areas of high variance due to experience with standard planimetric registration point-pairs. A minimum number of pairs is therefore required by the procedure (8-20 depending upon grid size and density).

For example, Figure D.14 shows the sequence of capturing uncertainty in the slide digitised in Figure D.12. Frame (a) shows the user defining the working area for the skewing grid. Frame (b) shows the point-pair entry (the number of pairs has been minimised for legibility), and (c) shows the line variability (epsilon distance) displayed using a simple fill.

Merge and Update Topology

A standard planimetric database editor is used to clean up digitised linework and join it with original data. At this point polygon topology can be updated if desired. In Figure D.15 the update of existing polygons using the ODFS is illustrated. Uncertainty data from the above procedure are now stored in the vector database, and can be accessed with the uncertainty visualisation routines.

Verification

The validity of the linework generated by the ODFS was verified through several means. Ground survey, comparison with other representations, and several statistical techniques were utilised. During initial trials high-resolution (1:10,000) orthophotos were used to generate a database of

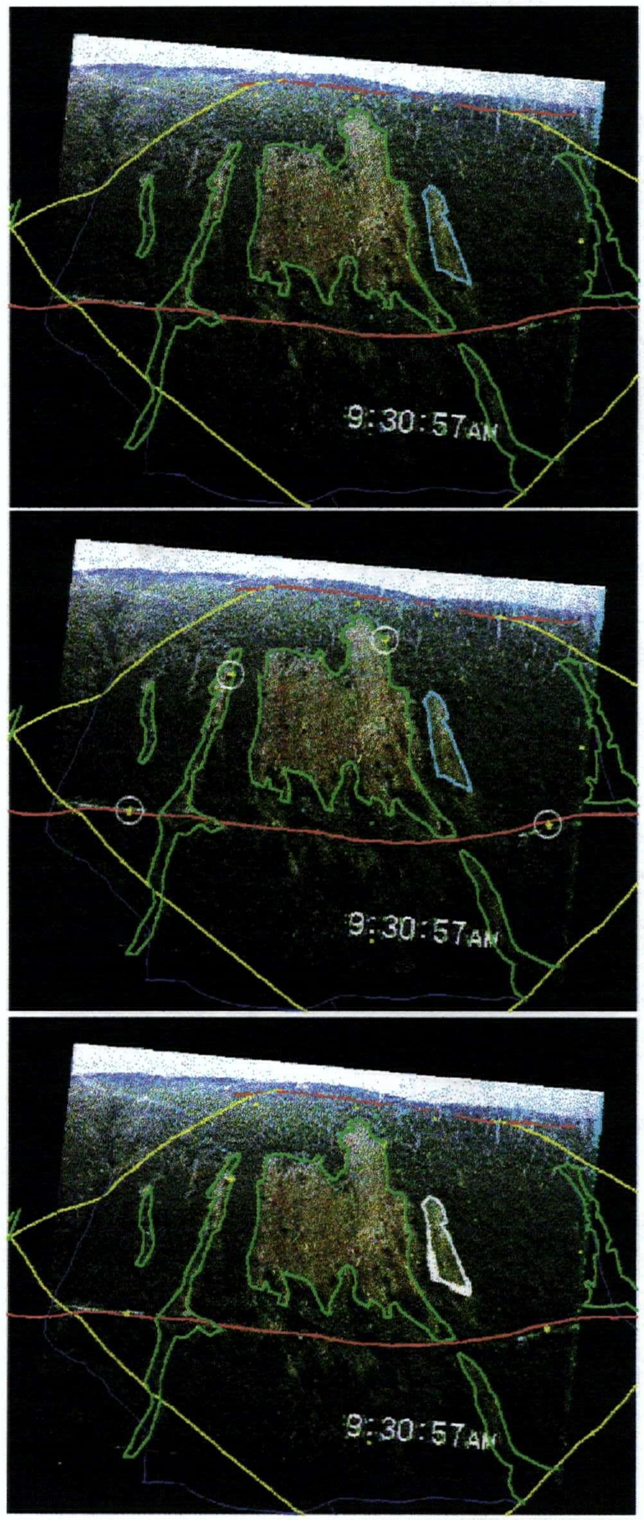


Figure D.14. Steps involved in determining local skewing error. (a) Boundary points are defined for the working area. (b) Point pairs defining from-to positions are defined in all areas of the image. (c) Line blurring displays the results of the procedure. All lines have been widened for illustrative purposes.

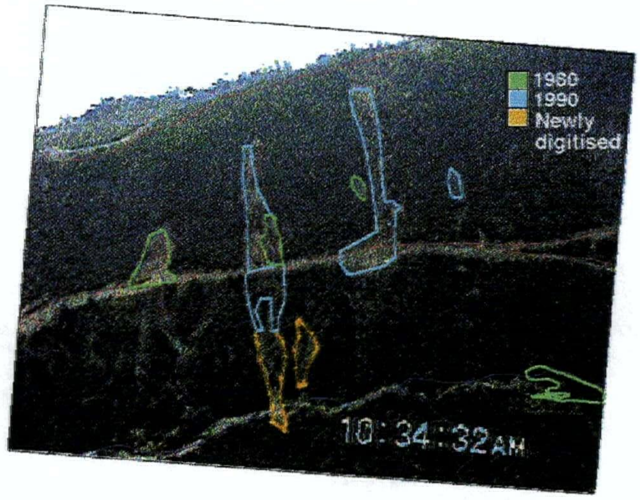


Figure D.15. Using the ODFS to add vectors to existing polygons. The new lines are displayed with skewing uncertainty saturation bands.

mass wastage polygons for comparison. In Appendix E the construction of a substantial temporal database of mass wastage polygons from digital orthophotos will be described. Here, a subset of these polygons is used for comparison with the ODFS generated data. The orthophotos are registered to the provincial base mapping system: TRIM data (British Columbia terrain and planimetric inventory data). The RMS error in generating the relevant orthophotos is ~ 2.3 m.

Although the methods used to derive the orthophoto uncertainty value are described in Appendix E, presentation of some details at this stage is appropriate in order to facilitate understanding of the verification procedures. The software used to generate the orthophotos ('Orthoengine') reports error values for each control point, describing the mathematical convergence of the orthophoto model relative to the reference data. An overall summary is also provided. In registering the orthophotos, although summary (mean) error was held to less than four metres, it was apparent that this number was inflated by poor control point correspondence in heavily forested regions. Therefore, trials were undertaken in which orthophotos were generated for specified subsets (typical mass wastage regions) of the study area. The areas relevant to this study were registered with a maximum RMS error of ~2.3m. This number is therefore offered as the skewing uncertainty of locations derived from the digital orthophotos.

Ground Survey

Surveys were conducted in the test area using differential GPS to establish ground control, and standard survey techniques to complete all measurements. Seven medium to large well-defined mass wastage zones were surveyed. Two of these areas were withdrawn from verification work due to poor definition from aerial surveys (which would lead to misleading comparisons), and one area was truncated for a similar reason. The survey ground control was tested for accuracy against a known survey station, and a point location accuracy level of 2.8 metres (circular map accuracy standard; CMA) was derived (Appendix B details the methods used).

Differential GPS ground control was utilised for the surveys due to a lack of readily available control points. The only existing provincial survey markers in the survey area are located at least three kilometres from the survey sites across difficult terrain. Ground control was established for each survey site individually using a GPS shot at the instrument location and a second shot at some distance along a opportunistically established baseline (see Figure D.16). Using a total station, a series of sideshots were then taken as the flag-person made a circuit of the landslide. This method was utilised due to the extreme nature of the terrain; a closed traverse, though more accurate, was considered too dangerous.

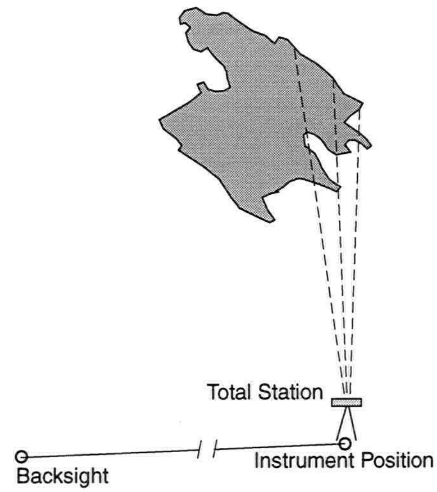


Figure D.16. Survey methods utilised for surveys of mass wastage boundaries.

In this situation, distances from the instrument to the sideshots and relative angles *between* the sideshots are high precision measurements (limited by instrument accuracy and operator error). However, the accuracy of the baseline used to establish *absolute* position of the survey points is much lower. For example (Figure D.17), for a given baseline and an equidistant target, the positional error of the target is equal to the accuracy level of the GPS positions. An increased baseline length serves to reduce this number. Yet in the extreme terrain being surveyed it was not always possible to establish a baseline substantially larger than the target distance. In fact, with the target distance commonly over one kilometre, it was at times necessary to establish a baseline that was shorter than the target distance.

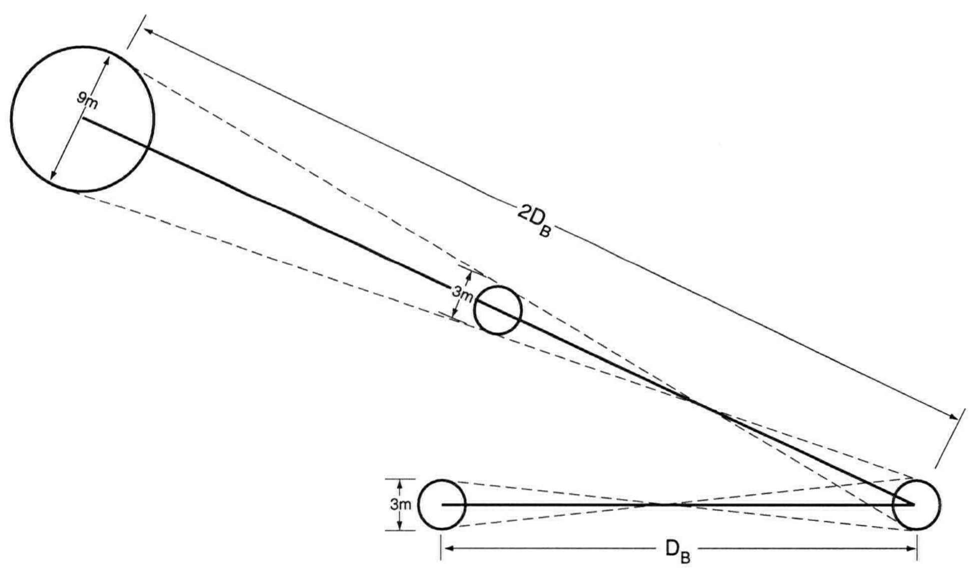


Figure D.17. Uncertainty in the survey baseline can lead to substantial uncertainty for survey shots, but only in reference to the overall relative angle of the survey. Distances and relative angles remain as precise as the instruments and operator error allow.

This problem was addressed in co-ordinate processing through rotation of the entire survey. The co-ordinates were rotated to minimise variability between the survey co-ordinates and the planimetric database. As the target of the overall procedure is to maximise the *relative* accuracy of lines, this method was deemed acceptable for this application. All rotations used were well within the limits of angular error for each baseline situation.

The epsilon bandwidth statistic is used to compare fusion system digitised lines and surveyed lines. All indistinct slide boundaries (i.e., not easily defined from imagery or aerial photos) are ignored; all others are compared in two ways. First, survey points are plotted as perpendicular offsets from the fusion system digitised lines (Figure D.18). These results indicate that 90% of points fall within ±2.5m. However, the points are not chosen randomly; they are intended to act as vertices of straight arcs that approximate the slide boundaries. A more useful statistic will be one that compares the *arcs* rather than just the points. The epsilon bandwidth statistic (Blakemore 1983) is an appropriate way of performing this comparison. A extensive study of polygon boundary accuracy measures by the BC Ministry of Forests (Ministry of Forests 1995) noted that epsilon bandwidth was the most appropriate and useful method of summarising boundary accuracy. The statistic is generated by standardising the entire polygon boundary and the surveyed boundary to the same length, subsampling at appropriate intervals, and plotting the offsets. A line on the graph that contains 90% of the offsets is defined as the epsilon bandwidth. Other statistics for boundary inaccuracy such as MacDougall's (1975) *H* statistic (based on standard error, line length and map area) rely on an assumed normal distribution of error. As this is not the case, the epsilon bandwidth is considered the most robust (Chrisman 1989).

The epsilon bandwidth statistic is calculated for the survey and polygon data, using one metre subsampling. Figure D.19 shows a graphic representation of these offsets and the epsilon value. While the epsilon value approximates one standard deviation for a Gaussian distribution, it is more commonly used for data with variable distributions due to its robust nature (Lodin and Skea 1996). A histogram of the data indicates that the distribution is non-Gaussian (Figure D.20). This is supported by the differing values for epsilon bandwidth (2.3m) and the standard deviation (1.7m).

This calculated epsilon value refers to the deviation of the fusion system digitised lines from the best available information regarding where the slide boundaries are actually located. Survey, differential GPS and datum errors are also factors, but have not been

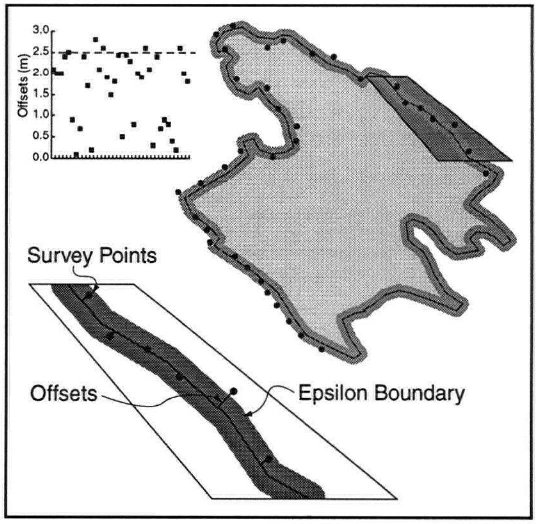


Figure D.18. Survey points plotted as perpendicular offsets from the digitised lines.

quantified in this calculation. Although such quantification would be helpful in determining the overall uncertainty level, it is the relative errors that are important in determining the utility of the ODFS. Overall, the epsilon value of 2.3m is comparable to the CMA standard for differential GPS of 2.8m determined through field tests. This indicates that, when optimal images are taken of well-defined (i.e., visible and good contrast) linear ground features, and registration is not constrained by a lack of visible tie-points, the accuracy of the ODFS is approximately that of a differential GPS survey.

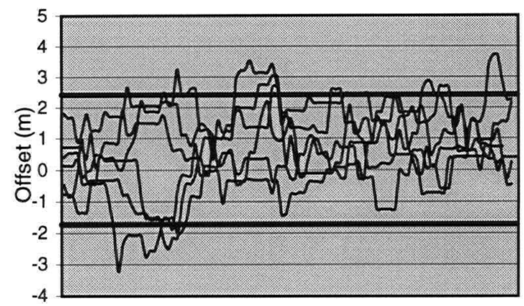


Figure D.19.. Epsilon values defined based on all perpendicular offsets.

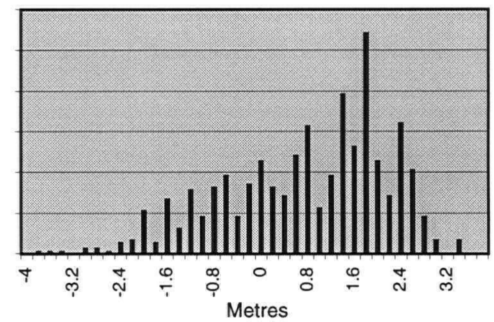


Figure D.20. The frequency distribution of offset magnitude.

Comparison of ODFS and Digital Orthophotos

Although statistical certainty tests for epsilon bandwidth have not been developed, it is clear that the number of comparisons available from the ground survey is insufficient to establish the validity of these statistics (57 original samples and ~400 line sub-samples at a 1m spacing). The TRIM data used for ground reference do not include mass wastage polygons, and are therefore of no use in evaluating the system in this application. Therefore, a second stage of evaluation was undertaken. A random series of mass wastage polygons were chosen from the orthophoto-based database. Any polygons containing sections that were not clear in the original data (e.g., indistinct mass wastage boundaries due to low contrast ratios) or that had changed between the orthophoto date and the video imaging date were removed from the test. These polygons were then hidden, and the ODFS was used to digitise their boundaries from the video data (these data are hereafter referred to as 'confirmation polygons'). The resulting linework was then compared with the original (orthophoto-based) data using the same procedures as above. In this case, over 3000 sub-samples were available for comparison. Note that this comparison is between polygons generated from the high-resolution digital orthophotography and polygons generated from the oblique fusion procedure applied to video frames. The comparison procedure is the same as that described in the previous section, with the exception of digital orthophoto generated polygons being substituted for ground survey-based polygons.

The results are detailed in Figure D.21. This comparison results in an epsilon bandwidth of 3.1m. This indicates that, for well-defined features, clear video imagery and properly established registration, 90% of digitised points fall within ±3.1m (at a 1:5,000 scale this represents 0.62 mm—the width of a typical line.) Again, distribution of the offsets is non-Gaussian. A median value of +0.7m

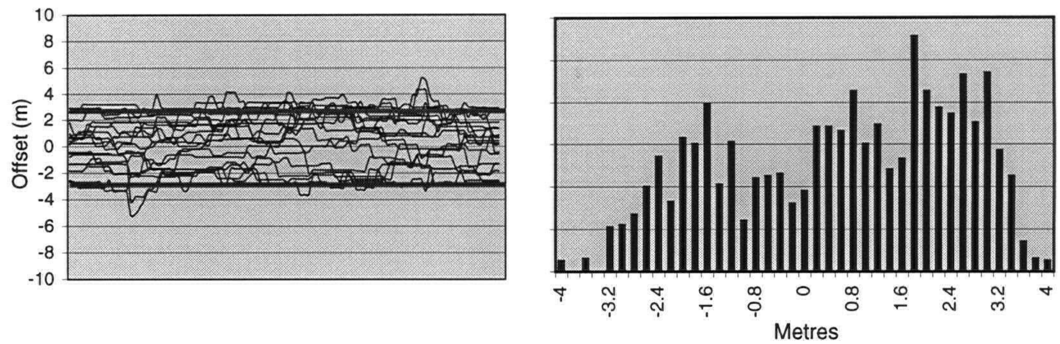


Figure D.21. Epsilon bandwidth and frequency distribution of re-digitised polygons.

indicates that there was a small bias in favour of overestimating polygon area in the ODFS procedure relative to the orthophoto data. This may be due to digitising style or part of the registration process. It may also be caused by the errors in slope value resulting from TRIM data inaccuracies, as noted in Chapter Four.

This accuracy value will obviously vary with distance to target, image quality and other variables. The slides chosen for this confirmation exercise had target distances that varied from 200m to 1600m. This epsilon value therefore represents a rough approximation for typical circumstances. It is unlikely that target distances would vary far outside of these values in most situations.

Area Comparisons

Comparisons of polygon area between ground surveyed polygons, confirmation polygons and original (orthophoto-based) polygons were also compiled. Although these statistics are primarily descriptive, they focus on one of the crucial measurements used in many tasks such as forest inventory, terrain inventory and cutblock delineation. Area comparisons do not necessarily indicate accuracy of registration, since only 'stretching' will be highlighted. Offsets, shears and other transformations will not significantly affect area calculations.

The areas of the four ground surveyed polygons that were eligible for area comparisons were all within 15% of the orthophoto-based polygons (the surveyed polygons were not included in the confirmation set). The use of just four polygons does not allow any conclusive comparisons. In contrast, the confirmation polygons provide considerably more information.

Two tests are utilised. The first directly compares areas of the ODFS-derived confirmation polygons (new) and the orthophoto-based polygons (reference) graphically (Figure D.22). A regression line has been fitted to the data, confirming the above zero median value noted in the section above. The second (Figure D.23) shows the absolute difference as a percentage plotted against the original

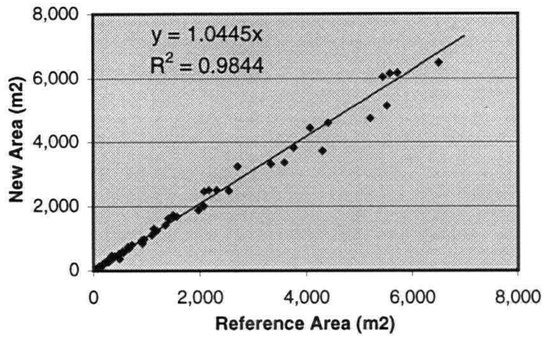


Figure D.22. Confirmation polygon areas compared with original values (derived from orthophotographs).

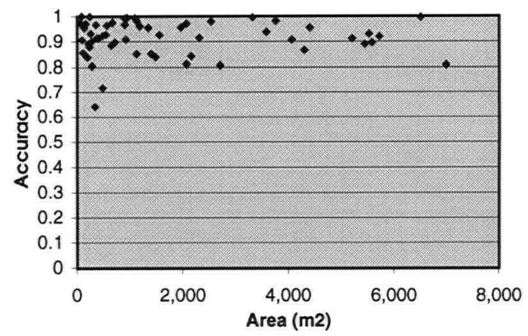


Figure D.23. Percentage accuracy of confirmation polygon areas as a function of area digitised.

area for each test polygon. This indicates that there is no obvious bias in area accuracy based on area magnitude (i.e., the percentage accuracy is not a function of absolute size).

As a summary, 90% of all confirmation polygons fall within 17% of the area as determined from orthophoto polygons. The mean difference is 5%, with a standard deviation of 10.2%. As noted above in reference to digitising offsets, there is no evidence that the areas are normally distributed (Ministry of Forests 1995), so a robust estimate such as the former (17%) provides more information about variance than the latter.

Line Crossings

Area comparison is a rather brute-force method of determining the accuracy of the ODFS process. If, for example, the two polygons being compared are offset to a significant degree, the ODFS database update will be inaccurate, even though the areas of the two may be in complete agreement. The individual lines that make up the polygon cannot be directly compared, because, even if the two polygons are virtually identical, the vertices that construct them may be in different locations and have a different density. Other methods are necessary to estimate ODFS accuracy through polygon comparison.

There are two major factors in comparing polygons in this application: how many vertices do they have (relative to each other), and how far are they offset relative to each other. The first of these factors—the ratio between the number of segments in the original and the re-digitised polygons—is an indicator of complexity. Assuming a similar input process and logic, a high ratio of original to new indicates a decrease in resolution; for example, the landslide or other object may have been digitised from an image taken a considerable distance from the target. A low ratio indicates an increase in detail.

The second factor—polygon offsets—is difficult to measure directly because the two polygons will typically have different structures. One possible surrogate measure is based on the number of line crossings, here termed the 'crossings index'. If the two polygons are accurately registered to each other, then the number of crossings will give a relative indication of offset. For example, the shaded polygon in Figure D.24a has another polygon (dotted lines) overlaid, but very closely registered to the first (both have 70 vertices). There is a large number of crossings between the two. However, in most cases mis-registration will reduce this number. For example, Figure D.24b demonstrates the relative reduction (68 to 19) in crossings due to a linear displacement in registration. In (c) the number is reduced to 12 when the registration has a scale displacement (in the ODFS this is might be due to a foreshortened target distance). If the number of crossings is presented as a ratio to the number of vertices, this 'crossings index' can be used to compare mis-registration of different polygon-pairs, and therefore act as an index of mis-registration

Note that this reduction in crossings due to mis-registration is not a completely consistent process. In certain limited cases the number may actually increase. In extremely convoluted polygons it is possible that a very small rotation or translation would cause such an increase. A series of extremely sharp corners (e.g., a 3-vertex triangle) can also lead to increases. However, most polygons encountered in forest or terrain inventory are relatively simple in structure, and very sharp corners are not common.

Prior to implementing this 'crossings index' for the ODFS-derived confirmation polygons, it is first necessary to determine how the first complexity factor mentioned above—the ratio of vertex counts—impacts on the index. If, for example, a decrease in resolution causes significant mis-registration (a consistently low crossings index), it will be necessary to account for the resolution change in determining ODFS accuracy. Figure D.25 compares the crossings index and the resolution difference between polygons on a scatter chart. The randomness apparent in this chart indicates that there is no relation between the polygons that are offset (or zoomed) and the level of precision in

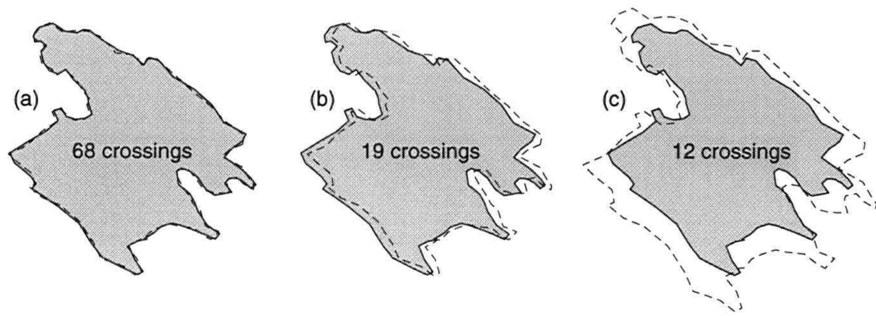


Figure D.24. Misregistration can reduce the number of line crossings between two vector representations of an object. (a) represents two closely registered polygons; (b) shows an offset between the two, while (c) demonstrates a foreshortened target distance.

digitising. In other words, when the image had considerable registration error (low Y), there is no reason to suspect that it was caused specifically by a low-resolution image or long target distance.

Therefore, the number of crossings relative to the number of vertices should provide a rough indication of the mis-registration present. Figure D.26 shows the distribution of this index relative to polygon area. A one (1) would indicate perfect correspondence between the old and new polygon, while numbers approaching zero would indicate large offsets. This must remain a rough indicator, for it is unlikely that there is a linear relationship between the zero and one in this index (this question could make for an interesting digitising accuracy study). Overall, the data illustrated in Figure D.26 demonstrates that 25% of confirmation polygons have very high correspondence with their orthophoto-generated counterparts (using 0.9 as a cut-off), and none of the pairs have a crossings ratio much below 0.4 (where Fig. D.24b has a ratio of ~0.3). Given that this index is a rough indicator of registration accuracy, further tests to determine the actual implications of this value were not warranted (however, this is again an area for further study).

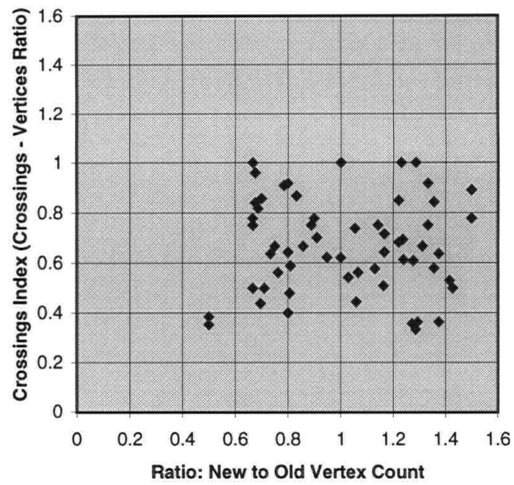


Figure D.25. Precision in digitising (as measured by the ratio between vertices in the old and new representations) is compared with a measure of offset error: the 'crossings index' (the ratio between the number of crossings and the number of vertices).

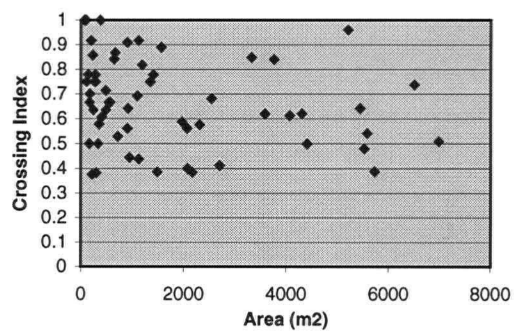


Figure D.26. The 'crossings index' (the ratio of crossings to vertices) for each polygon tested is plotted against the area of the polygon.

Finally, one other factor that may influence these figures is the absolute size of the polygon. If, for example, a relatively large polygon is being digitised, the fixed image size used in this fusion system will cause it to appear at a smaller scale than a small polygon. This may influence digitising precision, as determined by the crossing index. However, as shown in Figure D.26, there is no apparent relation between the crossings index and relative polygon size, so this factor can be discounted.

Results Summary

Overall, this confirmation digitising exercise has indicated that:

1. 90% of all ODFS digitised lines are accurate to within 3.1m of high-resolution orthophoto-generated lines;
2. 90% of all ODFS-based polygon areas are within 17% of orthophoto areas;
3. the mean area offset demonstrated by the ODFS is +5%; and
4. line crossing analysis shows that 25% of confirmation polygons have nearly perfect correspondence (>0.9 index) with the originals; none were offset to any significant degree (<~0.4 index).

Discussion

The magnitude of the area variance (17%) may at first appear relatively high. However, in a relative sense this system does quite well. Trials by the provincial Ministry of Forests have indicated that errors in area derived from photogrammetry (1;15000) are in the range of 6-20%, with considerable regional variability (Everitt *et al.* 1991). Secondly, there is approximately a 7% error in area on the average slide in the study area (which has a slope of 21°) due to the slope-planimetric correction. While the oblique registration procedure accounts for some of this, undulations in the terrain still create an underlying uncertainty that cannot be reasonably corrected.

This magnitude of area uncertainty is therefore typical of existing forestry analysis tools. Most studies of error in forestry mensuration note such errors, but indicate that overall totals are more trustworthy due to cancellation of over- and under-estimates. The mean of +5% found in this confirmation procedure is therefore more indicative of the magnitude of error that might be found in practise.

Limitations

The Oblique Data Fusion System is designed to be a low-cost, medium accuracy update tool, operating in areas of either high data density or considerable topographic relief. Given these constraints, the system displays some obvious limitations in operation. These include:

Reliance on local registration: The system relies totally on the three-dimensional configuration of database objects draped onto an elevation model. It requires a certain density of data to establish reasonable registration accuracy. The registration can rely on either topographic arrangement or key points in other data layers, or both. If both are lacking in a certain area, then registration is impeded, and the resulting linework will be questionable. Although the uncertainty is documented, a severe lack of registration points provides little data for even the uncertainty definition process.

General increase in uncertainty: The system relies upon previously registered images and data for its own registration. Therefore, overall data uncertainty can only increase during an update procedure. A series of updates, each relying on the former for registration, would eventually produce a highly uncertain product through error compounding. This system is therefore best suited for interim updates between larger inventories or other primary data gathering tasks.

Output of spatially variable metadata: While this type of data represents an overall increase in information content, most existing metadata management utilises single measures for entire data layers. For example, a data layer is often defined using a single scale, such as 1:20,000, to define its range of application. The ODFS output may be somewhat variable in effective scale across a single layer. Further processing of the data must explicitly take this limitation into account.

Automated Registration: Ideally, registration point-pairs could be set up for each image, and a batch procedure run to complete all registration (overnight or in the background while digitising). An efficient search algorithm is required in order for registration to complete in reasonable time. Robotic/computer vision algorithms offer a starting point; however, such algorithms (e.g., Keller 1997; Buzug *et al.* 1997) typically utilise one or two fixed positions and a search for shape matching in order to perform 3-D reconstruction. The search required here uses the reverse: a variable position and a fixed 3-D target. Definition of new search algorithms to solve this problem will quite possibly require some degree of understanding of the human vision process; therefore, expert systems and other learning algorithms may have some application. This is clearly an area for further research and development.

Potential Utility

There are several application areas where this system may be of use. Tests described above have focused on terrain inventory: delineating mass wastage zones and updating changes. However, any linear feature can be mapped with this system.

Forest Openings and Updates: This system is particularly suited to mapping forest openings and updating changes on a regular basis. In a typical scenario, where a series of cutblocks are gradually created in one general area over a period of years, this system would allow an extremely short update interval—keeping databases far more current than is presently the case.

Road Networks: Any changes or additions to a road network can be easily captured with this system. New roads into virgin forest would present more of a registration difficulty; success would be dependent on other terrain or planimetric detail.

Mensuration Data: Data gathering typically performed with manual cruising, such as tree heights, stem counts and species ID, is within the capabilities of this system. The oblique viewpoint, registration geometry and 3-D input make interactive on-screen measurement of ground distance and heights a possibility. Given sufficient resolution, species identification is also possible directly from the image.

Buffer Size: Other ground measurement parameters (e.g., stream buffers required by forestry codes) can be directly measured from the image in a manner similar to tree height delineation. The registered elevation model makes ground distance measurements possible. This application has potential utility for code monitoring and other regulatory enforcement.

The potential utility of the type of temporal database that the ODFS helps to make possible is illustrated in the latter sections of Appendix E. There, some initial temporal analysis is performed, and various components of the database are juxtaposed in a preliminary exploratory analysis in order to highlight the type of research these data can enhance.

Conclusions

This appendix has included a discussion of the development of an oblique data fusion system for input of oblique data to GIS-based planimetric databases. The system is made possible by uncertainty management techniques, which allow the variable registration distortion to be captured in the database. By tracking uncertainty, the utility of the registration and the data generated can be easily determined. A number of possible applications were discussed, including the capture of area objects. The costs of the system are far below air photo costs for comparable coverage; however, the data has different utility, and so direct comparisons are difficult.

Appendix E

Development of the Mass Wastage Database

Introduction

This appendix discusses the development of the mass wastage database used in Chapter Five to verify the output of the uncertainty model. An orthophoto-based temporal database is developed, and the Oblique Data Fusion System discussed in Appendix D is used to implement an update. The methods used for both tasks are described in detail. The development of such a database is not a trivial task; therefore, the latter section of this appendix includes an exploratory descriptive analysis of some of the information contained in this database, with the purpose of highlighting some of the possible applications of a temporal database based on the tools developed in Appendix D.

This appendix also includes a detailed description and background of the study area used (and briefly introduced) in Chapter Six: Lyell Island. The background begins with an introduction to the issue of mass wastage in the general study area, the Queen Charlotte Islands (Haida Gwaii).

Background

Mass Wastage in the QCI

Mass wasting constitutes the dominant geomorphic process in the coastal regions of British Columbia (Clague 1989). On the Queen Charlotte Islands, a combination of high rainfall, strong winds and steep slopes result in an unusually high intensity of mass wasting. The term 'mass wasting' is used to encompass a variety of processes by which masses of soil, rock and debris are transported downslope primarily by gravity (Gimbarzevsky 1988). Mass wasting in the Queen Charlottes can take a variety of forms, including rock and debris slides, debris avalanches, debris flows, debris torrents, and slump-earth flows (Clague 1989). There are two principle consequences of these processes. The act of transporting the debris can scour hillsides—at times down to the bedrock—and alter the site's physical characteristics. In commercial terms, site productivity can be considerably reduced. The second consequence occurs in the debris deposition zone, where rock, earth and woody debris can also alter site characteristics. Once a slide has occurred the altered flow characteristics of the site can trigger further slide and slumping events (Gimbarzevsky

1988). Debris accumulation, siltation and changes in flow are of particular importance to fish-bearing streams.

The high intensity of mass wasting found in the Queen Charlottes is not simply a function of structural and climatic characteristics. Two other major factors contribute: seismic activity and human activity.

Earthquakes have played a major role in shaping the physiography of the Queen Charlotte Islands, primarily through displacement and landslides. Currently (i.e., over the past 10,000 years), seismicity levels have been much higher than elsewhere in on-shore British Columbia, with the possible exception of Vancouver Island (Clague 1989). The impacts of these events have been greater than anywhere else in the province. Several of the region's major faults run directly under Lyell Island (Figure E.1). The Queen Charlotte fault, found on the outer west coast of Moresby Island, is particularly active.

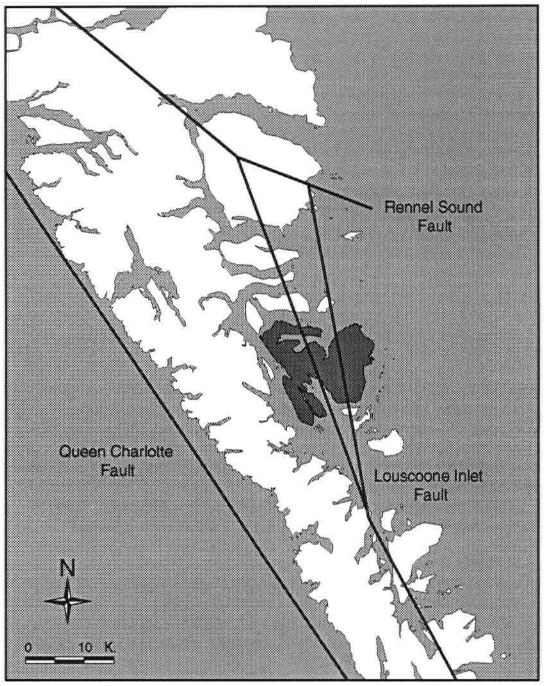


Figure E.1. Faults of the Queen Charlotte Archipelago (Source: Clague 1989)

Human activity—particularly forest harvesting and associated road construction—has considerable impact on the rate of mass wastage in the Queen Charlottes. This fact has been noted by several studies of the region (Rood 1984; Gimbarzevsky 1988); however, the relation has not been adequately quantified in either of these studies. The primary problem with establishing such a relation is the scale of the mass wastage relative to the land base. An analysis of an area as large as the QCI utilising a database that can maintain (for example) 5m wide slides requires a technology that has only become common in recent years. For example, the Gimbarzevsky (1988) study utilised 1-km^2 cells and air photo analysis. Such a study can provide general indications of slide frequency; however, the large cells generalise land use to a considerable degree (e.g., a 1-km^2 cell either contains roads or does not). Drawing conclusions about relations between mass wastage and road building is inappropriate in such an analytical environment.

It is also difficult to draw conclusions about the relation between mass wastage and forest harvesting in general, given the methodologies used in these studies (and others in the general region). For example, a study in the Cascade Mountains (Morrison 1975) found that 7% of all slides occurred in undisturbed areas, with the remaining 93% occurring in logged zones. In contrast, the

Gimbarzevsky study reported the reverse: 10% in cut areas and 90% in undisturbed zones. This discrepancy does not necessarily reflect differences in regional character or logging technology; it is primarily a matter of differing methodologies. The former summary only looked at zones that actually *contained* slides and utilised air photos with a resolution of 1-2 m, while the latter looked at *all* areas, but based the work on 1-km^2 grid cells. Other studies of mass wastage typically demonstrate similar variations. One might focus solely on disturbed areas and use satellite images at 10m resolution, while another might focus on ground data, but use watersheds as the unit of analysis. Wide-ranging studies that utilise appropriate control zones to help establish causality are rare, and are non-existent in the QCI region. The amount of time and data required to perform such a study would be substantial.

Despite these criticisms, the Gimbarzevsky study proves adequate as a general overview. The major conclusions of the study (relevant to this discussion) are that:

1. The average number of failures in active slide areas is 2.6 / km^2. The rate falls to 0.8/ km^2 if the entire QCI land area is utilised.
2. The estimated total area denuded was 12,500 ha, or about 1.25% of the QCI land area.

In contrast, the more specific study by Rood (1984) found that failure rates could be as high as 30 failures per km^2. Again, it is difficult to compare between studies due to differing methodologies and scale of analysis. Nevertheless, it is clear from both scientific and anecdotal evidence that road construction and clearcut logging affect mass wastage rates substantially. Although most authors assume that the QCI are the most mass wastage-prone area in the province, an inability to compare studies from different regions makes proof difficult, despite considerable visual evidence.

This work focuses specifically on Lyell Island. As such, it only generates comparisons between disturbed and undisturbed areas in this relatively small area. However, through the use of extensive metadata, uncertainty tracking and precise georeferencing, databases should be comparable with similar studies in other regions of the Charlottes or elsewhere. This fine resolution will also allow the influence of logging and road building on mass wastage rates to be extracted from the data.

Study Area—Lyell History

Lyell Island is located on the east side of Moresby Island, in the southern Queen Charlottes (Figure 5.1). Lyell consists of ~19,000 ha of originally forested land. The central valleys and eastern sides contain(ed) the bulk of the merchantable timber, while much of the western side consists of smaller trees on thinner soil. While the west side of the island is relatively protected, the east faces directly onto Hecate Strait. This strait plays host to numerous winter storms, and is considered by mariners to be one of the more dangerous passages in North America. These storms typically strike Lyell from the south-east, and provide the rain and winds that trigger most mass wastage events.

The island has been the site of extensive clearcut logging since about 1920. Originally, logging was conducted principally along the coasts. After a significant pause, the establishment of TFL 24 (which covers the entire island and much of the surrounding region) in 1958 led to renewed cutting activity. The most recent activities began in 1976. In the 10 years between then and 1986, approximately 20% of the island's land base was clearcut. Over 140-km of roads were constructed to reach over 3,000 ha of timber, with a scaled volume extracted of just under 2 million m^3. This timber would represent a gross market revenue of between $97 and $125 million (Ecosat Geobotanical Surveys 1989). Although cutting in the earlier part of the century had utilised temporary or floating camps, this more recent activity used a permanent camp at the south end of Powrivco Bay (central north shore).

In 1985 the Haida Nation designated Gwaii Haanas as a Heritage Site under the Haida Constitution, and in July 1987 Canada and British Columbia signed the *South Moresby Memorandum of Understanding*, which later led to the creation of Gwaii Haanas National Park Reserve / Haida Heritage Site. Gwaii Haanas includes all of Lyell Island, within a land area representing approximately 15% of the entire Queen Charlotte Archipelago. Lyell Island was the principle focal point of the Haida campaign that brought about the creation of this protected area. Following the establishment of Gwaii Haanas, it also became the focus of an extensive rehabilitation program. The agency responsible for the area is the Archipelago Management Board, or AMB.

The most recent logging activities on Lyell have led to extensive slope destabilisation, resulting in numerous large landslides—many of which are visible from considerable distances. The camp and dump detritus in Powrivco Bay were both an eyesore and a source of substantial soil contamination. The forestry company's figures showed that over 750 ha still required planting. The Ministry of Forests absolved the company of all responsibility for cleanup and restocking; the responsibility was passed to the AMB. They initiated a substantial rehabilitation program in 1989 and, over the next three years, spent over two million dollars on various components of the program. The database discussed in this appendix will also be used, over the long-term, to evaluate the success of this program.

Methodology

The overall purpose of the work discussed in this Appendix is to compile a mass wastage database at the best possible resolution (both temporal and spatial) for use in evaluating the output of the slope stability uncertainty model. A secondary purpose is to provide a case study of the effectiveness of the Oblique Data Fusion Tool discussed in Appendix D (note that the development tests used a limited subset of the case study data—a compromise necessitated by acquisition costs).

The general steps in the methodology are:

1. Evaluate the utility of available existing data sources for production of a high-resolution georeferenced temporal database;
2. Develop a 'baseline' database using existing sources, performing all corrections (e.g., referencing and orthocorrection) required;
3. Gather aerial data for the ODFS tool; and
4. Update the baseline database with the ODFS tool.

Initial tests occurred in the summer of 1996. Existing aerial photographs were used to generate a landslide and cutblock baseline database for four different periods (1974, 77, 80 and 90) with a spatial resolution of under 1 m. In 1997, a single helicopter flight procured the video data, and several weeks of ground-truthing were performed for the purpose of system development. Data were extracted from the video frames using the ODFS.

Data Sources

The most important issue in developing the baseline database is that data must be carefully georeferenced. Studies that provide area summaries, or those that use non-standard referencing techniques, cannot be revisited at a later date to determine either study accuracy or change over time (other than through general area summaries). Only when all relevant items can be spatially located and later relocated can a proper temporal database be created, allowing analysis such as change over time in relation to other spatial variables.

Although considerable data are available through the AMB (and other sources), very little historical data were explicitly georeferenced. Therefore, the first stage of this project involved determining the usability of the many data sources. These sources include:

- air photos of varying resolution and coverage, dated 1964, 1974, 1977, 1980, 1986, 1989, and 1990;
- BC TRIM data (1:20,000, based on 1986 photos) including roads, rivers, cutblocks, coastline and photogrametrically-produced elevation points (~30m intervals);
- inventory maps (forest cover) hand sketched onto mylar, dated 1975 and 1980;
- AMB base map and thematic databases, based primarily on 1986 TRIM photos (1:65,000) and compiled by Westland Resources in 1994;

- summaries of a 1988 LANDSAT TM image (Ecosat Geobotanical Surveys 1989);
- a 1991 LANDSAT TM image (30m, 7 band);
- a 1992 SPOT image (10m); and
- several timber harvesting history maps from Western Forest Products (1979 and 1985 plans).

Aerial Photography

The various aerial photography series provide the most detailed, unbiased data on cutting and landslide history. However, the resolution and coverage vary between series. Only the following series were used:

1974 series: only exist as a compiled air photo mosaic at approximately 1:60,000. The mosaic was not intended for analysis, and therefore has not been orthocorrected or even lined up to match other data. However, it represents the only complete coverage of Lyell prior to the 1976 commencement of logging. It was therefore scanned and run through georeferencing and orthocorrection software. Area estimates could not be derived from this map, but the existence (or non-existence) of all major slides could be ascertained. This series therefore assisted in determining the approximate natural slide rates.

1977 series: provided partial coverage at ~1:18,000 in B&W. This series was digitised and orthocorrected (see below for georeferencing and elevation sources).

1980 series: provided partial coverage at ~1:18,000 in B&W. This series was also digitised and orthocorrected.

1990 series: partial coverage in colour at 1:10,000, including all cut areas. This series was digitised in colour and orthocorrected. Figure E.2 details the extent of the three partial coverages.

All photos were digitised with pixel sizes of approximately 1m (1.2m for B&W, 0.9m for colour). Georeferencing and orthocorrection were accomplished using 'Orthoengine' software by PCI. The

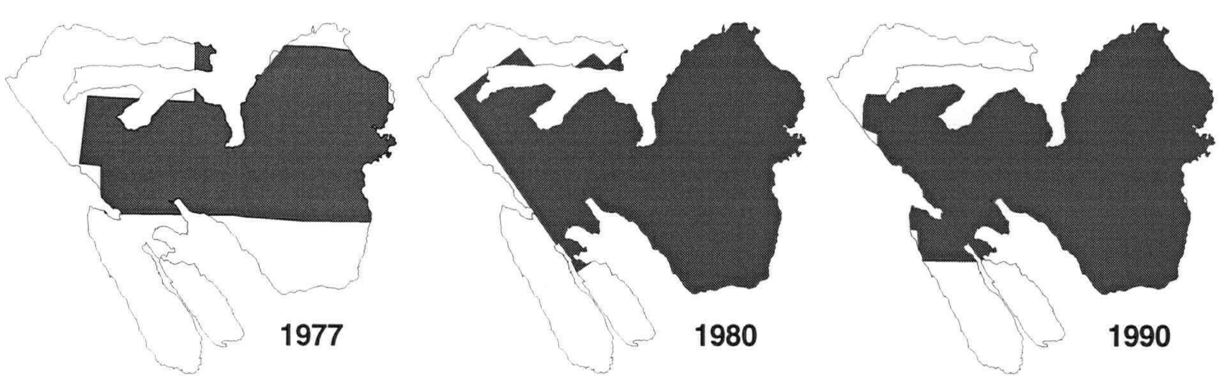


Figure E.2. Areal coverage of digital orthophotos.

1990 series consumed 2/3 of the working and processing time due to the detail involved at the higher photo resolution. An average overall RMS error of 4m was achieved. However, most of this error is confined to areas with standing timber due to a lack of georeferencing points. The primary landslide zones have an overall estimated RMS error of ~ 2.3 m.

Elevation and Planimetric Data

BC Terrain Resource Inventory Mapping (TRIM) data were utilised as ground reference for all orthocorrection work. TRIM data are extracted from 1:65,000 photos, but are given a resolution rating of 1:20,000. The quoted accuracy levels of TRIM planimetry are 90% of all well-defined features co-ordinated to within 10m of their true position. This corresponds to a circular standard error of 2.15m (SRMB 1990). These numbers represent the *minimum* accuracy standard for the data (quality can vary considerably between areas of the province due to the large number of contractors used to create these data). This level of accuracy might be a problem in dealing with individual *points* during photo correction and subsequent analysis; however, lines and area features also have a relative placement that helps to mitigate any inaccuracies. Even if the entire map sheet (TRIM data file) is shifted 5m from its true position, the features within will exhibit an accuracy relative to each other higher than 5m. Other accuracy concerns relating to the analysis of landslides will be discussed below.

The digital elevation model (DEM) was developed from TRIM data as well. Points generated from digital photogrammetry are provided at approximately 30m intervals. A drainage-corrected digital elevation model was created for use in orthocorrection, landscape visualisation and analysis. The TRIM specification for these elevation points is 90% accurate to within 5m of true elevation.

AMB Databases

Relevant databases made available by the AMB include planimetry (TRIM data - 1986 vintage) and thematic databases for terrain and biophysical features. The thematic databases are based on the 1:65,000 original TRIM photos (working resolution of 1:20,000), coupled with field observations within derived polygons. Data were collected in accordance with the *Terrain Classification System for British Columbia* (BC Ministry of Environment 1988). Surficial material and geomorphic processes were derived directly from observations. Several types of terrain hazards were assessed based on these observed data, as well as derived values (such as slope and aspect).

The biophysical database was derived in much the same way; however, the classification system used was developed specifically for Gwaii Haanas. Both it and the specifics of thematic database collection methodologies are described in detail in the Gwaii Haanas Ecological Land Classification (Westland Resources Group 1994). This project utilises portions of both databases in order to analyse slide characteristics.

Other Sources

Other available data sources include satellite coverages (LANDSAT, SPOT) and a variety of engineering and thematic maps generated by both forest companies and rehabilitation contractors. Satellite data sources were considered too coarse-grained to be utilised in this study other than for general reference. The original 1:5,000 mylar sketch maps used by the forest company for their initial inventory in 1975 were examined and tested. However, they did not utilise any standard georeferencing method. With a great deal of effort these maps might be skewed to fit existing data, and the effort would be justified if details of the original forest cover were required for a study. This effort was not deemed worthwhile for this project.

Maps of cutting activity (covering several periods) generated by the forest company and in subsequent studies also lack georeferencing. Attempts to digitise and skew them in to fit existing data proved futile; it appears that these maps were based on road engineering sketches of the island and were never intended to be combined with other data. They are useful for general reference, but not for analysis. A series of strip charts generated by rehabilitation contractors were based on the same base-maps and suffer from the same deficiencies. Although numerous efforts were made, these maps proved impossible to accurately georeference.

Baseline Database

The three airphoto series whose coverages are illustrated in Figure E.2 (1977, 1980 and 1990) were used in combination with TRIM data to generate three orthophoto mosaics. These were then used in combination with other data discussed above to generate GIS vector layers detailing visible mass wastage, visible cutblocks and road updates. This section provides details of their lineage, accuracy, contents and structure.

Description

Orthophoto Mosaics: The coverage of the three mosaics was described above (Fig. E.2). Georeferencing was accomplished with TRIM data and a TRIM-based elevation model. Problems encountered in the georeferencing process were primarily due to inaccuracies in the TRIM data. Coastlines and road intersections were generally trouble-free. Road endpoints were commonly found to be troublesome (i.e., inaccurately georeferenced and of limited utility), as were some of the rivers. Taking these issues into account, the final product represents a compromise between skewing accuracy and visual accuracy. Unfortunately, even though TRIM planimetry is known to be error-prone, it is the current de facto standard baseline; there is no other 'truth' to refer to at present.

The 1990 colour series is the most detailed and accurate of the three due to the larger scale of the photos. Over 300 ground reference points were used to reference the images, and over 500 tie points were used to knit them together. For all images the overall RMS error is under 4 metres, but much of this error is due to a lack of reference points in unlogged areas. Points in the cut areas have average errors less than this value. This number was estimated at ~2.3m through separate trials as discussed in Appendix D.

Mass Wastage: All visible mass wastage zones were digitised from each mosaic into (initially) three separate databases. In order to ensure the accuracy of these data, each air photo used in the mosaic was exhaustively examined by two independent analysts using a stereoscope. All mass wastage polygons required agreement by both. These polygons were dated by mosaic date, and then overlaid into a single coverage.

Each polygon was then examined with the following data in-hand: rehabilitation slide-seeding records and maps, forest company reports detailing slide occurrence dates (with rough sketch maps), pre-1976 air photos, and records of extreme weather events. From these, each slide's date of occurrence and, if applicable, date of disappearance (i.e., regrowth to an 'non-disturbed' status) were estimated and entered. It should be noted that these values were entered conservatively. For example, many of the slides dated with a 1990 occurrence no doubt happened during the numerous storm events in the mid-1980's. However, if no concurring evidence was available, dates were left as 'first noticed'. This has certain implications for analysis, and will be discussed in later sections.

Only obvious artefact slivers were removed from these polygons in order to facilitate analysis. Some of the remaining slivers are due to overlay (i.e., they represent the same line), while others represent growth (or shrinkage) of slides. These were dealt with on an individual basis during analysis.

Subsequently, each slide was examined in light of other data, and fields were coded based on the following criteria: source (natural or logging); impact on streams (scouring, side impact, or direct scouring of potential fish habitat); type of event (mass movement or road sidecast); and relation to roads (triggered by road or not). These items were *not* coded based on automated analytical criteria, but were coded individually taking all available data into consideration. Unfortunately, as there existed no standard reference or numbering system prior to 1990, much of these incidental data were not useful in database construction. For example, knowing that 20 slides occurred in Nov., 1988—but not *which* slides—does not assist greatly in database coding.

Aerial Video Update

Aerial Data Gathering

The aerial video data gathering for the 1997 database update took place in late July of that year. The weather was overcast, with numerous low clouds and occasional rain. A Bell 206 helicopter, rear left door removed, spent approximately 2 hours over Lyell gathering the data. All slide areas and cutblocks were flown at least twice, and all other potential slide areas were surveyed. Given that the system was under development, a number of different speeds, elevations, and distances to target were utilised. In a practical exercise all relevant data could have been gathered in 45-60 minutes.

The camera utilised was a Sony Handicam—a high-end one-CCD consumer model with optical image stabilisation. Earlier tests had indicated that the higher resolution offered by professional (three-CCD) cameras was not required, and the lack of standard image stabilisation technology in these cameras actually reduced the quality of most frames relative to the one-CCD model. The GPS data were captured to a laptop computer during the flight; however subsequent flights have utilised internal GPS memory only. Differential correction was *not* utilised in this test. It has been used in subsequent work, and found to speed up the georeferencing process to a small degree. The use of differential GPS did not increase overall accuracy of the captured linework.

Image Registration and Data Capture

Over 420 video images were captured from the tape and registered using the ODFS system described in Appendix D. Images were chosen based on minimal blur, target (landslide) centring, and optimal viewpoint/target geometry. Landslides were digitised using the ODFS system, and any with questionable geometry (e.g., not centred in frame or captured from too low/high an angle) were checked using a frame with a different viewpoint.

Several additional factors were incorporated or dealt with during this procedure that were not part of the initial testing.

1. It was found that the image was often rolled forward or backward relative to the reference ground plane (i.e., helicopter pitch or camera roll). Although the database could be rotated, a much simpler approach was used: interactive rotation of the image became one of the final steps in the registration procedure.
2. Tree height values were added to ground elevation values through the use of an 'age' parameter in the cutblock database. In retrospect, a more sophisticated ground cover database might have increased the accuracy of the registration process. This change would have to account for variations in tree height around watercourses, variations due to soil quality, etc. Forest inventory maps coupled with cutblock information could be utilised if available.
3. In a GIS-based perspective view the more accurate the rendered surface is, the slower the image is to display. The most effective method employed in the tie-in process was gradually working up from a coarse to a fine display as parameters tightened.

4. The outer edges of both images and perspective views—particularly foreground areas—are prone to greater error due to both lens distortion and perspective algorithm inaccuracies. Images were chosen that did not have crucial data in these areas.

5. It was found that errors in TRIM planimetric data and errors in orthophoto mosaic registration could create discrepancies. These problems are dealt with in more detail in the following subsection.

Overall, comparative work has shown that the high relief environment is of particular advantage in tying-in the image data. The presence of numerous ridgelines and incised streams and valleys greatly increased the speed and accuracy of the process. If a planimetric vector (e.g., a road) were visible in the foreground, the height of land visible above the target, and at least one peak or ridge visible in the background, then precise registration could take place very quickly. This level of precision refers to locating the aircraft to within 3-5 metres of its true x,y,z co-ordinates. The levels of precision of the baseline data and of the display process itself preclude tuning the parameters to a finer degree.

Ground Truthing

Ground truthing for this update was performed in concert with system development. All relevant details are described in Appendix D.

Results

The principle result of the above procedures is a detailed database of landslides in the study area, covering over 22 years of history (likely closer to 30 years based on slide visibility persistence), as well as an elevation and slope model. These data are used in Chapter Six in validation tests of the slope stability uncertainty model. However, a detailed temporal database such as this has a number of other uses. The following sections present a brief exploratory analysis of these data, with the purpose of highlighting the potential utility of the ODFS tool, and the level of detail available from these data gathering method. This information is not intended to represent a thorough analysis; instead, it is meant to offer possible directions for further study. It is also intended to point out why regular incremental update of information (rather than very occasional full re-inventory) is of considerable use in temporal analysis.

Current Conditions

An overall summary of the change over time in various categories is presented in Table E.1. For each of the result columns that are based on orthophotos (1997, 1980 and 1990), the 'study area' refers to the orthophoto coverage. In the final column it refers to the area covered by the aerial video update, which encompassed the entire island. The orthophoto coverages include all areas that had been logged to that date; therefore, no logging-induced mass wastage has been excluded from any of the study areas.

Overall, there have been 144 ha of mass wastage observed on Lyell since 1977 (0.75 percent of the land base), including areas visible at the onset of logging. Currently, 115 of those hectares remain visible. The 1997 update indicates that 21 ha of new slides have occurred since 1990, although approximately six of these occurred between 1990 and the onset of treatment in 1991 (based on differences between the 1990 data and a set of maps generated at the outset of the rehabilitation project) . Of this 144 ha total, 11% are classified as naturally occurring; and 21% of these natural slides were in place prior to the 1976 commencement of logging.

Note that the final column of Table E.1 was generated with a different methodology than the first three columns. Accuracy statistics for the ODFS tool (used to generate the final column) compiled in Appendix D indicate that 90% of all areas should be within 17% of the area established with high-resolution orthophotos (for a given range of operational scale). However, in summary statistics most errors in area will cancel out (a +0.7m area bias was observed during testing in Appendix D).

			1977	1980	1990	1997
Study Area: (ha)			10390.4	13990.2	14550.9	19036.8
Total Area Logged (visible)			214.7	991.8	3517.2	3517.2
as % of Lyell			1.1%	5.2%	18.5%	18.5%
as % of study area			2.1%	7.1%	24.2%	18.5%
Area logged during prev period			214.7	777.1	2525.4	0.0
as % of Lyell			1.1%	4.1%	13.3%	0.0%
Slide area show ing			9.1	27.0	97.7	115.3
as % of Lyell			0.05%	0.14%	0.51%	0.61%
as % of study area			**0.09%**	**0.19%**	**0.67%**	**0.61%**
as % of total cut area			4.25%	2.72%	2.78%	3.28%
Recovery since prev date				5.3	19.2	3.9
New Slides			9.1	23.2	89.8	21.6
as % of study area			**0.09%**	**0.17%**	**0.62%**	**0.11%**
Treated Slides						75.0
as % of 1990 slides						**76.8%**
New Slides						
	Cause	Logging	5.8	17.7	83.4	21.1
		Natural	3.3	5.4	6.8	0.5
		(as % of unlogged)	**0.03%**	**0.04%**	**0.06%**	**0.00%**
	Type	Mass W	8.8	22.0	85.9	20.3
			96.8%	91.7%	95.6%	92.6%
		Sideslip	0.3	2.0	7.2	4.9
	Stream	Scouring	7.3	19.7	54.2	4.3
			80.3%	85.0%	60.4%	19.7%
		Main (fish habitiat)	0.0	0.2	0.0	0.0
	Road Induced		1.3	6.5	18.2	7.3
			14.6%	28.1%	20.2%	33.8%
Total Slides						
	Cause	Logging	5.8	18.6	89.1	106.6
		Natural	3.3	8.4	8.6	8.7
	Type	Mass W	8.8	25.9	93.3	109.6
		Sideslip	0.3	2.0	0.8	12.5
	Stream	Scouring	7.3	22.7	60.3	62.1
		Main	0	0.2	0.0	0.7
	Road Induced		1.3	6.5	20.3	26.7
Treated Slides						
	Cause	Logging				75.0
		Natural				0.0
	Type	Mass W				74.5
		Sideslip				3.5
	Stream	Scouring				49.4
		Main				0.0
	Road Induced					16.0

Table E.1. Summary of the rehabilitation study results.

Change Over Time

The natural slide rate (i.e., restricting the analysis to only those slides occurring in unlogged areas) as interpolated from four data points is presented in Figure E.3. The 'uncorrected' line represents the data as observed, while the other represents a per-year estimate. It is obviously difficult to establish a natural rate on a land base that is constantly having natural areas removed. This drop in rates may be due to a) saturation - where most potential slide zones have experienced failures in the past and are in the process of regeneration; b) the removal of most potential slide zones from 'natural' to 'logged' areas; or c) a drop in natural slide rates due to weather, tectonic or other factors. Nevertheless, this 20-25 year window of time is likely insufficient to determine slide rates, even given the best of data.

This possible weather influence on slide rates is explored in Figure E.4. Major precipitation events appear to taper through the 1980's, in-line with the decreasing rate of natural slides. This graph provides evidence for reason (c); however, the limited amount of data makes this evidence merely circumstantial.

Figure E.5 details values for all slides, regardless of origin. Chart (a) shows *new* slides using the same method as Figure E.3. The 'trend' curve represents the best approximation of slide rate available from the data (i.e., new area exposed per year). The cumulative values (minus grow-back) are shown in Figure E.5b. This curve represents the estimated visible area of mass wastage at any given time.

It is apparent that the new slide rate has certainly changed since the cessation of logging operations in 1986. Research indicates that the slide rate will typically peak at 10-15 years post-harvest (assuming sufficient regeneration). After this point, new root growth will typically overtake the old root decay, stabilising the slope (Hammond *et al.* 1992). This is supported by the database, as shown in Figure E.6. In this graph the two initial peaks are likely erroneous, since they represent

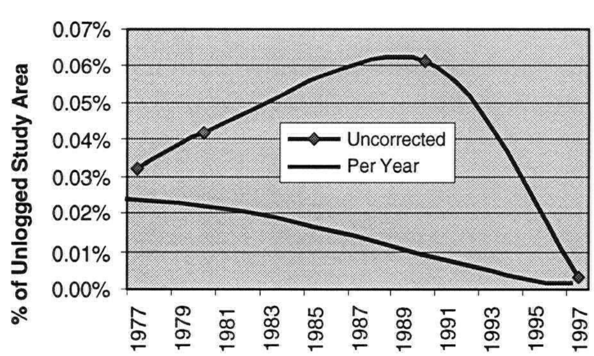


Figure E.3. The natural slide rate. The uncorrected points delineate totals for the intervening period. The second line shows the approximate rate per year.

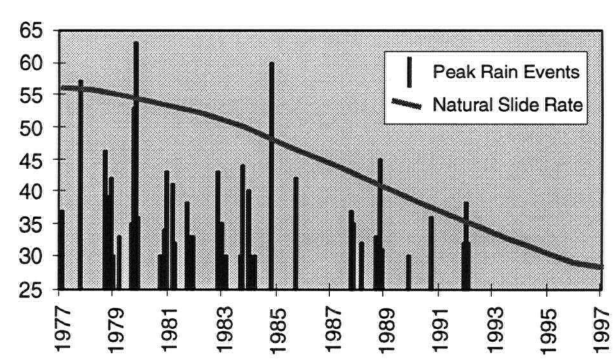


Figure E.4. Estimated natural slide rate plotted with peak rain events.

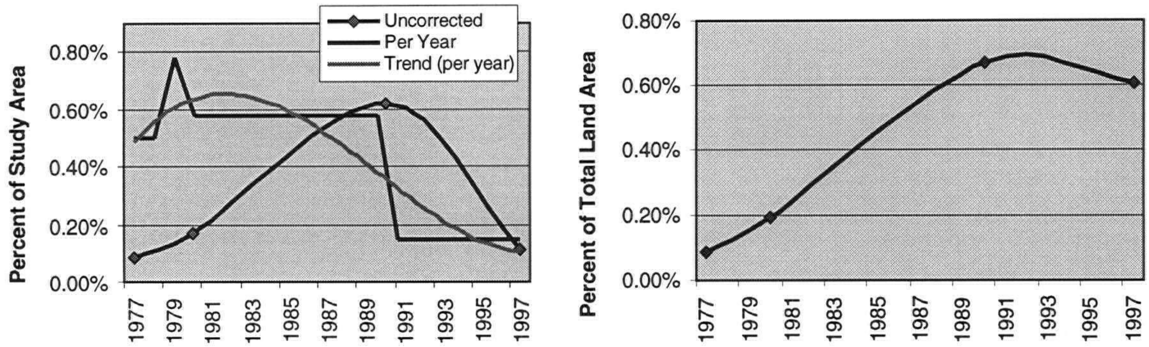


Figure E.5. (a) New slide rate (all sources), and (b) cumulative slide area minus grow-back (area visible in the year specified).

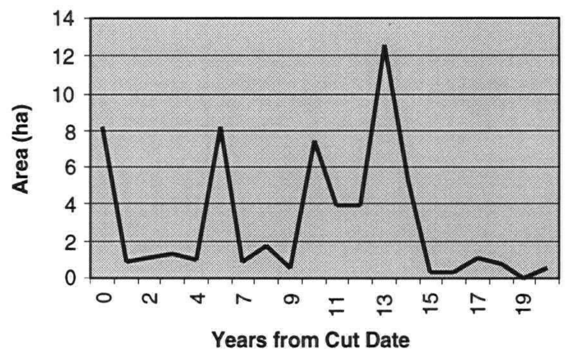


Figure E.6. Slide initiation delay.

all the 1981-90 slides that were lumped into a 1990 start date due to insufficient data. Otherwise, slides appear to typically commence right on schedule.

Limitations: It is difficult to make any broad statements regarding mass wastage given the limited data available. Although the mass wastage database represents all available useable knowledge, it falls short in many areas. First, there are only four major data points. It is difficult to calculate rates and averages with these few data. Although slide initiation dates have been supplemented through records, it is obvious that many of the slides labelled (for example) START=1990 actually occurred during major storms in the mid-1980s. However, in an absence of other information, the conservative value has been left in place. Secondly, mass wastage is generally driven by extreme rainfall events (cumulative over a short period of time). There is, therefore, no true value for 'slide rate'. There can only be an average slide rate over a period of time. The values presented here must be interpreted with these limitations in mind.

Slide Characteristics

The development of this database presents an excellent opportunity to study mass wastage events themselves. There are many possible applications, including comparisons with other areas, relation to terrain form, relation to soil or weather, and assessments of predictive models. This section delves into several of these topics—principally to demonstrate the possible utility of the type of spatio-temporal mass wastage database that can be generated using the ODFS and uncertainty management techniques.

Slope and Aspect

The existence of a DEM makes it possible to easily catalogue mass wastage areas based on terrain characteristics. The areas are broken down into grid cells and compared with the cells of the DEM. Figure E.7 compares the slope (average) of all slides with the overall distribution of slopes on the entire island. Between 30° and 44° the incidence is higher than expected (the expected value would occur if slides were placed randomly). Below 30° the incidence is lower, and above 44° it is approximately the same. This information on expected values is detailed in Figure E.8.

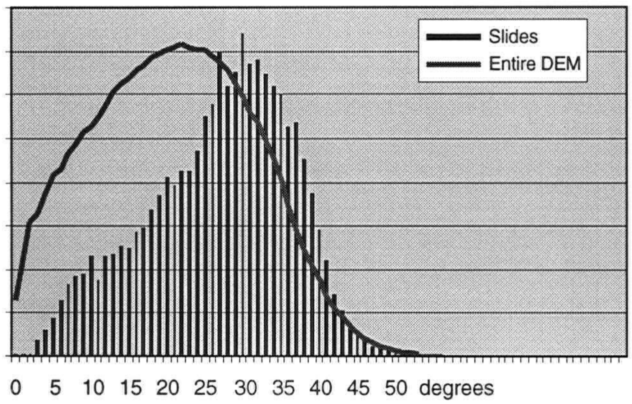


Figure E.7. A comparison of the frequency distribution of slope in slide areas with the distribution of slope over the entire study area.

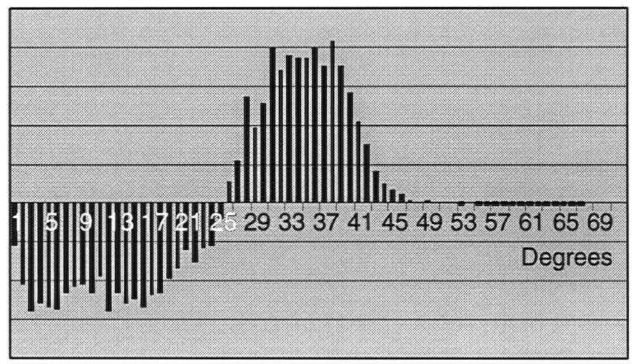


Figure E.8. Slide slope frequency distribution compared with the distribution that would be expected if slides were located randomly (observed minus expected).

The distribution of mass wastage aspect relative to expected is graphed in Figure E.9. This figure has been corrected based on the overall aspect of the DEM (i.e., a circle would indicate a random distribution of slides on the island). The influence of southeasterly storms is evident. The decrease in expected value in a northeast direction may be due to the fact that the NE corner of the island is the only substantially unlogged portion facing the winds of Hecate Strait.

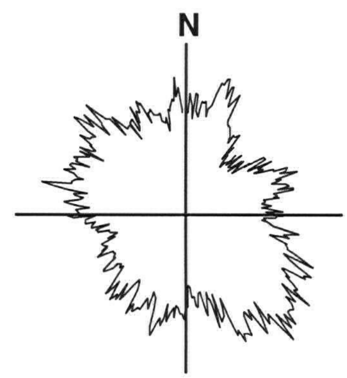


Figure E.9. The frequency distribution of slide aspect (corrected using the total aspect distribution for the study area) showing a SE directional bias for slides, and a negative bias to the NE.

Weather

The apparent correlation between natural slide rates and major weather events has been discussed earlier. Here, in Figure E.10, major wind events are added to the chart. It is possible that the overall decline in general weather severity in the past 20 years has contributed to the decline in natural slide rates. Only with detailed regular data updates could such a correlation be properly estab-

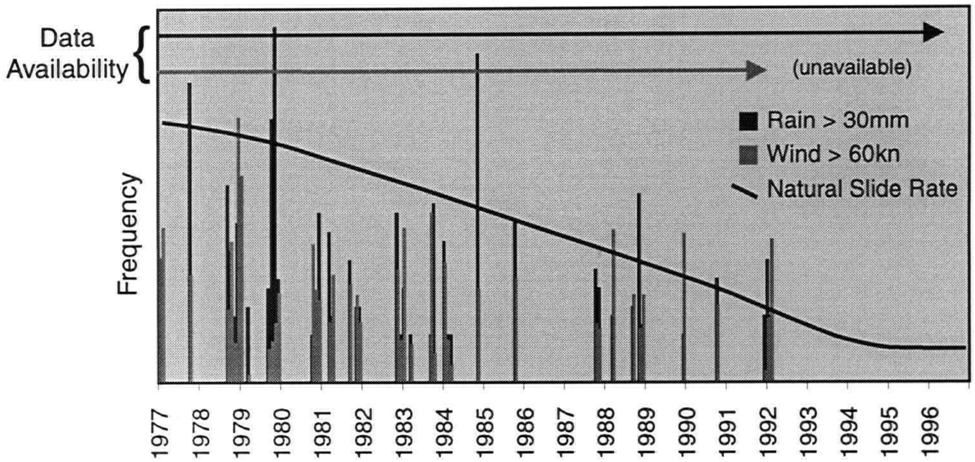


Figure E.10. General peak weather patterns compared with the natural slide rate. Figures are based on daily peaks (wind) and daily totals (rain) over the 20 year period (Sources: rain data - Environment Canada published statistics for Juan Perez Sound; wind data - Environment Canada database of daily wind peaks for Juan Perez Sound as compiled by Dr. S. Tuller).

lished. There is quite probably a triggering threshold that involves a certain amount of precipitation over a certain period of time, applied to a slope of > *x*. Properly populating this temporal database will make it possible to determine these values and to better predict slide occurrence.

Discussion

This test of the ODFS in a real-world scenario has shown that, for extreme and remote environments such as this, oblique video monitoring represents a highly useful data source. It is possible to capture seasonal and event influences by gathering data whenever new slide events are suspected to have occurred. Oblique monitoring does not require perfect weather conditions, and so could take place at nearly any time of year. Also, video data can be gathered opportunistically. For example, when a helicopter is ferrying equipment or personnel through the area, one extra hour's worth of fuel would accomplish the task. Database integration could take place at any time.

The integration system functioned well for a wide selection of images. Very few of the 420 images were difficult to register. Spot checks, in which a newly digitised area was viewed from a different angle, led to very few changes in linework. Line uncertainty varied, as the system was designed to do, with very few vertices registering an epsilon of more than 4m. Due to this inherent variability within the system, it is not possible to state overall uncertainty statistics.

Conclusions

The work discussed in this appendix has brought together information from a number of sources in order to build a detailed spatio-temporal database of mass wastage on Lyell Island. A secondary focus has been on evaluating the ODFS, and exploring possible secondary research areas that such a database makes possible. Specifically, the mass wastage history and current status of Lyell were determined through the use of orthophotographic interpretation and through the development of the ODFS—which integrates oblique photographs with an existing database. The utility of the database generated in this project was explored through comparisons with other data—including spatial continuous data (slope, aspect), spatial discrete data (soil, terrain) and non-spatial data (precipitation and winds). The work concludes with a discussion of changes made to the ODFS due to this study.