LANDSCAPE CLASSIFICATION USING GIS AND NATIONAL DIGITAL DATABASES

A thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy
in the University of Canterbury
by
Lars Kevin Brabyn

Department of Geography
University of Canterbury
New Zealand
May 1996

ABSTRACT

This study considers whether visual landscape character can be classified using GIS. Landscape classification is needed to give landscape researchers and planners a frame of reference for communicating and comparing their research. Such classification is difficult because of the complex nature of landscapes and because it must be explicit. Classification needs to be based on theory, but there is a distinct lack of landscape theory. It is argued that to effectively develop landscape theory a classification is required and that a classification evolves with theory. GIS provides a suitable platform to facilitate this evolution.

A set of criteria is established to which a landscape classification should adhere. To be useful for evaluative and cognitive research, a landscape classification needs to distinguish the important characteristics that affect landscape. These characteristics are identified from what little landscape theory exists: a landscape classification needs to incorporate landform, vegetation, naturalness, and water; the classes should be based on the public's perception; the classes should be general and involve compositions; and the classes should incorporate movement and exploration. Besides these criteria, more general criteria that have been used on other land based classifications also apply, particularly the need for a classification to be repeatable.

GIS and national digital databases can incorporate these criteria in a landscape classification and this is demonstrated on a transect of the South Island of New Zealand, using mainly a 1:250,000 topographical database and a vegetation database. Difficulties associated with these databases are discussed. A three-phase landscape classification process is developed:

1) Selection of attributes,

2) Definition and classification of the attributes to six levels of generalisation, and

3) Creation of landscape classes from compositions of the attributes.

The sensitivity of the process to different operational definitions is considered, and it was significant in some cases. An important analysis function that enables GIS to classify landscapes is the focal neighbourhood function. This in effect analyses the study area from many different points. Once a landscape classification is developed, it can be used with GIS for description, mapping, and inventory purposes. Uniqueness and variety of landscapes can also be determined. A range of observer perspectives can be recognized in the classification by using an application of fuzzy set theory that incorporates entropy.

Automating landscape classification requires developing appropriate operational definitions that balance the human concept model of landscapes, the characteristics of national digital databases, and GIS capabilities. Operational definitions can be formulated using four abstractions: classification, generalisation, association, and aggregation, and then represented using GIS analysis techniques. Classifying landscapes automatically is an exercise in generalisation, as there is a considerable amount of information to consider. The challenge is to produce a meaningful generalised classification, rather than a very detailed classification. Expressing association is also important because landscapes are a composition of different landscape components. Focal neighbourhood functions enable the spatial influence of different components to be expressed and from this landscape compositions can be identified.

The national digital databases used in this study do not contain conceptualised information on morphological landforms. Height contour databases are available from which it is possible to classify landforms and a substantial part of this study investigates this. Hammond's manual landform classification was automated and applied to the study area. Some problems were identified and a modified process was subsequently developed.

ACKNOWLEDGEMENTS

With sincere gratitude I would like to thank a number of individuals and organisations. Firstly, a special thanks to my supervisors Prof. Pip Forer, followed by Dr Burn Hockey. Pip's early enthusiasm towards this study was particularly appreciated, as well as his ideas and feedback. Burn took on the particularly difficult challenge of supervising me after Pip's departure to Auckland. He met this challenge very well and the quality of the final product can be attributed much to him. Thanks Burn. I would also like to thank Assoc. Prof. Doug Pearce for his time in helping me to formulate a topic, and Dr Eric Pawson for reviewing the first couple of chapters.

A special thanks to the technical staff in the Geography Department, and in particular John Thyne, James Guard, and Graeme Glen for computer support, and Janet Bray, Fiona Clark and Michelle Rogan for their assistance.

Many individuals from outside the Department also showed interest and gave valuable advice. This included Dr Simon Swaffield from the Department of Landscape Architecture (Lincoln University), Steve Thompson from the Ministry of Forestry, and Allan Rackham from Boffa Miskell.

Part of my research was conducted from the Department of Physical Geography, University of Oslo, Norway, and I am very grateful to this Department for accommodating me. It was during a visit to this Department that I met Dr Richard Dikau, and his work subsequently influenced the direction of my Ph.D.

Landcare Research Ltd. provided many of the databases used in this research in exchange for nominal costs and a copy of this thesis. I am very grateful to them for this assistance. The Ministry of Forestry, and Statistics New Zealand also provided useful databases. The support and generosity from these organisations is very much appreciated.

Lastly, a special thanks to my friends and family for their support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

AML	ARC Macro Language
ARA	Auckland Regional Authority
AVHRR	Advanced Very High Resolution Radiometer
DEM	Digital Elevation Model
DOSLI	Department of Survey and Land Information
DCW	Digital Chart of the World
ESRI	Environmental Systems Research Institute
GIS	Geographical Information Systems
GPS	Global Positioning Systems
GRID	Global Resource Information Database
LINZ	Land Information New Zealand
LRI	Land Resource Inventory
NAW	Neighbourhood Analysis Window
NDDB	National digital databases
NOAA	National Oceanographic Atmospheric Administration
NZ	New Zealand
PNA	Protected Natural Areas
RMA	Resource Management Act
SER	State of the Environment Reporting
TIN	Triangulated Irregular Networks

CHAPTER 1

INTRODUCTION

The use of manual methods for classifying the important characteristics of visual landscape has been well documented (Countryside Commission, 1970). They have been driven by the need for landscape evaluation.

> "It is only following the identification and organisation of these diagnostic characteristics [of landscapes] into a system that consideration can be given to questions of evaluation. ... [C]lassification is an essential first step to the evaluation of any resource, including landscape" (Countryside Commission, 1970, p.27).

A landscape character classification is fundamental to landscape research because it provides an important frame of reference for researchers to communicate and compare their work. Landscape research is needed not only to understand landscapes but also for landuse planning. In particular, planners need to know how development can be incorporated within the landscape so that it does not unduly compromise the perceptual quality of the landscape. Despite this need, manual landscape classification has had very little success because of technical and cost issues. The classification of landscapes is a complex problem that has yet to be sufficiently resolved because of the complex nature of landscapes. The principal research question that will be investigated in this thesis is whether Geographical Information Systems (GIS) and national digital databases (NDDB) can be used to classify landscape character. This study will focus on the classification problem, rather than on issues of landscape evaluation.

It appears viable from past work (such as Duffield and Coppock, 1971, Dikau et al., 1991, and Lay, 1991) that recent developments in GIS can partly solve this landscape classification problem. This will be investigated by exploring different options with the use of GIS tools, and by developing an automated process that classifies landscapes. This process will be demonstrated on a transect of the South Island of New Zealand that has a wide range of landscapes. A set of criteria will be established for assessing the validity of a landscape classification. This will consider the important characteristics of landscapes, as well as general classification principles. This thesis shows how GIS and NDDB can revolutionize landscape modelling, and explores some interesting theoretical issues.

Inadequate information on the visual landscape is now a major concern in New Zealand as impacts on the landscape are one of the most controversial environmental issues resulting from development initiatives (Jackman, 1988). This is particularly the case with respect to two of New Zealand's main growth industries - tourism (Collier, 1991), and forestry (Kilvert and Hartsough, 1993). It can be argued that landscape perception needs to be integrated with other landuses to maximize the total value to society. The value of the landscape can be easily compromised by different landuses. In New Zealand, the booming tourism industry, although dependent on the landscape, is actually changing it through the construction of hotels, gondolas, roads, and other infrastructure. If this is not carefully planned, it could diminish the landscape resource that it is dependent on. Commercial forestry is another example of humans altering the landscape on a large scale. Although the scale of indigenous logging in New Zealand has substantially diminished in the last decade, exotic plantation forestry is expanding. The establishment of exotic plantations changes the character of the landscape. This may be having significant consequences on the landscape and its associated values. The Marlborough Sounds is an example where this is happening, and the Mackenzie basin is an example of where it could happen if proposed forestry plans are accepted (Boffa Miskell, 1993). Research and monitoring are required. Whether landscape values are significantly compromised by different landuses depends on the landuse in question, the landscape context, the spatial context, and the observers of the landscape. Some landscapes are more sensitive to development than others due to their proximity to tourist circuits or urban recreational areas, or because they are regarded as natural. This sensitivity to development depends on whose perspective, for instance the developer or the conservationist. Because of all these considerations, research on landscape values is complex, yet essential. Leopold (1969) argues that quantitative data on landscapes are required in order to empower their protection from conflicting landuses. Often landuses that conflict with landscape values are proposed by developers who employ strong quantitative arguments, while the value of landscapes has been dependent on emotional pleas from environmentalists. Leopold's view in 1969 was that environmentalists should begin to support their arguments with numbers. This view is still valid today. The Resource Management Act 1991 makes it a statutory requirement for regional councils to monitor and provide information on New Zealand's significant landscapes, and makes provisions for their protection. Thus, resource managers, developers, and conservationists require landscape information. The utilization of Geographical Information Systems (GIS) and national digital databases appears to offer an effective method for providing parts of this information.

In the last ten years there has been a dramatic change in the utility and power of Geographical Information Systems (GIS), because of advancements in computer hardware, as well as improvements in the GIS software. Closely linked with this advancing GIS technology is the increase in the amount of digital data available to be analysed. This is often referred to as the "fire hose" of data (Maguire, 1991). Significant improvements in automated data capturing devices, such as satellite scanners, airborne scanners, Global Positioning Systems (GPS), and office scanners and digitisers, have dramatically increased the amount of digital data available for describing and monitoring the environment. An inventory of available digital databases in New Zealand was compiled by the Department of Statistics (1992). This inventory reveals the significant amount of data available for reporting on the state of the environment. The challenge is to analyse and present this data so that it becomes useful information for decision makers. GIS can play an important role in this.

Perhaps the most significant databases now available are the topographical databases developed by mapping agencies all over the world. Because of advances in GIS and automated cartography, standard topographic maps are now being produced in digital format. This means that topographical maps, covering extensive regions, can now be analysed using GIS. In fact, the whole world can be analysed using global databases, such as the Digital Chart of the World (Environmental Systems Research Institute, 1993). Complex spatial queries over extensive areas can now be implemented automatically with a computer, as in Dikau et al.'s (1991) attempt to classify the landforms of the state of New Mexico. From reviewing such works and from personal experience with GIS, any measurement that can be derived manually from assessing a map can now be derived automatically. Moreover, because GIS can do billions of spatial measurements in short periods, there are some parameters that a GIS can obtain quantitatively from a map that would be impossible to obtain manually because of practical constraints. Considering the importance of maps and the spatial analysis of maps to geography, such technology ought to be a powerful tool for landscape classification. This thesis develops and demonstrates this tool.

Landscape evaluation is an important end use for a landscape classification. Classification is important for the implementation of public preference surveys that ascertain landscape quality, because it provides a frame of reference that enables different research initiatives to be communicated and compared. A landscape classification can also be used for assessing landscape variety and uniqueness, which will be demonstrated using GIS once a classification has been devised. In fact, landscape classification is important to all forms of landscape research because it helps organise our understanding of landscapes and provides a means for communicating about different types of landscapes (Countryside Commission for Scotland, 1970). The basic rationale for this study is to compare the amenity values of scenery against other resource considerations. Landscape research is necessary for improving resource inventories, making carrying capacity decisions, and assessing environmental impacts.

The classification of the landscape is a particularly difficult spatial analysis problem. Landscape is defined as the appearance of the land (Swaffield, 1991). On the one hand, the landscape is a generalisation of the environment because only the larger objects are perceived. However, it also includes the composition of objects, and this makes landscape considerably diverse and complex (Jackson, 1984, and Robinson et al., 1976). This is further complicated by the fact that different observers view the landscape differently (Bourassa, 1991). In addition, classification must be based on explicit definitions (Rhind and Hudson, 1980). Even though landscapes are heterogeneous in nature, it is necessary to identify homogeneity in order to classify them. This is in common with all resources. People identify homogeneity to make sense of reality, and to describe and communicate realities. Evidence of people's cognitive landscape classification is demonstrated by common words, such as "coastal", "mountainous", or "flat", which are, in effect, describing landscape classes.

To attempt to define landscape classes explicitly to a level of sophistication that incorporates the important characteristics of landscapes requires sophisticated quantitative definitions that are too difficult to implement manually. Quantitative manual methods instead have used simple definitions that do not capture the important attributes of the landscape. For instance, the Manchester evaluation method attempted to classify landforms by counting the number of contours in a one kilometre grid cell (Penning-Rowsell and Searle, 1977). It will be shown that the landform features important for landscape classification cannot be accurately defined in this way. More commonly, landscape classification practitioners have avoided quantitative definitions, and instead used more intuitive approaches, as in the Auckland Regional Authority (ARA) landscape study (ARA, 1982). The intuitive approach suffers because it cannot be repeated by different practitioners, making it difficult to compare landscapes in different regions. Considering that the main purpose of landscape classification is to provide a frame of reference for communication and for describing and comparing landscapes, this is a severe limitation. In comparison, GIS approaches are totally explicit and repeatable.

With GIS, it appears that sophisticated quantitative definitions of important landscape characteristics can be implemented and applied to extensive areas. GIS has been used for analysing related phenomena such as cliffs and farms (Barbanente et al., 1992), visibility (Miller, et al., 1994), wilderness (Lesslie et al., 1988, and Kliskey and Kearsley, 1993), and extracting terrain information (Lay, 1991, Cowen, 1993, Dikau, 1989, Tang, 1992, and Weibel and DeLotto, 1988). These works are useful, not only because the features studied are an important part of landscape, but also because the techniques and structural frameworks that they use can be applied to landscapes. However, when landscapes are classified as a whole, generalisation becomes a complex issue. Past research that has concentrated on individual components has not had to develop classes that are overall impressions of many different components, therefore many issues remain unresolved. Automated cartography literature on generalisation (Shea, 1991) and semantic data models (Nyerges, 1991) also provides useful frameworks that can be incorporated in an automated landscape classification. Since automated landscape classification is relatively new, dating from the release of commercial GIS in the late 1980s, it is necessary to bring together many fields of study that have some expertise in different aspects of automation. Mitchell (1993), and the Countryside Commission (1988) have commented on GIS as a possible future direction for landscape research, but there does not appear to have been any research initiative that tackles the application of GIS to landscape classification directly and fully.

The information in NDDB that can be used for landscape classification includes roads, railways, transmission lines, rivers, lakes, coastlines, and contours, which are all available from topographic databases. Also obtainable are vegetation classes from specialized vegetation databases, and population information from census databases (Supermap2). If GIS and NDDB prove to be valuable tools for landscape classification, then this could have important implications for the development and use of NDDB. The amount of information (in different layers) within NDDB has generally been kept to a level that can be adequately displayed at the scale mapping agencies publish their hard copy maps as these have often been the primary source of information. Yet, GIS can analyse information that is much more detailed. The data models used for NDDB have mostly been in vector format, but GIS can also use raster format, which is perhaps better for some spatial analyses within a GIS. It appears that significant improvements can be made to NDDB to realise the full potential of automated spatial analysis.

Chapter 2 presents the research problem of this project. This includes the meaning and complex nature of landscapes, a consideration of why landscapes need to be researched and classified, a brief outline of landscape research, and a list of criteria that a landscape classification should comply with.

Chapter 3 frames automated landscape classification as an operational definition problem. In a GIS context, Lay (1991) identifies three factors that need to be balanced with operational definitions: the human concept model (for landscapes this is discussed in chapter 2), characteristics of the digital databases, and GIS capabilities. A brief overview of GIS capabilities is given, followed by a detailed description of focal neighbourhood functions as these are important for landscape classification. Appropriate databases are then discussed and identified. This discussion on operational definitions incorporates theoretical input from automated cartography literature. Although this has a different objective to landscape classification, both are concerned with automated abstraction of structural geographical meaning. Nyerges (1991a and 1991b) identifies four important types of abstraction: classification, association, generalisation, and aggregation. To classify landscapes, these abstractions need to be represented using GIS functions. National digital databases contain geographical meaning, though the objects within them can be further formulated to identify even more complex geographical meaning, such as landscape classes. With landscape classification the most difficult abstractions to represent are generalisation and association, and these are discussed in detail. Chapter 3 also describes the method of investigation, discusses validity, and introduces the study area.

The process for classifying landscapes is subdivided to simplify the task. Vegetation, naturalness, and water are classified in chapter 4. Many characteristics of these three landscape attributes are already conceptualised in existing databases. In chapter 5, landform is classified. This is a complex process because a contour coverage must be conceptualised. Chapter 5 also introduces an application of fuzzy set theory. Chapter 6 combines the vegetation, naturalness, water, and landform classifications to produce a landscape classification. The validity of this resulting landscape classification is then discussed using criteria established earlier in the thesis. Conclusions then follow in the last chapter.

CHAPTER 2

THE NATURE OF LANDSCAPES

2.1 Introduction

This chapter explores the nature of landscapes by addressing the definition of the term "landscape", discussing landscape planning issues, reviewing landscape assessment research, and identifying important characteristics of perceived landscape. This serves several purposes. First, it provides a rationale for this study by arguing that landscape is a resource with values that need to be reconciled along with other landuses. It demonstrates the need for landscape research and highlights the importance of classification. Lastly, the identification of important characteristics of landscape enables the establishment of a set of criteria that a landscape classification should incorporate.

2.2 Definition

"Landscape as a concept is bedevilled by semantic differences, misunderstanding, and controversies" (Countryside Commission for Scotland, 1970, p.1). Despite this, three common interpretations of the word landscape can be deciphered. Landscape Ecologists and some Landscape Architects use the word as if it is synonymous with the word "environment". In this context it has been defined as "the total spatial and visual entity of human living space, integrating the geosphere with the biosphere and the noospheric man-made artifacts" (Naveh and Lieberman, 1994, p.4). Fabos (1979, p.4) defined this understanding of the term landscape as "a homogeneous segment of the environment (including the surface of the land, the air, and all useful resources) which support all living creatures."

Within physical geography, "landscape" has often been used in relation to the physiographic, geological, and geomorphological features of the earth's crust (Naveh and Lieberman, 1994). In this context the word "landform" or "topography" would be more exact and is used by most physical geographers.

The third meaning of "landscape" is the "environment perceived, especially visually perceived" (Appleton, 1980, p.14), or the appearance of the land (Swaffield, 1991). The Countryside Commission (1970) used the phrase "the spectacle presented by the countryside"(p.2). This is the meaning intended in this thesis. More precisely, it can be defined as the overall impression obtained from viewing the land (environment) from a reasonable distance. Land includes the flora, fauna, cultural developments, surface soil and rock, landform, and water, but not all these can be perceived from a distance. Often this perceptual connotation of landscape is called scenery. The term landscape used in this way can also be found in early literature. A very early reference is the Book of Psalms, where it is used in reference to the beautiful overall view of Jerusalem (Naveh and Lieberman, 1994). The nature of perceived landscapes is discussed in detail in section 2.9.

2.2.1 Natural versus cultural landscapes

Within landscape assessment, it is common for natural and cultural landscapes to be treated separately (Auckland Regional Authority 1984). Cultural landscapes incorporate human modification and heritage links (Jones, 1991), while natural landscapes focus more on natural components. Although it is very difficult to completely separate the two, as they are very much interlinked, certain landscapes are heavily modified, such as urban areas, while others are not. With natural landscapes, landform and vegetation are important (a point that will be discussed in more depth in section 2.9), while with urban landscapes, the architectural style and layout are important. It can perhaps be said that with natural landscapes people are orientated towards the attractiveness of the existing landscape, while with urban landscapes people look to see what improvement or renovations can be made (Auckland Regional Authority 1984). Although this study is concerned mainly with natural landscapes, there is an interface between natural landscapes and cultural landscapes, albeit not very well defined, which is important. The degree of human modification in natural landscapes affects the character of natural landscapes. Therefore, the influence of cultural landscapes will be considered in this study by means of a classification of naturalness. The detailed composition and classification of cultural landscapes will not be a part of this study.

2.3 Landscape as a resource and a landuse

Landscapes have been considered a resource by many authors, such as Cloke and Park (1985), and Mitchell (1993). They can be seen as a resource within the classic framework proposed by Zimmermann (1951) which provides a functional interpretation of resources as relevant today as when it was first proposed in 1951 (Mitchell, 1993). Zimmermann argued that parts of the environment are not a resource until they can satisfy human needs and are therefore valued. For example, coal was not a resource until people found utility for it and wanted it. With this interpretation the landscape is a resource since people value landscapes.

It is useful to divide landscape values into three categories - economic, environmental, and ethical (Jackman, 1988). The economic value system is perhaps the most widely recognised, but could be argued to be the least important. There can be no doubt that landscapes have considerable economic value. The obvious example of this value can be seen in the tourism industry. New Zealand's tourism industry is totally dependent on its landscape because it is this that attracts the tourist. Currently, international and domestic tourism in New Zealand is an $8.2 billion per year industry (Statistics New Zealand, 1994). The economic value of landscapes may also be realised in other ways, for instance from the flow of wealthy immigrants attracted partly by the quality of New Zealand's landscapes, or the country's "clean green" marketing image.

The environmental value system is associated with the quality of the environment and the quality of life. Although difficult to measure, this quality of life value is particularly significant because landscapes are all around us and are experienced daily. It is possibly of higher value than the economic value. How happy would New Zealanders be if New Zealand was totally flat, urban, unforested, and landlocked?

The ethical value system is defined as the expression of the culture of the people. It includes society's spiritual or religious beliefs, and may include the perceived relationship that a culture may have with the land. Land ethics, such as the right to own land, and concepts like sustainability are all part of this system. With landscapes, the ethical values can be significant, as demonstrated through experiential research such as Hay (1990). People over time often develop a cultural and spiritual bond with landscapes that they have become familiar with, either through work, leisure, or home environment.

Zimmermann explains that resources are dynamic because they become available to people through a combination of increased knowledge, expanding technology, and changing individual and societal objectives. This dynamic is evident with landscapes. There has been a growing awareness of the significance of the aesthetic value of different landscapes. This is evident by the growing number of amenity groups, preservation societies, and general environmental lobbyists (Lowe, 1977). This awareness includes all values - economic, environmental, and ethical, although important documents such as the "Brundtland Report" and "Agenda 21", and legislation such as the Resource Management Act do put emphasis on environmental and ethical values. Not only do landscape characteristics change over time, but "the way humans view and value landscapes changes over time. Therefore, the human "measuring instrument" for observing landscape change is not fixed" (Cary, 1995, p.1). The implication of this for landscape classification is that the classification must be flexible - a criterion given in section 2.8.

2.4 Landscape planning issues

The impact on the landscape is now a major problem for many development initiatives. The planning process for the construction of the new Bealey Hotel, near Arthur's Pass in New Zealand, took four years to get the necessary planning consent (Brabyn, 1991). The main issue was whether it was preferable to have sporadically developed landscapes or to have development intensified in particular locations, like towns or cities, and leave the rest of the landscape in a relatively natural state. Collier (1991) identified that the impact on the landscape was now the most controversial environmental impact of tourism development in New Zealand. This controversy is likely to continue. Kearsley and Gray (1993) in their review of infrastructure requirements to meet the demands of the increasing tourism industry in New Zealand, drew attention to several new road links that may be needed, such as a direct link between Queenstown and the Milford road. The impact on the landscape is also a problem for the forestry industry. This is evident by the planning required for establishing plantations in the Mackenzie Basin (Boffa Miskell, 1993) and work on the public perception of forestry operations (Kilvert and Hartsough, 1993). This landscape issue is not confined to New Zealand but is a major global problem. Large amounts of landscape planning and research have been undertaken in Great Britain (Countryside Commission, 1988), and in the United States (Itami, 1989) for various local and national government organisations.

Landscape planning issues can be seen as a conflict of landuses, which can also be interpreted as a conflict of values, and a conflict of scales. Landscape is inextricable linked with other landuses, such as preservation, forestry, and farming, because these landuses are a part of the perceived environment. Such landuses can add or detract to the quality of the landscape, depending on whose perception is considered, the landscape context, spatial context, and the landuse in question (Amedeo et al., 1989). Such compromises can be within a particular value system, or between different value systems. The economic gain from commercial forestry could compromise the economic gain from tourist viewing the landscape, similarly it could affect the quality of life and spiritual fulfilment the landscape offers. Landscape perception can, as a landuse itself, cause conflict because landscape perception through tourism can provide economic value, but the facilities to provide for this, such as roads and viewing towers, can generate costs economically, environmentally, and ethically. The values provided by the landscape and different landuses can vary with scale resulting in conflict being even more difficult to resolve. On the west coast of the South Island, the locals may value the indigenous forests more for their timber than their aesthetics. While on a national scale, the forests may be more valued for their aesthetics. The question that then needs to be addressed is, whose values are more important? Whatever the answer, decision making will be more informed if information on the importance of landscape values at all scales is known. Therefore landscape research needs to be conducted at all scales, with appropriate levels of generalisation.

The landscape and spatial context are also an important consideration. With the Bealey Hotel example previously given, the landscape context was mountainous and reasonably natural as it was close to Arthur's Pass National Park. If the landscape context had been different, such as the Canterbury Plains which are flat and developed, it is likely that there would have been considerably less controversy. The Bealey Hotel is also located next to a busy road frequented often by tourists. This spatial context means that the hotel is highly visible. If the hotel had been located in another spatial context that was less visited, this too could have reduced the controversy. However, these alternate locations would probably be unacceptable to the tourism industry.

Throughout the world, planning agencies have been forced to consider landscape values because of statutory laws. In the United States there is the "National Environmental Policy Act 1962", in Norway there is the "Nature and Conservation Act 1970", and in New Zealand there is the Reserves Act (1977), which establishes provisions for the "preservation of representative samples of all classes of natural ecosystems and landscapes..." (section 3 (1) (b)), the Conservation Act, 1987, which gives power to the Department of Conservation to advocate conservation, and the

Resource Management Act (RMA), 1991.

In the RMA, landscape values can be considered under the general umbrella of environmental values, which are provided for throughout the Act. However, in several sections specific reference is given to landscape values. For example, in Section 6 - matters of national importance, resource planners need to provide for:

> "6(a) The preservation of the natural character of the coastal environment", and
>
> "6(b) The protection of outstanding natural features and landscapes from inappropriate subdivision, use, and development."

Under section 7(c), particular regard to "the maintenance and enhancement of amenity values" is also required.

2.4.1 Landuse information modelling - the path to conflict resolution

To satisfy the requirements under sections 6, and 7 of the RMA, a comprehensive landscape assessment programme is required. The Resource Management Act, Section 35 -Duty to gather information, monitor, and keep records, makes this explicitly clear. It requires that:

> "(2) Every local authority shall monitor -
>
> (a) The state of the whole or any part of the environment of its region or district to the extent that is appropriate to enable the local authority to effectively carry out its functions under this Act;
>
> (b) The suitability and effectiveness of any policy statement or plan for its region or district."

An example of a response to these statutory requirements is the Canterbury Regional Council's (1995) policy regarding landscapes. This is to "protect landscapes' aesthetic values" (p.94), which they intend to do this through information provision and monitoring of trends. A landscape study has already been completed (Canterbury Regional Council, 1993).

The Department of Statistics (now Statistics New Zealand) and Ministry for the Environment (1990) have pushed the concept of State of the Environment Reporting (SER). SER is defined as "the systematic analysis, description, and presentation of credible, scientifically based information on environmental conditions and trends, and their significance to human activity and its effects on the biosphere" (p.12). The main product from this is a national State of the Environment report. Several countries now produce regularly such reports, notably Canada and the Netherlands. The OECD reports five-yearly (OECD, 1991). By comparing SERs over a period, trends can be identified.

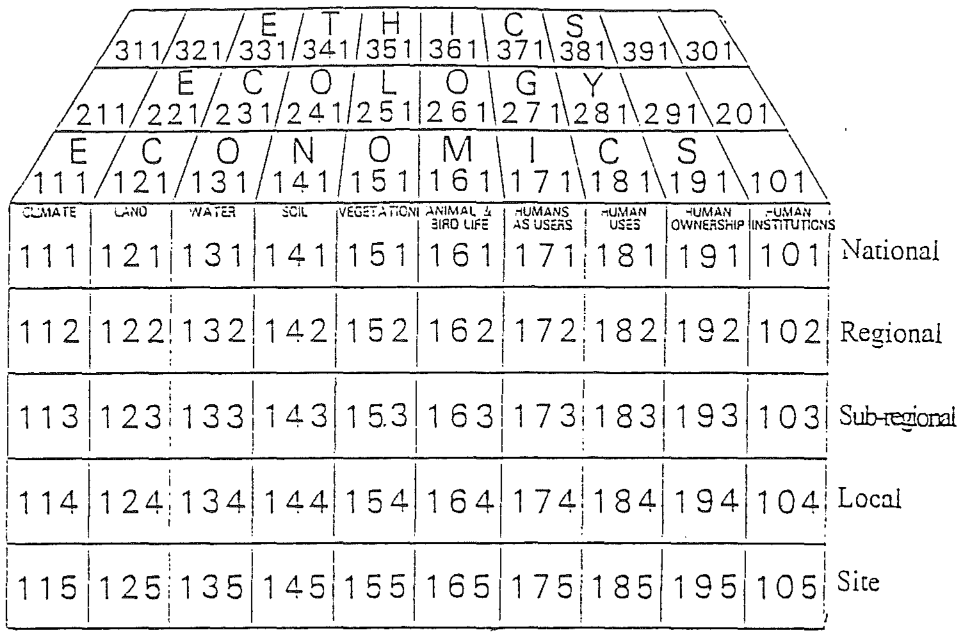
To implement SER in NZ, environmental monitoring is required (Ward, 1991). Ward and Beanland (1992) have consequently determined a list of appropriate environmental indicators to be used for monitoring the environment. The affect on aesthetics is listed as an issue, but no indicator is suggested. As will be discussed in section 2.5.1, public preference can be used as an indicator of landscape quality using psychophysical assessment. It is important that the indicator used is standardised so that national reports can be aggregated from regional reports.

The United Nations, through its environmental programme, is developing a Global Environmental Monitoring System (GEMS). One of the most important tasks is the harmonization of environmental data so that national and global assessments can be implemented. The data are being stored in digital form and the intention is to establish a central global resource information database (GRID) that can be accessed from every country through computer networks (United Nations Environmental Programme, 1990).

Jackman (1988) proposed a comprehensive framework on which landuse information could be structured for decision makers (refer Figure 2.1). The framework groups landuse into ten components. Landscape fits in the category "humans as users". For each component, economic, ecological, and ethical values are determined for five different scales ranging from national to site. It is recognized that each component and value system is connected, but separation is required to determine the linkages between them. From this parametric approach to planning, tradeoffs between different landuses, value systems, and scales can be assessed enabling planners to be proactive, rather than reactive to planning issues. The development of such a model may seem a formidable task, but computers, in particular GIS, are providing useful support for capture, storing, analysing, and retrieving such information. It is a model that planners can work towards for identifying information deficiencies. The Canterbury Regional Council, for example, has a GIS that contains information on most of the ten components. However, information discerning the different value systems at different scales still needs to be developed. Over time, research and the development of planning tools will enable planning authorities to use increasingly sophisticated models of landuse that one day may approach the model proposed by Jackman. The computer aided study for optimizing the location of transmission lines in New Zealand (Electricity Corporation of New Zealand Ltd, 1988) is a good example of a study that approaches this model.

For such parametric models of landuse to work, information on all the different landuses is needed. Jackman (1988) showed that there were major deficiencies in information on landscape values in New Zealand. Compared to the amount of research done on other landuses there is very little assessment of landscapes. Yet landscapes may be of equal or more value to society and more vulnerable than those other landuses. The benefits of research on landscapes may be just as productive for enhancing or maintaining value to society as research in other landuses. This leads to the question: what does landscape research involve?

Figure 2.1 Jackman's (1988) Landuse information framework



2.5 Landscape research

Zube, Sell and Taylor (1982) derived a landscape research framework after reviewing 160 landscape articles, covering 20 different research journals. They identified four different research paradigms. These have been labelled expert, psychophysical, cognitive, and experiential. The expert and psychophysical paradigms are concerned with applied landscape assessment and seek to determine what landscapes are significant or beautiful, and which are not. They are strongly motivated by the pragmatic concerns of resource planners. The cognitive and experiential are more concerned with theoretical issues, such as the nature of landscapes, why people have preferences for particular landscapes, and the meaning people attach to particular landscapes.

Daniel and Vining (1983) developed a similar framework using five groups - ecological, formal aesthetic, psychophysical, psychological, and phenomenological. The main difference between these frameworks is that the latter has an extra paradigm called "ecological", which has been included within the expert paradigm in

Zube, Sell and Taylor's framework. Otherwise, both frameworks are very similar, except the labels used. The frameworks have been accepted by landscape researchers as a valuable frame of reference for assessing different approaches to landscape assessment and research, and are often cited, for example Dearden (1989).

Steinitz (1993) also provides a useful research framework for addressing landscape issues. With his 25 years experience working with GIS and landscape planning he realized that there was an "overwhelming (and perhaps necessary) structural similarity among the questions asked by and of landscape planners and other environmental design professionals" (p.42). His proposed framework consists of six questions:

1. How should the state of the landscape be described: in context, space and, time?

2. How does the landscape operate? What are the functional and structural relationships among its elements?

3. Is the current landscape functioning well?

4. How might the landscape be altered: by what actions, where and when?

5. What predictable differences might the changes cause?

6. Should the landscape be changed? How is a comparative evaluation of the impacts of the alternative changes to be made?

These questions summarise the breadth of landscape research that is being implemented. The first two questions are more theoretical and concerned with the nature of landscapes, while the remaining questions are more concerned with landscape assessment, which is discussed in the following section. The nature of landscapes is discussed in section 2.9.

2.5.1 Landscape assessment

There has been much discussion on methodological issues regarding landscape assessment. In particular, what is the preferred approach - the expert or the psychophysical? The main factor that distinguishes the expert approach is that it is based on the judgement of experts who have been trained in the field of aesthetics. There are two kinds of experts, one that has had training in the field of fine art perspectives, and the other expert who is more ecologically orientated.

The fine art expert uses the formal artistic properties, which have been defined as form, lines, colours, and textures (Bennett, 1985). From these basic elements, landscape architects determine what are called the principle determinants of landscape quality. Daniel and Vining (1983) have called these variety, harmony, unity, and contrast. These determinants vary with different experts. Boffa Miskell (1993) and Ministry of Works and Development (1987) used vividness, intactness, and coherence to determine beauty. In the Boffa Miskell study of the Mackenzie Basin "vividness" was defined as the memorability of the visual expression received from contrasting landscape elements. "Intactness" was defined as the integrity of visual order in the natural and man-built landscape, and the extent to which the landscape is free from visual encroachment. "Coherence" was referred to as the way the landscape "hangs together" and is explicable, particularly in terms of its natural formative processes. Even when definitions are given, as with the Boffa Miskell study, the definitions of such criteria are still ambiguous. What is meant by "hang together" or "integrity of visual order"?

The ecological expert assumes that particular attributes of the landscape, such as naturalness and mountains, are of high value and uses the presence or absence of these attributes to judge beauty. Linton (1970) exemplifies such an approach. These are bold assumptions that are questionable, although they may be proven correct with cognitive research using preference surveys. These attributes are reasonably objectively defined, using parameters such as relative relief, presence of roads, and particular vegetation. They are also familiar concepts to the general public.

The psychophysical approach is based on the work of classical psychophysics, which sought to establish quantitative relationships between physical features of environmental stimuli and human perceptual responses (Daniel and Vining, 1983). Landscape quality is determined by the public as opposed to experts. This approach typically involves two stages. The first stage is the classification of the landscape into reasonable homogeneous classes, which are usually based on aspects of landform, landcover, naturalness, and water (eg. The Auckland Regional Authority study, 1984). The second stage is the evaluation of quality of the different classes within the classification. This is done with a public preference survey and with the classification acting as a frame of reference. With public preference surveys, typically, samples of the public are questioned about how they rank each landscape class for its visual quality. The survey could require a sample of the public to visit landscape sites, but usually photographs of the classes are used as a surrogate, and these photographs are ranked. The psychophysical approach might identify a consensus within society, but it could also show that different groups in society value landscapes differently.

Both the expert approach and the psychophysical approach have their strengths and weaknesses. The expert approach sacrifices reliability (repeatability) for utility and is more concerned with getting the job done. It is the more practical approach but is considered invalid because it claims that quality is inherent in the landscape (Daniel and Vining, 1983). The psychophysical approach, on the other hand, claims that quality is related to both the landscape and the observer, which is consistent with landscape theory. The psychophysical approach achieves a high level of precision and consistency but at the expense of generality and resources. It is the more scientific method as it is more concerned with measurement that is free from the bias of the researcher. Daniel and Vining (1983, p.79) concluded that;

"...no other approach has come so close to meeting the criteria of the ideal assessment system".

The psychophysical approach attempts to combine the cognitive research on the subject (ie. the viewer) with the object (the physical landscape). For scholarly enquiry, factors such as reliability and validity are more important considerations than generality and utility. If GIS can be used to classify landscape character, it is hoped that the cost of the psychophysical approach will be substantially reduced, and that it can be applied readily to a wide range of landscapes.

2.6 The purpose of landscape classification

This study is principally concerned with developing a methodology to classify landscape character. This is an important task that requires specialization because of technical developments in information and its processing.

There is a distinction between landscape classification and landscape description. Classification groups objects into categories, while description does not (Countryside Commission, 1988). Description has traditionally been the common means of communicating about landscapes. It describes a particular landscape in a way that conveys a clear picture. This approach is commonly used by Landscape Architects for analysing a site for a proposed development, or by writers who try to evoke the character of particular landscape in literature. However, description can be an inefficient means of communicating about a resource. If landscapes are similar then description can be repetitious. Descriptions are also difficult to further analyse and evaluate.

Classification is important to science because it provides a frame of reference that enables different researchers to communicate their results effectively. It also helps order and structure what is known (Haines-Young and Petch, 1986). In fact, classification is an important part of cognition (Langridge, 1992). The importance of classification for landscape research is no exception (Countryside Commission, 1970). The psychophysical approach to landscape assessment, as discussed previously, is based on a landscape classification. Without a classification the approach would be of little use. Without a landscape classification, landscape researchers are unable to effectively communicate their discoveries, and as a result a body of theoretical knowledge will be slow to develop.

To derive a landscape classification one must first consider the specific purposes to which such a classification is to be used. Objects can be classified in many ways. Objects are assigned to classes according to the characteristics that they have in common, but even the simplest objects have many features that could be used for this purpose. A simple table can be classified by its size, colour, design, style, etc. The choice of classification criteria is related to purpose (Langridge, 1992).

It is intended that the landscape classification developed in this study be used for landscape research and planning. The information demanded by planners was discussed in section 2.4. This includes the relative qualities (values) of different landscapes and the tradeoff in landscape quality associated with landuse change. If it is quality that is the focus of planning, then the landscape classes within a classification must distinguish this. This is not saying that the classification should identify quality. It is hoped that by dividing the landscape into homogeneous character classes, then the quality is also being divided into homogeneous classes. This may or may not be the case. It is possible that someone may perceive differently the quality of two landscapes that are reasonably identical in character, but located in separate areas. For example, people may value the area where they live more than another area that is similar in character, but far away from where they live. However, it is fair to say that the perception of landscape quality is dominated by landscape character and that this should therefore form the basis of a landscape classification.

2.7 Manual landscape classification

Conventional methods for classifying landscape character have relied on manual techniques (eg. Auckland Regional Authority, 1984), whereby maps, photographs (aerial and ground based) and field observations are used. Usually the classification process has not been explicitly stated. This is because of the large amount of information that needs to be analysed. Practitioners have avoided using computational and quantifiable methods and have instead used subjective approaches that Densem (1980, p.8) calls, "gut reaction". Without the use of GIS and detailed databases this has really been the only feasible option.

The manual approach, if done properly by operating with strict definitions of attributes measured on maps or in the field, becomes incredibly labourious and tedious. For example, relative relief can be used to differentiate between a mountain and a hill, but relative relief is dependent on scale. So it is necessary to use relative relief within a certain area, say a grid cell, examine a topographic map of the whole study area with a grid template, and then within each grid cell calculate the relative relief. This is time consuming and may still give the wrong answer as a flat area on top of a mountain may occupy an entire grid cell and thus get classified as flat when it should be classified as mountainous. Possible solutions are to consider the neighbouring grids cells as well, but this would make the process even slower. Alternatively use a larger spaced grid, but then accuracy will be lost. As of yet, no universal classification of landscape character exists that has clearly defined definitions. This is probably because the technology to implement these definitions has not existed until recently.

Manual techniques instead have generally used ambiguous definition involving more intuitive methods, whereby, for example, a contour map is viewed and areas that appear to be mountainous are defined as mountainous and areas that appear hilly are defined as hilly, or areas that look natural defined as natural and areas that look forested defined as forest. The problem with such an approach is that it would be impossible for other researchers to replicate it exactly, as the definitions of the classes would not be known. The inconsistency that results would mean that two independent studies could not be compared. This means that such landscape classifications can only be used as a frame of reference for the particular study where they were developed. For example, if a landscape classification was completed for the North Island of New Zealand by one person and a classification of the South Island was completed by another person independently using different definitions, it would then not be possible to say, based solely from these two classifications, that the South Island is more mountainous than the North Island.

Even when all the attributes have been mapped for the study area, compositions of these attributes need to be considered. Manual overlays are often used but this can be a time consuming task with many problems. It is often necessary to weed out unnecessary classes, or rework the generalisation process to cut down on the number of classes. With the Auckland Regional Authority (ARA) (1984) classification, 85 classes were derived. Yet, with the combination of different attributes used, clearly more classes would have been identified. How this reworking was done was not reported. Again, explicitness was compromised for expediency.

Another problem with the manual approach has been the high costs involved in time and resources. If it is necessary to do field observations of every landscape then this is going to cost a considerable amount. Because of these high costs only sporadic areas have been classified in New Zealand.

Landscape classification requires identifying areal units or enclosed areas that aid analysis. With manual classification the options available for this have been (Robinson et al., 1976):

1. the character tract (usually based on macro landforms),
2. the viewpoint and its associated visual envelope (viewshed),
3. the grid square, and
4. the whole study area.

With GIS, a new areal analysis option is available that is not practical to implement manually. This is the focal neighbourhood function, which opens new possibilities for landscape classification, and will be discussed in detail in the following chapter.

With GIS and a computer automated approach, many problems encountered with manual classification can be considerably reduced. Computers force one to define exactly what procedures are being followed, there is a high degree of consistency between applications, different information layers can be easily integrated, the results can be easily displayed and used for further analysis, and all this can probably be done at less cost than the manual approach. A fuller discussion of this is given in the following chapter. It remains possible that the manual approach, by using field observations and intuitive methods, can capture more of the subtleties of landscapes than an automated GIS approach. Therefore, a comparison of the two approaches is required. This is done in section 6.4.4, after an automated process has been demonstrated.

2.8 Landscape classification criteria

The classification of landscapes has many problems. One of the main problems is that researchers involved in landscape classification have originated from many different professions and do not classify landscapes from the perspective of the general public. The classifications they produced were often too detailed, especially regarding information relating to their original profession. This problem is also compounded by the fact that "landscape" is an ambiguous term and people have different interpretations of what should be in a landscape classification. As a result many classifications that claim to be landscape classification are not landscape classifications using the definition of landscape adopted in this thesis. The Protected Natural Areas (PNA) program is an example of this. Here the protection of representative landscapes was an objective and a landscape assessment was required (Myers, et al., 1987), but a detailed geomorphological and botanical classification was developed. The landscape classification produced by Canterbury Regional Council (1993) also falls far short of being a landscape classification, since a general land inventory was produced rather than a classification based on the appearance of the land. It is therefore necessary to establish criteria for a landscape classification. This should consist of general criteria that apply to all classifications, and specific criteria that apply only to landscape classification. Specific criteria will be listed in the following section after the important characteristics of landscape are discussed.

General classification criteria are listed below. They have been adapted from Rhind and Hudson (1980) who have used them as criteria for a landuse classification.

(i) The classes must be exhaustive and mutually exclusive, i.e. all geographical individuals must be classified, but no individual must fall into more than one class.

(ii) It has to be easily understood and applied.

(iii) It has to produce repeatable results that are independent of the researcher.

(iv) It has to be hierarchical, to cope with needs at different levels of resolution in different areas.

(v) It has to be sufficiently flexible for new interests and tasks to be met from a modified, rather than a completely new, classification.

(vi) It must incorporate some recognition of seasonal or other cyclical changes.

In order for a classification to be repeatable by different researchers, it is necessary that the classification be totally explicit. This does not necessarily require the process to be quantitative, however, this is usually the most efficient means of being explicit. If a classification process has been automated with computers then this criterion will be met.

Regarding criteria (vi), elements of seasonality can be incorporated into a landscape classification by ensuring that within a class the whole class changes similarly through the seasons. For example, low land plains and a highland plateau need to be in different classes as these areas change differently with the seasons because of the differences in altitude. A plateau may get covered in snow in the winter, while a lowland plain may not. This is generally the case in New Zealand but may not be the case in colder or warmer parts of the world. It is therefore necessary to restrict the domain of the classification being developed in this study to New Zealand.

Besides these general criteria a landscape classification must also incorporate the nature of landscapes. This is needed for a landscape classification to be useful for psychophysical research, and for the identification of quality - an important use of a classification. Haines-Young and Petch (1986) say that classification needs to be undertaken in the context of theory. This begs the question, what is the current theory behind the nature of landscapes?

2.9 The important characteristics of landscape

The definition of landscape in section 2.2 focuses the intent of this thesis, but leaves unanswered questions about the important characteristics of landscape. What are the important components of landscape, and what does viewing the land entail?

The environment is different to the appearance of the land (landscape). The environment contains entities that are not commonly perceived, while landscape is restricted to objects above a certain scale. The perception of landscape is dominated by the larger, more visible entities. Many components of the environment can be seen if one cares to look closely, for example small fauna such as lizards, snails, and hedgehogs. However, these cannot be seen from a reasonable distance, and would not contribute significantly to the overall spectacle of the landscape. The definition of landscape in this thesis excludes such visually insignificant components. If, however, there were a sufficient number of a particular small component that together could be seen from a distance, then this group could be part of the landscape. A blade of grass on its own is not a part of the landscape, but a paddock is likely to be. Landscape does not include small isolated patches of mosses, lichens, and small shrubs as they are unlikely to contribute significantly to the overall spectacle of the countryside. Often the surface rock and soil are not a part of the landscape as they are hidden from view, usually by vegetation, or buildings. So, for a landscape classification, it is not necessary to do a detailed assessment of soil, geology, entomology, grasses, or forest undercover. The notion that "landscapes are perceived from a distance" is useful for indicating the minimum degree of generalisation involved.

Is the weather, or the polluted air a part of the landscape? Or is the sun setting on the horizon a part of the landscape? Certainly, weather, air pollution, and the position of the sun affect the view of the land. However, they are highly variable components that are normally held constant in landscape studies, as in the ARA (1984) landscape study. Consequently, they are not generally included in a landscape classification, so will not be considered in this study.

To discuss the important characteristics of landscapes, it is necessary to have a means for conceptualising landscapes. As discussed in section 2.5.1, there are two lines of thought by experts about how landscape components can be conceptualised -either the fine arts or ecological perspective. If two areas are similar in physical components, then there is a high probability that they should be similar in terms of form, colour, lines and texture. It therefore should not make too much difference which approaches are used for identifying landscapes. However, there is likely to be a difference resulting from implementation. The major problem with using the formal artistic approach is the lack of definition about how these artistic principles can be defined. The approach is usually very intuitive and not explicit. This study uses physical components because they can be described explicitly. Common language also suggests that this is how most people conceptualise landscapes, for example mountainous, forested, and coastal. As will is discussed in this section, there has also been research to suggest which of these physical components are important for landscape classification.

The landscape is derived from an interaction between aspects of the environment (landscape components), and human perception processes. Landscape perception depends not only on these physical landscape components, which may be extremely diverse, interrelated, and complex, but also on the values, experience, and social-cultural conditioning of the observers (Dearden, 1989). Jones (1991) describes this complexity as the "elusive reality of landscape"(p. 229) and adds, that in the past, the lack of recognition that landscapes are both a physical reality and a social or cultural construct has led to an "academic battlefield", with different disciplines and schools concentrating on either the physical landscapes or on the observer.

Identification of the important characteristics of landscapes therefore depends on whose perspective is considered. "Beauty is in the eye of the beholder". One of the most important points to consider is whether the observer is attached or disengaged from the landscape (Bourassa, 1991). For example, an urbanite is likely to perceive a farm differently from the farmer, and a farmer is likely to perceive a city differently from an urbanite. Linked to this, is the observer's familiarity with the landscape. For example, a Geomorphologist who has had training in detecting subtleties in the landscape that give clues to the formation of landforms may perceive an interesting diversity in a landscape, while somebody who has not had this training may think that the same landscape is quite monotonous. According to Appleton (1975a), humans have innate preferences for particular landscapes that are common to us all. This has been suggested to be a savanna type landscape that offers both refuge and prospect, because this would have been an important consideration when humans lived as hunters and gatherers.

Since it is the public's perception of landscape quality that is important to planners, then it is the public's perception of landscape character that should be considered in a landscape classification, not the perceptions of specialized scientists (Zube, 1984a). Specialists are a small minority. The total value that they derive from landscapes would be far less than the total value that the general public derive from a landscape. It could also be argued that for many scientists it is not the landscape but the scientific information that is of value.

The difficult questions that then arise are, what is the general public, and what is the nature of their perception? The general public is not a homogeneous group, but as a notion it eliminates extreme perceptions and is therefore a useful notion. A classification should be based on the lowest common denominator within this group. For example, not all people perceive the difference between exotic and indigenous vegetation (international tourists may be included in this group), but many members of the public do, therefore the distinction between exotic and indigenous vegetation is necessary. This lowest common dominator level is not always obvious. Therefore, a classification should be hierarchical, in terms of detail, to cope with a range of different perceptions. Observer perspectives can also be considered using fuzzy set theory. This is discussed in section 5.4.2.

So what are the main components of landscapes from a general public's perspective? These are assumed to be landform, vegetation, naturalness, and water. This assumption is substantiated by a branch of cognitive landscape research, known as content category identification (Amedeo, Pit, and Zube, 1989), concerned with identifying important components of the landscape that explain differences in how the public perceive quality. It is generally understood that people when determining scenic quality, organize sets of landscape components into classes or categories. Categorization of landscapes is a mental process that proceeds when people actively discriminate among landscapes. How people categorize landscape is important information for determining the appropriate attributes to use for a landscape classification. An ultimate goal of content category identification research is to eventually delineate a landscape classification system resulting from public perceptual assessments, rather than professional judgements of scenic value (Amedeo, Pit, and Zube, 1989). Content category identification studies generally determine features perceived as having negative impacts on quality, and those that have positive impacts. This research typically involves using some form of statistical analysis, for example Q-sorting (Amedeo et al., 1989). Often preferences for different photographs of landscape are determined by asking samples of the public to rank them. Many attributes of the landscapes in the photographs are identified and quantified, and related to the preference of the photographs with some form of regression analysis.

Pomeroy, FitzGibbon and Green (1989) used personal construct theory, the repertory grid, and multidimensional scaling to ascertain these attributes.

The Countryside Commission of Scotland (1988) reviewed research involved with identifying those physical attributes of the landscape that determine quality and presented a summary table (refer to Table 2.1). From this, there are some attributes that are consistently identified as determining quality. Man made structures have been identified in all the studies, especially as having a negative effect on quality. This suggests that the degree of naturalness be an important component. Vegetation is identified in five of the six studies, and water features, such as sea, lake and streams have been identified in four out of the six studies. Landform is, surprisingly, only identified in two of the studies. Zube, Sell and Taylor (1982) also identified relative relief, landuse diversity, water, and naturalness, as determinants of quality in their review of this type of research. A recent study by Amedeo et al. (1989) also identified aspects of vegetation, landuse, influence of water, and topography. In this thesis, landform, vegetation, naturalness, and water will be called landscape attributes. The characteristics of these attributes, such as forests, hills, lakes will be called components.

For New Zealand landscapes, there has been no content category research per se. Since landscape quality is specific to cultures, it may be unwise to apply the above results to New Zealand. However, the four attributes identified above have been used for classifying landscapes in Auckland (ARA, 1984), and have been used in various other landscape studies (Mosley, 1989, Kliskey and Kearsley, 1993, and Fairweather and Swaffield, 1994). The components also relate well to common language used in New Zealand to describe landscapes. It appears that these attributes have become fairly standard for describing landscapes.

Table 2.1 Summary of content category research
Source: Countryside Commission (1988)

RESULTS OF "STATISTICAL" EVALUATION METHODS - COMPOSITION OF REGRESSION EQUATIONS

STUDY AND STUDY AREA	NO. OF LANDSCAPE ELEMENTS MEASURED	VARIABLES USED IN EQUATION + effect on quality	- effect on quality
Dearden 1980 Saanich Peninsula, British Columbia, Canada	30	Undeveloped coastline Rocky coastline Unpaved roads Rivers and other watercourses	Industry Airports/cemeteries/firing ranges Greenhouses Residential areas Power lines, Highway Parks
Clamp 1976 Survey line from Colchester to Carlisle (Assessment by public)	34 (subdivided into 301 "content categories") plus a number of "general features" such as undulation, complexitiy	Man-made features (attractive) " " " (indifferent) Woodland Heath and rough land Water features Sky (all weathers), grassland	Arable Man-made (unattractive)
Clwyd County Council 1978 Clwyd, Wales	35	Measures of: steepslopes, deciduous or mixed woodland streams, rocks, bracken and gorse scrub parkland	Urban influence flatness
Briggs & France 1981 South Yorkshire	15	Permanent pasture Deciduous woodland	Industrial land Residential land Hedges Wasteland Railways
Robinson et al 1976 Macclesfield, Cheshire	44	FACTORS: Topography : Upland Urbanisation : Non-urban Cultivation : Marginal VARIABLES: "Other field boundaries" single trees	 Lowland Urban Farmland Industrial buildings in countryside
Durham County Council	36	Deciduous & coniferous woodland Lakes, reservoirs, sea	Quarries, power lines Motorways

Landscapes can not be classified by simply dividing the land into areas that reflect boundaries between different landscape components because it is these boundaries that are important characteristics of landscapes (Jackson, 1984, and Robinson et al., 1976). The interaction of components is sometimes more important than the components themselves (Arthur et al., 1977). An analysis of landscape paintings in the Canterbury region showed a clear preference by painters for landscapes that were compositions of plains and mountains (Canterbury Regional Council, 1993). A landscape classification therefore needs to incorporate this juxtaposition of boundaries between different components of the landscape in order to distinguish landscape quality. This consideration makes landscape classification particularly difficult, but if it is not included the classification will be of little use for researching perceived landscapes. In one respect landscape is a generalisation of the environment, because only the large visual entities are perceived. However, when compositions of these generalised entities are considered, virtually millions of different combinations become apparent. The perceived environment therefore becomes very complex. There can be no doubt that the perceived environment (landscape) is different to the environment that Landscape Ecologists study. Completely different types of classifications are therefore required.

The geometrical perspective from which landscape components are perceived is also important in determining landscape quality (Higuchi, 1988). For example, a view of plains from a high point, such as the top of a hill, is very different from a view from a point that is at the same height as the plains. The field of view will obviously be greater from the higher vantage point. It has also been suggested that the geometric perspective of the observer is dynamic in the sense that landscape is viewed from a multiple of points, surrounds the observer, and is experienced from movement and exploration (Zube, Sell, and Taylor, 1982). For example, when someone looks from the top of a hill into a valley, they may have just spent the last hour driving through that valley. This experience will be with them when they view that valley from above and will affect the way that landscape is perceived. Many components that cannot be seen from the top of the hill, perhaps because of visual obstruction, but had been seen previously from exploration, will still be a part of the perceived landscape. Such memories may be visually disturbing, such as a rubbish dump, or perhaps more pleasing. When people view a landscape from a point, it is not just the area that is directly visible that gives an impression, but also peripheral information that has been previously experienced (Zube, Sell, and Taylor, 1982).

Landscapes have also been identified as being perceived through multiple senses, and not just sight (Zube, Sell, and Taylor, 1982). Porteous (1990) refers to the "smellscape" and "soundscape". However, sight is considered the dominant sense that provides information on landscapes. Porteous (1990) says that "it yields more than 80 percent of our knowledge of the external world" (p.4). It is necessary to simplify landscapes to visual information to enable the study of landscapes to be feasible. This simplification can be further justified by the assumption that the other sensory inputs are likely to vary consistently with sight.

The above discussion shows that there is some consensus on the important characteristics of landscapes. These should be incorporated in a list of landscape classification criteria. It should be noted that landscape character classification is concerned with the stimulus properties of the landscape, not the outcome of landscape perception, which may be meanings, actions, or values. The following criteria need to be adhered to in order to incorporate the important characteristics of landscapes:

1. The classification should incorporate landform, vegetation, naturalness, and water.

2. The classes should be based on the general public's perception of the above attributes.

3. The classes should be based on an overall impression of the above attributes in an area from a distance, and involve generalisation and composition.

4. The classes should recognize that landscapes surround and are experienced from a multiple of geometrical perspectives that can be obtained from movement and exploration.

The list of specific criteria may appear to be over simplified considering the complex nature of landscapes. Many questions are left unanswered - what is the exact nature of the landscape components and their relationship with each other, and what is the exact nature of the observer? If a greater number of more detailed specific criteria were stated, their validity would be questionable. This is because they would not have been substantiated by research and/or there would not be the consensus among researchers. More research is required. However, to do this effectively requires a landscape classification so that researchers can communicate their results. The dilemma is that it is necessary to know the important characteristics of landscapes to classify, yet to know the important characteristics of landscape it is necessary to classify. A classification needs to be developed with what information is available, and then reassessed as information is forthcoming. It is in this way that classifications evolve. Our understanding of landscapes is at a superficial level - "theoretical vacuum" (Appleton, 1975b, p.2), and needs to be based on many assumptions. Despite these criteria being an over simplification, if a classification can meet all of them, then significant advancement will have been made. As of yet, this has not been achieved using manual techniques.

The specific landscape classification criteria listed above and the general classification criteria given in section 2.8 focus the issues that this thesis will address. It is worthwhile stressing that the following considerations will not be addressed in this thesis:

1. Cultural landscapes per se;

2. Entities that are not highly visible such as underlying soils and geology, and micro landforms;

3. Highly variable factors such as the weather, atmospheric pollution, and the position of the sun;

4. Non visual senses such as sound, smell, taste, and touch.

The classification criteria will be used later in the thesis for discussing the validity of an automated landscape classification that will be developed. These criteria are only a minimum that a landscape classification must meet to be valid. For a landscape classification to be valid, it also needs to be verified by many independent researchers that have used it in research. Since the criteria include the need to incorporate seasonal variation, then a classification will only be valid for the particular climatic region that it was designed for. The classification being researched in this study is intended for New Zealand conditions. Although the components of each of the four attributes are not part of the criteria, suitable components will be defined and discussed in chapters 4 and 5.

CHAPTER 3

GIS MODELLING OF THE LANDSCAPE

3.1 Introduction

Rhind (1988, p.26) stated that,

> "existing GIS systems do not contain the ability to express high level geographic concepts. Instead they are entirely or very substantially based upon storage of coordinate data and their attributes - essentially low level conceptualizations of the objects under consideration. Human beings evidently store multiple levels of conceptualization of objects, sometimes in a "soft" or "fuzzy" fashion... ."

From the previous chapter it is apparent that the concept of landscape is complex and so Rhind's statement is questioning whether GIS can express landscapes adequately. This chapter addresses the challenge of classifying landscapes using GIS by framing it as an operational definition problem. In a GIS context, Lay (1991) identifies three factors that need to be balanced with operational definitions: the human concept model, characteristics of the digital databases, and GIS capabilities. The previous chapter has provided information on the human concept model of landscape, and the first part of this chapter will provide a brief overview of GIS capabilities and in particular focal neighbourhood functions as they are important in later chapters. The available digital databases will also be introduced. Nyerges (1991a and 1991b) provides a more sophisticated framework for developing operational definitions by dividing geographical meaning into four abstractions: classification, generalisation, association, and aggregation. The challenge is to represent these abstractions using GIS. This is discussed, especially with regard to representing generalisation and association as these are more difficult. This chapter also introduces the study area, the method of research, and a means for validation. Lastly, past research in automated landscape classification is briefly reviewed but the detail of this is left to later chapters.

3.2 GIS Overview

GIS is a collective term commonly accepted for describing computer systems that can manipulate geographic data. This includes the following operations:

. acquisition and verification
. compilation
. storage
. updating and editing
. management and exchange
. retrieval and presentation
. analysis and combination.

Geographical data can be defined as consisting of information on the qualities of and the relationships between objects that are uniquely georeferenced (Bernhardsen, 1992).

GIS is a relatively new technology that has only become well recognized and utilized with the development of commercial GIS software in the 1980s, although the basic principles were conceived in the early 1960s with the first system, the Canadian Geographic Information System (Maguire, 1989). The key to their enormous value is that they offer users the opportunity to analyse and manipulate large databases, select data by theme, search for particular features in particular areas, and update databases quickly. Also, they can produce a variety outputs, ranging from maps, graphs, data lists, and summary statistics. The benefits, components, and functions of GIS have been thoroughly reviewed elsewhere (Maguire, 1989, Aronoff, 1991, and Cassettari, 1993).

Data models for GIS can be divided into two categories: vector and raster. In brief, a vector data model is represented by points, lines, and polygons, while a raster data model is represented by pixels (commonly called grids). Both data models have their advantages and disadvantages. Raster (or grid) format is a simpler data structure, while a vector data structure can be complex but provides an accurate representation of boundaries and linear and point features. Different data models suit different analysis functions. Overlay and neighbourhood analysis functions are easily computed with raster data models, while vector format is more efficient with network analysis (Aronoff, 1991). With GIS it is now common to have functions that convert data from vector to raster and vice versa. Such functions will be used in this research. Both data models will be used at different stages in this study depending on the analysis functions being used.

A major part of GIS is cartographic modelling (or GIS modelling). This is concerned with how data are used rather than with the gathering, maintaining and conveying of data. It is, as the term suggests, the development of models (or representations) expressed in a cartographic form (Tomlin, 1990), but is more concerned with process rather than a product. Tomlin (1990) identifies two types of cartographic modelling - descriptive and prescriptive. Descriptive modelling describes "what is" or perhaps "what could be", and uses analysis of form and position with synthesis of cartographic characteristics. Prescriptive modelling is concerned with "what should be", and is problem solving, especially regarding allocation (eg. selecting locations to satisfy stated objectives). Landscape classification is dominantly descriptive as it is concerned with describing "what is there".

An important objective of cartographic modelling is to derive meaningful information from what can be an overwhelming amount of data (Cassettari, 1993, Maguire, 1989). Planners need clear single theme models that can then be incorporated in a landuse information model (as discussed in section 2.4.1). A theme identified as important for planners is the landscape. There is a wealth of different databases that could provide information relating to this theme, but these are not helpful to planners who do not have the time to interpret such databases. Automated landscape classification is about converting large quantities of data to useful information.

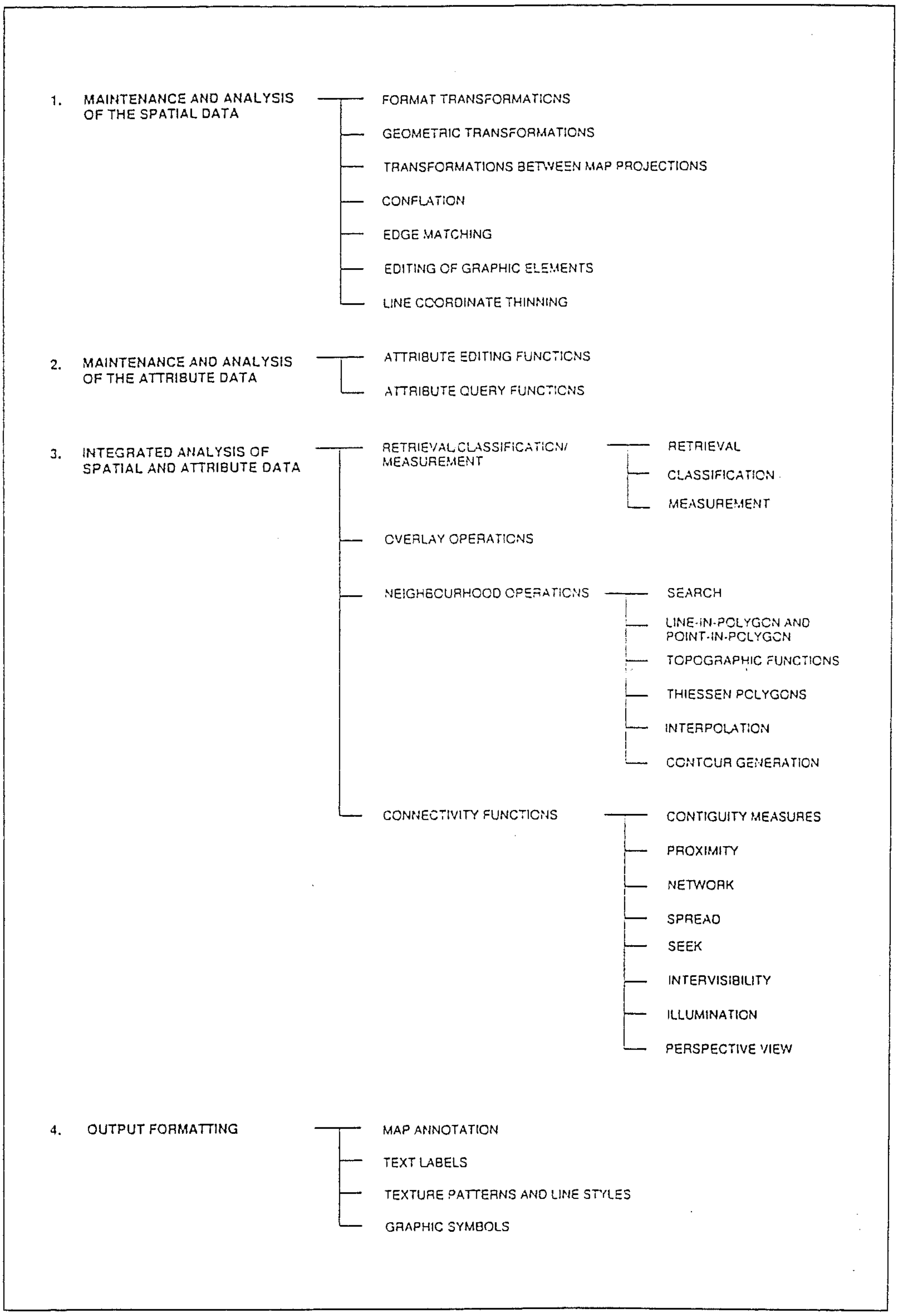
3.2.1 GIS analysis functions

The processing of digital data into information with GIS requires the use of analysis functions. Aronoff (1991) provides a useful classification of GIS analysis functions, which has also been adopted by Cassettari (1993) (refer to Figure 3.1). This research will use and/or discuss many of them. For example it will use: various functions for the maintenance of spatial and attribute data; retrieval, classification and measurement functions; overlay functions; various neighbourhood functions; topographic functions; interpolation functions; some connectivity functions such as, proximity measures, and intervisibility (viewshed); and the output formatting functions. As mentioned previously, raster data models (or grids) are particularly useful for spatial analysis. This is particularly so with neighbourhood analysis. In a vector data model, neighbourhood analysis is virtually limited to the use of buffer zones. With grids there are other possibilities. The more promising of these is, what is commonly called, "focal neighbourhood functions" (Tomlin, 1993). Since this function is an important part of this study it will be discussed in detail.

3.2.1.1 Focal neighbourhood functions

With focal neighbourhood functions each cell within the specified coverage becomes in turn the centre for processing (Figure 3.2). When a cell is being processed, the cell values within the specified neighbourhood of that central cell are included in the processing. The process could, for example, be to calculate the mean of all cell values with the neighbourhood. The result of this process is then assigned to the cell of a new grid with the same position as the central cell. The next cell in the grid coverage then becomes the centre of the processing. This continues for all the cells that are available for processing. As can be imagined, this can involve a lot of processing, especially when the specified neighbourhood consists of many cells, and the grid coverage contains many cells. The neighbourhood can be of any shape which can be directly specified in some GIS software or custom designed. The focal

Figure 3.1 Classification of GIS functions
Source: Aronoff (1991)



function has been used extensively for image processing in regard to remote sensing (Mather, 1987). For this application it is used for different types of filters (kernels).

Figure 3.2 Focal neighbourhood functions
Source: ESRI (1991)

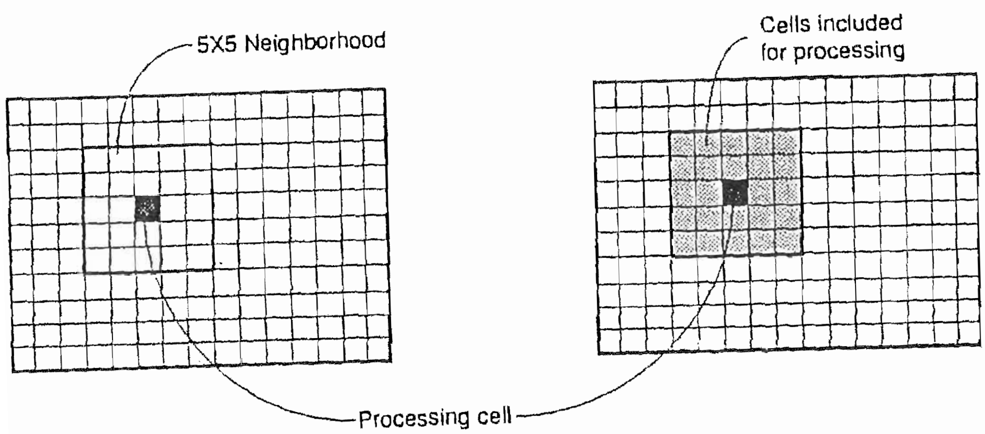
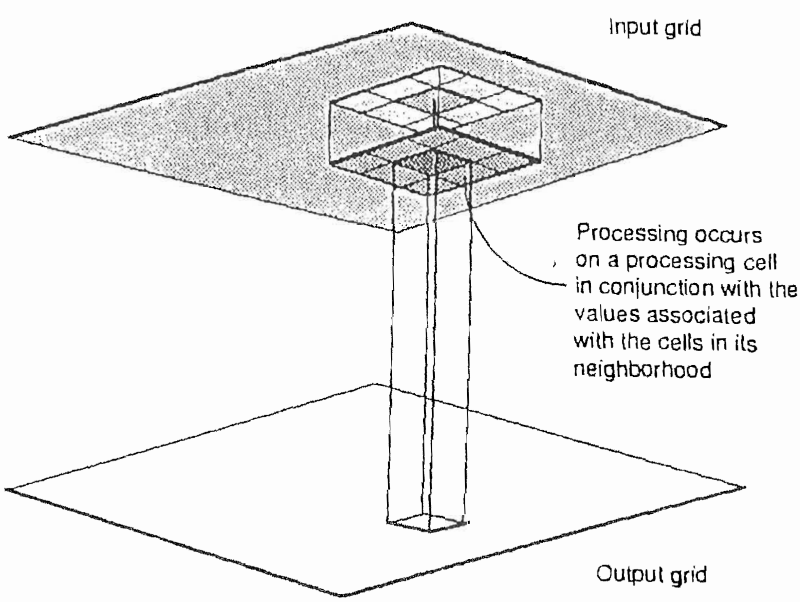


Figure 1.



In a good GIS many different focal functions are available. The focal mean has been mentioned above. Other focal functions that will be used in this study are as follows:

Focal sum - which is the sum of the values within the specified neighbourhood;

Focal maximum - which is the highest value within the specified neighbourhood;

Focal range - which is the difference between the focal maximum and the focal minimum;

Focal majority - which is the most frequent value within the neighbourhood; and

Focal variety - which is the number of unique classes within the neighbourhood.

There are other focal functions and these are reviewed in GIS manuals, and Tomlin (1993). Focal functions can be applied to both discrete and continuous data. However, some functions are better suited to certain data types, for instance focal majority and focal variety are more suited to discrete data.

When a landscape is assessed manually, the overall impression of an area is considered. Focal functions are particularly powerful for landscape classification because they can be used to capture the essence of the surrounding location of a particular point, and therefore capture some of the holistic (composition) qualities of landscapes. Duffield and Coppock (1975) used focal mean functions for identifying recreational landscapes, but since then it does not appear to have been used for landscape classification although it has considerable potential. The function has also been shown to be useful for automatic landform classification (Dikau, 1991) which will be discussed in chapter 5. It will be argued in section 3.4.2 that focal functions are now the most effective functions available in GIS for expressing spatial association for landscape classification.

3.3 National digital databases

The term digital database, for this study, refers to geographical databases that are in digital format and that can be incorporated in a GIS. Information stored in these databases is geographically referenced. The term national digital database (NDDB) refers to such a database that covers, or is intended to cover, the whole of a nation. They have been developed as a result of the development of GIS. The full utilization of these databases is yet to be realized. They are a recent technology whose full potential needs to be developed and experimented with. The construction of such databases can be a time consuming and expensive task and so it is preferable to utilize existing databases if they are appropriate.

Most Western countries have developed, or are in the process of developing, national digital databases of their environmental resource. In the United States, a digital topographic map of the whole country is covered at a scale of 1:100,000, and the United States Geological Survey aims to complete the digitising of a 1:24,000 scale map by the year 2000 (Southard, 1987). In Britain, a topographical database of the whole country is covered at a scale of 1:25,000, and the Ordnance Survey aims to complete digitizing of all the large scale maps (1:1,250 and 1:10,000) by about the year 2010 (Maguire, 1989). National databases are also being developed for the less developed countries (United Nations Environmental Monitoring Programme, 1990).

The development of digital databases is also being instigated at a global level (Clark et al., 1991). These are usually constructed by combining national digital databases. The Digital Chart of the World (DCW), developed by Environmental Systems Research Institute in the US, is an example of such a database. This gives digital information, which is stored on CD ROMs, of the whole globe, and provides a variety of information ranging from roads to political boundaries and waterways. The

Global Resource Information Database (GRID) being developed through the United Nations Environmental Monitoring Programme (1990) also has goals of developing a global digital database. Other well known global databases are the World Data Bank I and II files. These contain information on contours, river networks, and coastlines, which are all digitised from 1:1,000,000 maps (Maguire, 1989). This study will use mostly national databases, however, global databases can be more accessible than national databases and can therefore be useful for analysis at a national level. It is conceivable that a process, once developed, may be applied to the whole globe with the development of global databases and powerful GIS.

The databases suitable for landscape classification were mostly identified from database directories, such as The Department of Statistics (now Statistics New Zealand) (1992), and Newsome (1995). The following criteria were used for identifying the relevant databases:

1) The databases need to contain information on at least one of the four important landscape attributes at a national scale;

2) They need to have an appropriate level of spatial and attribute accuracy; and

3) They need to be accessible to the researcher.

Using these criteria, the databases described in Table 3.1 were identified and will be used in this study. Most NDDBs are derived from hard copy maps and the scale of these are specified in Table 3.1. Mostly DOSLI's 1:250,000 and Newsome's vegetation databases are used in this study, but where these were deficient for particular themes other databases were used. The advantages and disadvantages of these databases will be discussed when the different landscape attributes are classified in chapters 4 and 5. Many other databases are being developed by Regional Councils and DOC. However, some of these databases are not available or consistent at a

Digital Database	Base Map	Base Map Scale	Spatial Accuracy	Date of last revision	Format
DOSLI's topographical database	Infomap 262	1:250,000	150m circular radius for 90% of un-generalised points	1990	Vector
Vegetation cover database (Landcare)	Newsome (1987)	1:250,000	Locational precision 200m Min. map unit size 500 ha.	1981-1987	Vector
Land resource inventory (LRI) (Landcare)	LRI	1:63,360	Locational precision 35m Min. map unit 20-60 ha.	1984-1992	Vector
Supermap2 (Statistic New Zealand)	Census data	Not applicable	Variable size areal units	1991	Tabular
Ministry of Forestry's exotic and indigenous forest databases	NZFS's maps	1:250,000	Not specified	1992	Vector
Digital chart of the world (DCW) (Environmental Systems Research Institute, 1994)	ONC maps	1:1,000,000	Horizontal 2000-7100m Vertical 500- 2000m (for NZ)	1968-1991	Vector

Table 3.1 National digital databases used in this study

national scale, and some are not relevant. Unfortunately DOSLI's 1:50,000 topographic database was incomplete for the study area and would have also been financially inaccessible. Accessibility is a severe limiting factor that affects the use of DOSLI's topographical databases. The topographical data used in this study cost about $30,000 to purchase from DOSLI. Fortunately access was secured through Landcare Research LTD through a collaborative agreement. Without Landcare's support this research would have been severely limited.

It should be kept in mind that this thesis is investigating the potential for GIS to classify landscape. It is not intended that a current, usable classification is produced, and so no attempt will be made to identify and remove specific errors propagated from databases. If the classification produced in this thesis has substantial errors resulting from the databases, then with time this will be reduced as databases are upgraded. The real issue is whether GIS can incorporate the important compositional and generalised nature of landscapes. Despite this, the database errors will be discussed in section 6.4.3.1 to determine if there is a need for improvement.

3.3.1 Sources of national digital databases

There are two main sources of national digital databases, which are remote sensing, and the scanning and manual digitising of existing information. These two sources will be discussed separately. Global Positioning Systems (GPS) are another source of digital data. They are used for collecting spatial information usually in conjunction with either field work or scanners used for remote sensing.

3.3.1.1 Remote sensing

The term "remote sensing" refers to the observation of a target using a device located some distance away from it (Curran, 1985). This includes taking normal photographs, using aeroplanes to take stereoscopic photographs and scanning infra red images, and the use of satellites for scanning a wide variety of wavelengths. All these can be used as primary data sources for information on the landscape.

Of particular interest are the images obtained in digital format from scanners, as these can be analysed conveniently with computers using "image processing" techniques. Typically, for environmental sciences, these images are derived from scanners located on satellites, however, the use of scanners located underneath aeroplanes is becoming increasingly important.

The first unmanned satellite designed to provide systematic global coverage of the earth's resources was the Earth Resources Technology Satellite (ERTS-1, later named Landsat-1) (Aronoff, 1991). It was launched in 1972. Since then there has been an array of different satellites launched for remote sensing, ranging from geostationary satellites that are fixed above some point on the earth's surface and usually used for weather forecasting (eg. Meteosat-2), to sun-synchronous satellites that orbit the earth (eg. Landsat-5 and Spot-2). Continuous acquisition of digital scans of the earth's surface from these satellites has been prevented in practice by cloud cover and the lack of local ground receiving stations. The current generation of radar satellites will help to overcome the cloud problem. The resolutions of past images vary with the scanners from 10 x 10m for the panchromatic scanner on Spot-1, to 56 x 79m for the multispectral scanner on Landsat-5, and 1 km or more for the Advanced Very High Resolution Radiometer (AVHRR) on the National Oceanographic Atmospheric Administration (NOAA) satellite. From these images, it is possible, using image processing, to derive digital information on a range of environmental attributes, such as topography, vegetation, landuse, and influence of water. This information can then be incorporated within a GIS.

Scanners mounted on airborne platforms can provide even more detailed environmental information. The images are analysed in a similar way to satellite images using raster based image analysis software. However, for the same areal coverage as satellite images, this can be a more expensive option.

3.3.1.2 Scanning and manual digitising

A lot of environmental information has been gathered, either through field observation, topographic map interpretation, or photo interpretation. In this way, considerable information has been obtained on vegetation, soils, geology, landforms, fauna, landuse, archaeological sites, karst systems, and topography (Department of Statistics, 1992). An important means of conveying this information has been the map. With the development of sensitive office based scanners, many of these maps are being converted into digital form. Manually digitising these maps is also an option but this is tending to be less important as scanning technology improves. Scanning and digitising provides the spatial extent of different entities, however, it is also necessary to input attribute information that describes the different entities.

Many mapping agencies around the world are scanning the different layers of their maps so that they can be easily updated and republished. It appears that the main reason that these topographical databases are being developed is to aid cartography. It is perhaps a coincidence that these topographical databases can also be used for complex automated spatial analysis within GIS.

3.3.2 Classification of digital databases

Databases can be classified by many different data characteristics, for instance point or area, discrete or continuous, and integer or real. A useful classification could be based on the degree of input processing that they have had, and on whether they are specific or general purpose. Such a classification exists for data in general, and distinguishes between primary and secondary data (O'Brien, 1992).

Primary databases consist of crude data that has not yet been analysed, and does not necessarily present any meaningful information for a particular context. Digital databases that could be included in this category are remotely sensed images, such as from SPOT and LANDSAT, and also digital data obtained from field observations and GPS.

Secondary databases have already been processed to meet the needs of the collectors. Digital databases that fall into this category include digitised topographic maps, DCW, LRI, and Supermap 2. The agencies that supply these databases are in the information business and are therefore producing generalised databases that will suit a wide range of clients. They are usually derived from primary digital databases, or digitized or scanned from maps that were originally derived from field observations or remote sensing.

It would be useful if another category, here labelled tertiary digital databases, was distinguished to refer to digital databases that contain only relevant information for a specific issue. A tertiary database could be derived directly from processing a primary or secondary database, or digitised from maps. Landcare's digital vegetation map would be an example of such a tertiary digital database. A database could be secondary for some purposes and tertiary for others. For the landscape issue, being addresses in this study, Newsome's vegetation database would be regarded as a secondary database as further processing of this information is required.

This study is interested in developing a database that could also be categorized as tertiary. It is intended to do this by processing secondary databases. There is no point in deriving a landscape database from a primary digital database if it can be done more efficiently from a secondary database. However, there are disadvantages in using secondary databases because the processing used to derive them, which usually involves generalisation, is not often known, and therefore it is difficult to determine their quality.

It could be argued that it is better to derive a tertiary database by digitizing or scanning tertiary maps. Such a map, for a particular purpose, has to be available, and also, digitising can be expensive. Specific theme maps are usually not suitable for landscape classification purposes. For example, the landform map of Norway (Klemsdal and Sjulsen, 1988), is based on a genetic classification rather than a morphological classification. As will be discussed in chapter 5, it is the landform morphology that is important for a landscape classification.

3.4 Operational definitions

The automatic classification of landscapes is influenced by three factors: the human conceptual model, characteristics of the digital database, and GIS capabilities. As already mentioned, the balancing, or integration, of these factors has been labelled by Lay (1991) as "operational definition". Operational definition is not a foreign concept to geographical analysis (Mitchell, 1993), although Lay's interpretation is a slight variant because it is in regard to automation. Automation requires that the human concept model be formulated in a way that it can be "operationalized" with existing databases and GIS capabilities. With automation, the tradeoffs on the human concept model can be considerable, but this can be outweighed by the benefits of automation. Just because an automated approach may not represent a particular landscape precisely, is not a sufficient reason to discard the approach. The speed, explicitness, consistency, and repeatability of an automated representation may outweigh the disadvantages of misrepresentation. To classify landscapes automatically, it is necessary to understand the nature of landscapes, the available databases, and GIS functions. The former has been discussed in section 2.9, and the latter two have just been discussed in this chapter. The formulation of operational definitions now needs to be considered.

Kliskey and Kearsley's (1993) attempt to automate the mapping of wilderness also needed to address operational definition issues. They used a public perceptual survey to help determine more precisely the nature of wilderness so that definitions of this could be constructed. However, this does not appear to have been a useful method for deriving operational definitions, because public perceptual surveys still only provide general definitions. For example, some people identified remoteness as an important component of wilderness, but remoteness is ambiguous. What distance from huts, tracks, and roads, constitutes remote? For a definition to be precise, it really requires the use of numbers and mathematical relationships. Most people do not think in this way regarding landscape classes. Although Kliskey and Kearsley used a perceptual study of the term wilderness, when it came to implementing these definitions within a GIS, many arbitrary decisions regarding the mathematical interpretation of these definitions were required (for instance the extent of the buffer zones surrounding the tracks for identifying areas of different degrees of remoteness).

With automated classification, the issue that needs to be investigated is the transition from concepts (or geographical meaning) to operational definition. This is where the emphasis in this study will be, but obviously attention will be given to the meaning behind different concepts. A perceptual survey of the public's concept of different landscape attributes will not be conducted. The content category research discussed in section 2.9 provides some direction for a landscape classification. Definitions used by previous manual methods will also be used if appropriate, as well as definitions found in the literature, and if necessary personal judgement.

In the documented dialogue between Carlson (1977) and Ribe (1982) over the possibility of quantifying scenic beauty, Ribe (p.69) states that:

> "Numbers, when used for equations and statistics, provide a powerful means of rigorously describing, testing and analysing relationships in ways not possible through the use of only qualitative concepts and description".

With GIS, it is possible to extract from digital databases an almost unlimited number of different kinds of measurements on different aspects of the landscape. Not only can the quantity of different components be measured, for example length of road, and area of mountainous terrain, but this can be qualified in terms of different levels of scale, and can be combined with other measurements so that associations can be measured. This is a powerful advantage of GIS and digital databases, and it does not appear to have been utilized fully for landscape classification, especially with regard to identifying landcover. An important part of this research will be the identification of useful parameters that can be used for identifying different landscape classes. It can be argued that GIS can measure some parameters that are not practically possible to do manually, just because of the number of calculations involved. An example of such a parameter could be the density of roads within a given radius, calculated for 15 million points systematically located throughout New Zealand. This can be done within 10 minutes using modern computer hardware. To attempt to do this manually would not be practical. Such a measurement could be useful for constructing an operational definition of naturalness.

As discussed in section 2.9, landscape perception is a complex cognitive process that, among other things, involves generalisation, composition, and classification. Before an operational definition of landscape classes can be defined and implemented within GIS, it is necessary to know their exact nature in terms of mathematical relationships. Often landscapes are expressed in words rather than quantitatively. The challenge is to express the meaning of these words quantitatively. For example, how can a mountain be expressed mathematically. Nyerges (1991a) provides an interesting discussion on how geographical meaning (conceptualisation) can be represented, or formulated, in what he calls semantic data models. He argues that in order for computers to automate geographical models of reality it is necessary to include geographical meaning.

Four types of geographical abstractions are important in providing sufficient knowledge of meaning to perform structure identification. These are classification, generalization, aggregation, and association (Nyerges, 1991b). "A classification abstraction is created when one or more entities are assigned to an entity class" (p.1489). A generalisation abstraction is "created when a specific character of an entity class can be identified such that it is described as a subclass of the original class" (p.1490). Aggregation and association are both forms of geographical neighbourhood. "An aggregation is created when entities of the same or different entity classes form part of a more complex entity as a rigid structuring of parts" (p.1491). With aggregation there must be a substantive connection between entities. With association, entities are grouped as well, but this is based on looser relationships.

If the four above abstractions form the basis of structural geographic identification, how can these be represented within GIS? Nyerges (1991a) outlines a range of techniques for representing knowledge within a semantic data model. These are type hierarchy, functional dependency, domain role, definition, schema, attached procedure, and inference rule. The following summarizes these.

> Type hierarchy is the ordering of classes according to generality.
>
> Functional dependency indicates whether the entities are primary or secondary referents. Primary referents are independent, while secondary are functionally dependent on primary entities.
>
> Domain roles interpret the interaction of an entity in relation to another.
>
> Definition can be of three types: (1) classical - the use of conditions to show inclusion or exclusion; (2) prototype - the use of best examples to determine inclusion or exclusion; and (3) probabilistic - the use of statistical commonality to demonstrate inclusion or exclusion.
>
> Schema describes default (ie. normal) occurring roles that an entity type plays in relation to another type.
>
> Attached procedures are a set of external procedures that are used depending on a set of criteria.
>
> Inference rule represents reasoning based on explicitly stored knowledge of entity classes.

The formalisation of structural meaning into four abstraction types and then outlining representation techniques, is an important attempt to develop definitions for geographical meaning that can be implemented with computers. These building blocks can perhaps be implemented within a GIS. They now need to be tested in relation to particular entities, and an attempt to automate landscape classification will provide this test.

The four types of abstraction can be used to express the nature of landscapes. Landscape can be seen as an association of components. For example, a mountainous, forested landscape is an association of mountains and forest. Different components can be seen as an aggregation of sub-components. For example, a forest is an aggregation of large trees that may consist of a range of species. It has already been stated that landscapes are a generalisation, and the previous example is also an example of this. This demonstrates that abstractions are interrelated and complex. Describing landscapes using these abstractions raises questions as to the exact nature of the associations, and how components are aggregated, generalised, and classified. To answer these, it is necessary to express these abstraction types using representation techniques. This research will attempt to do this using representation techniques available within GIS. Nyerges' formulation of possible techniques provides a useful overview at a generalised level. To develop an automated approach requires the exact specification of GIS functions, such as overlays, conditional statements, and neighbourhood functions. The language used to express representation techniques in this thesis will therefore be at a GIS level rather than at the general level used by Nyerges.

One representation technique that Nyerges did not mention specifically was the use of fuzzy set theory, although this could be regarded as an inference rule. The foundations for fuzzy set theory were first laid by Zadeh (1965). Since then, it has been of growing research interest, especially with the development of GIS. Fuzzy set theory provides a strict mathematical framework in which imprecise conceptual phenomena can be studied. It can be thought of as a generalization of classical set theory, but instead of using the binary choice of two elements, weighted membership with more than two elements is used. This weighting of membership allows a continuum of possible choices that can be used to describe imprecise terms

(Zimmermann, 1992). For example, with landforms, there is not a clear distinction between mountains and hills. Some areas may be described as either a mountain or a hill. Such an area could be classified as 50 percent mountain and 50 percent hill, while areas that are clearly mountains or hills could be described as 100 percent mountain, or 100 percent hill, respectively. Landscapes are inherently fuzzy in nature because they are human constructs. Different people perceive landscapes differently and this needs to be incorporated in a classification. Fuzzy set theory provides a theoretical framework for expressing fuzziness. This now needs to be incorporated within operational definitions. How this can be done will be an aspect of this research, and is discussed further in section 5.4.2.

Of the four abstractions presented by Nyerges, classification and aggregation are easy to represent using GIS. Objects can be assigned to classes simply by selecting objects and naming them. Aggregation can be implemented by using overlay techniques. The representation of generalisation and association is more complex and will be discussed in detail in the following sections. Related to operational definition is the need to balance complexity with functionality and this will also be discussed separately.

3.4.1 Generalization

As discussed in chapter 2, landscape perception involves generalisation and it is necessary to incorporate this in a landscape classification. This generalisation is complex because it is an overall impression of an area obtained from exploration and movement. The question is: How can GIS incorporate this? This section discusses in more detail why generalisation is an issue, and also identifies techniques for resolving this issue.

Many existing databases have far more information than is needed in a landscape classification. The information in such databases cannot be perceived in reality from a reasonable distance. These databases may have been developed by researchers in specialized fields, such as botanist, soil scientists, and geomorphologists, with special purposes in mind, for instance to provide understanding on geomorphological process, protect species diversity, or determine the optimal crop. For deriving a landscape classification, it is not always optimal to import these databases directly without some form of generalisation. For example, Landcare's digital vegetation map has many more classes than can be normally perceived from a distance, as it was not developed for this purpose.

Generalisation is a contemporary problem resulting from developments in information technology. In the past, the degree of detail used in a model or classification has been limited by resources, especially finance (Jeffers, 1973). Classifications have contained as much information as can be obtained within the budget for the project. Converting the "firehose of data" that is available today to useful information is becoming ever more a generalisation problem. Techniques are now required to process this information and derive adequate generalisations. It appears that with GIS and national digital databases, it is the easy option to produce a detailed classification. The harder option is to produce a meaningful generalised classification.

With landscape classification, the composition of landscape components need to be considered. Since there are many different landscape components, there exists the potential for a very large number of possible compositions. These compositions need to be generalised to ensure the number of classes is at a useful level. It is difficult to know exactly what level of generalisation is appropriate, because of the different scales that the classification may be used for. It was concluded in section 2.8 that a hierarchical classification with a range of different level of generalisation is needed.

Nyerges (1991a) makes the distinction between cartographic generalisation, and the use of generalisation for geographical database abstraction. Cartographic generalisation commonly applies to selection, simplification, classification, induction, and symbolisation of maps. It is concerned with removing unwanted detail when a scale change takes place, and removing unwanted detail for thematic mapping (Armstrong, 1991). Newsome's (1987) vegetation map uses shading and symbols to express three levels of generalisation. When the map is viewed from different distances, different amounts of detail become apparent. Generalisation for database abstraction is concerned more with "a concept having a more general interpretation than some other concept with a more specific interpretation" (Nyerges, 1991a, p. 67). This is the way the term is used in the philosophy of science literature, and is the intended use in this thesis since the concern is with database abstraction. However, it appears that although cartographic generalisation may have different purposes to database abstraction, its generalisation can be similar. For instance, grouping trees into one symbol and calling the symbol a forest is an example of cartographic generalisation and generalisation for database abstraction.

Within GIS there is a range of different generalisation techniques available. Shea (1991) calls these "rule groups" and has provided a model that portrays them (refer to Figure 3.3). The model has been provided in relation to cartographic generalisation but may be useful for geographical abstraction. Conditional rules are the basic mechanisms for generalisation, of which there are five types:

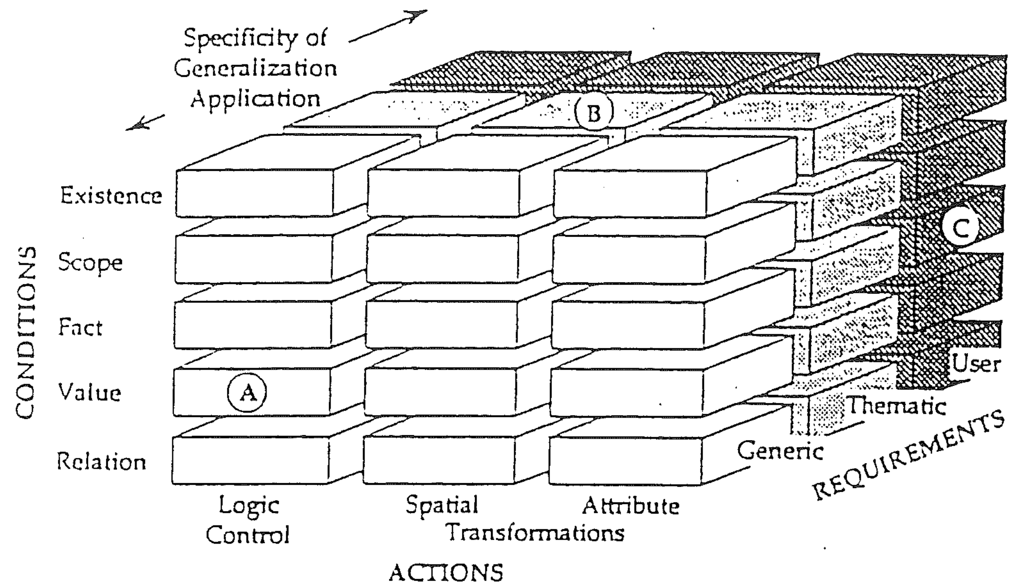
(1) existence, which test for the presence or absence;
(2) scope, which test for specific instances of some characteristic;
(3) fact, a test for truth or fallacy;
(4) value, which examine an entity's attribute values; and
(5) relation, which address cartographic and topographic relations.

These conditions can be applied within three types of actions:

(1) logic control, which directs the search and reasoning techniques;
(2) spatial transformations, which affect spatial data; and
(3) attribute transformation, which affect attribute data.

The relevancy of the logic control actions in this model is questionable as it specifies the type of generalisation rule that should be applied rather than being an actual generalisation rule.

Figure 3.3 Model of generalisation techniques
Source: Shea (1991)



The combination of conditional rules and actions can then be applied to various degrees of severity to suit requirements. In this case three levels are presented in Figure 3.3: generic, thematic, and user.

In designing a landscape classification process, different types of generalisations need to be considered. For example, should spatial information be generalised by deleting objects, or should attribute information be reclassified to more general classes. It also needs to be decided what type of conditional rule should be applied. Conditions can be complex involving many different objects and their values, or they can be simply based on the existence of one class. The importance of different types of rules will be demonstrated in chapters 4 and 5.

3.4.2 Association

Within GIS, there are many different methods for expressing association and it is necessary to determine which are the most appropriate. As previously mentioned, overlays can be used for expressing aggregation, but overlays could also be used to express association, since the distinction between these two abstractions is not that clear. Overlaying by itself is limited for expressing neighbourhood associations as it cannot identify whether two objects are within the vicinity of each other unless they occupy the same space. Other functions in conjunction with overlays have therefore been used for expressing wider neighbourhood associations. For landscape classification these have included buffer functions (Kliskey and Kearsley, 1993), nearest distance calculations (Lesslie et al., 1988), viewshed analysis (Bishop and Hulse, 1994), and focal neighbourhood functions (Duffield and Coppock, 1975). It will be argued that focal neighbourhood functions (described in section 3.2.1.1) are the most appropriate for this task.

A buffer function only indicates that a particular entity is present within a specified distance. It does not indicate how much of that entity is present, or how far away that entity is (except that it is within the buffer zone). Nearest distance calculations determine the distance to the nearest object in question but will not indicate the magnitude of the object. For example, if the spatial influence of roads from a particular point need to be determined, and from that point there is one road 10 km to the south and another road 11 km to the north. The use of nearest distance will give a value of 10, whether the road to the north existed or not. If a focal mean function was used then the output will be affected by both roads and is therefore more sensitive. However, if a road also went through the central point, then the nearest distance will be zero. The focal mean would be affected by this and also the roads in the distance. This may or may not be appropriate since the roads in the distance may be considered too far away. If it is desirable that roads too far in the distance not be included, then this can be achieved by limiting the neighbourhood search radius.

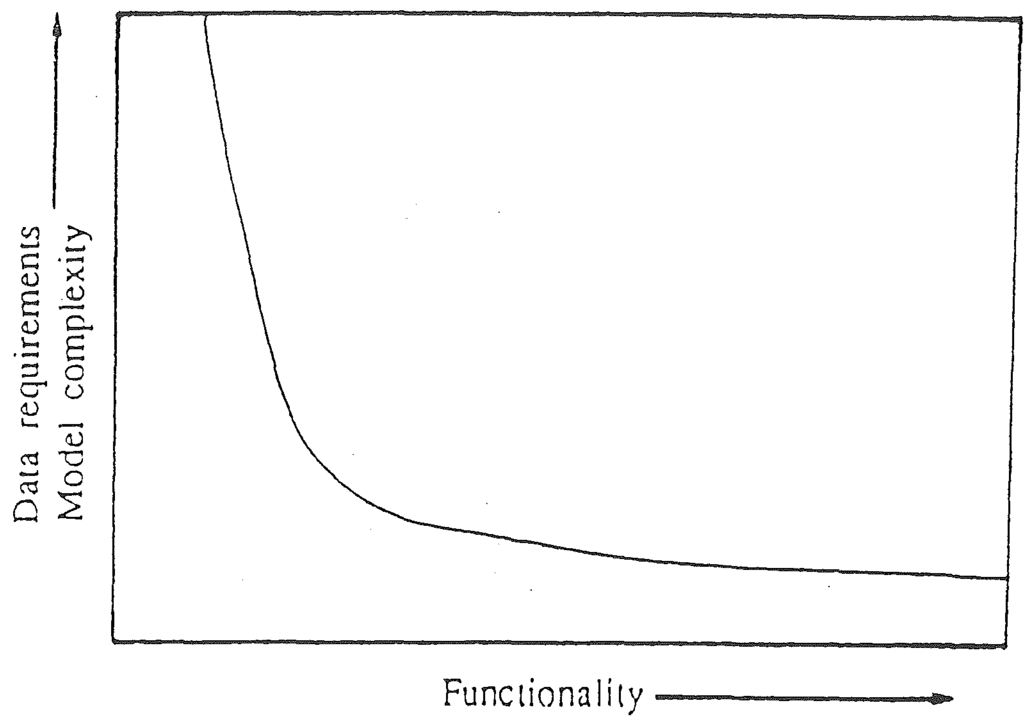
Although viewshed analysis is becoming a standard function within GIS, focal mean functions may be more appropriate for determining the spatial influence of different objects. This is because landscape perception is not just affected by what is directly visible, but also by what has been experienced through movement and exploration. This point has been made in section 2.9, and is based on the work of Zube, Sell, and Taylor (1982) who reviewed 160 landscape articles from 20 different journals. The need to incorporate movement and exploration has therefore been stated as a criterion for landscape classification. Focal mean functions can express the spatial influence of objects within the vicinity of a particular point regardless of whether or not it is in direct line of sight. Focal neighbourhood functions will therefore be the main GIS function used to express spatial association and their effect is demonstrated in chapters 4 and 5.

3.4.3 Complexity versus functionality

Operational definition requires that the process can be run adequately within the confines of GIS, and also provide adequate representation. The process needs to be complex enough to give useful results, but also needs to be functional within GIS. Moore et al. (1993) provide a graph showing the tradeoffs between complexity and functionality regarding mathematical modelling (refer to Figure 3.4). It shows that if a model is too complex, requiring substantial amount of data and processing, then the model will not function very well. This is because the demands on computation will be too great. It is difficult to know exactly what the dimensions of this negative relationship are. It is shown as an exponential curve, but it may not be. The figure is just an abstract illustration that is useful for discussing this important tradeoff.

Moore et al.'s figure, however, only regards functionality in terms of "ease of use" (p.198). Functionality should also consider how well the model depicts reality, which is also a function of complexity. If a landscape model is too simple and does not reveal the important subtleties that are present in reality, then the model is not functioning very well. It should be noted that complexity and degree of generalisation are not the same thing. A model can be quite general, with only a few broad classes, but the process for deriving this generalisation may be very complex involving large detailed databases and sophisticated calculations.

Figure 3.4 Model complexity versus functionality
Source: Moore et al. (1993)



To develop a functioning model, it is therefore necessary to choose an appropriate level of complexity that is computationally feasible, and that can identify important subtleties. It appears from past research that manual methods have been unable to balance these two criteria to produce a functioning model. This is either because the manual models were too computationally demanding, thus requiring considerable resources, or the resulting classification was not complex enough to be of any use. The question that will be addressed in this thesis is: Can GIS function acceptably at the required levels of complexity?

Moore et al. (1993) suggested two important principles that a model should follow - parsimony and modesty. A model should be parsimonious in that it should not be more complex than it needs to be, and should include only the smallest possible number of parameters. A model should be modest by not pretending to do too much.

3.5 Investigation method

It has already been said that this research is about investigating the application of GIS to landscape classification. The discussion so far has provided a general theoretical framework, and identified major issues, such as operational definition and generalisation. It is now necessary to put this theory into practice and actually apply GIS to this problem. This section outlines the method that will be used for doing this, as well as the study area and a discussion on validation issues.

It has been argued that landscape is composed of landform, vegetation, naturalness, and water. To simplify landscape classification, these attributes will be classified separately, and then the landscape classes can be constructed from the unique combinations of these four layers. When classifying the separate attributes, it will be necessary to consider that some of these attributes, for instance vegetation and naturalness, are interlinked.

The main method used for developing an automated landscape classification can be regarded as a kind of simulation. Simulation can be defined as the representation of the characteristics of one system through the use of another system, such as computers. The system being represented is manual landscape classification. The characteristics deemed important have been incorporated in the criteria listed in chapter 2. Simulation is a powerful tool within GIS that have macro language capabilities. A process, once developed, can be easily altered by simply changing parts of the program. The sensitivity of different parameters can be investigated by using a range of parameter settings and comparing the resulting outcomes either visually or quantitatively. Parameter settings can be changed using variables within a "Do Loop". In this way many different outcomes can be produced with relative ease. The GIS used for this investigation was ARC/INFO 6.1.2. and the hardware was a SUNSPARC 10 workstation.

Display of outcomes can be a problem because of the quantity produced. To facilitate comparison between maps, information on the hard copy outputs will be kept to a minimum. For instance, most of the maps produced will not contain north arrows and scale keys. For all maps the north direction is up. The scales vary, but Figure 3.5 has a scale bar, and from this the approximate scale of the other maps can be ascertained. For all maps that are raster based, the cell size will be given. Many maps will also display the main roads and hydrology layer for geographical reference purposes.

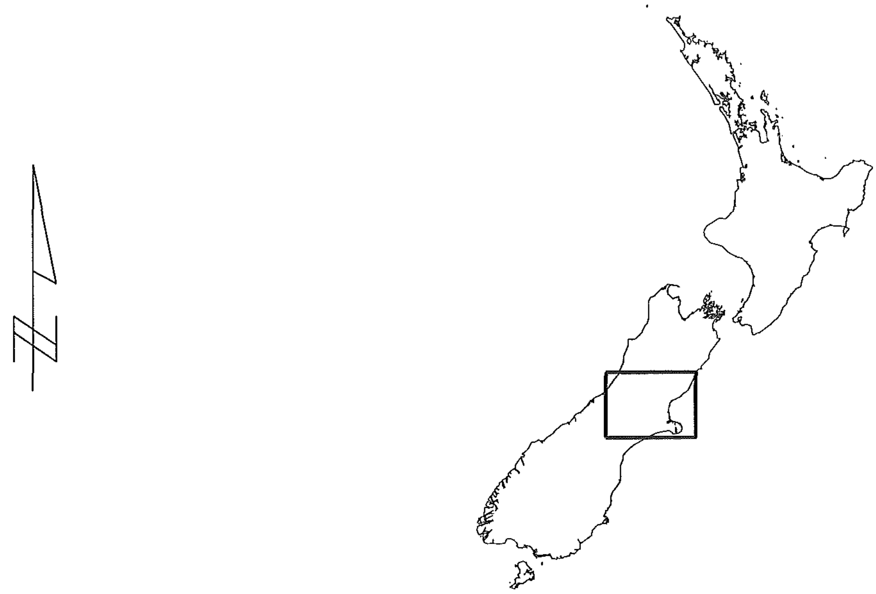
The classification process developed in this study will convert vector databases to raster databases to aid spatial analysis. With raster databases it is necessary to decide on an appropriate cell size. The effects of cell size are complex and will be a major part of this study. Consideration will be given to the processing speed, the spatial resolution of the NDDBs, and the objects that are being identified. A cell size of 500m will initially be used, but the effects of smaller and larger cell sizes will be investigated.

3.5.1 Study area

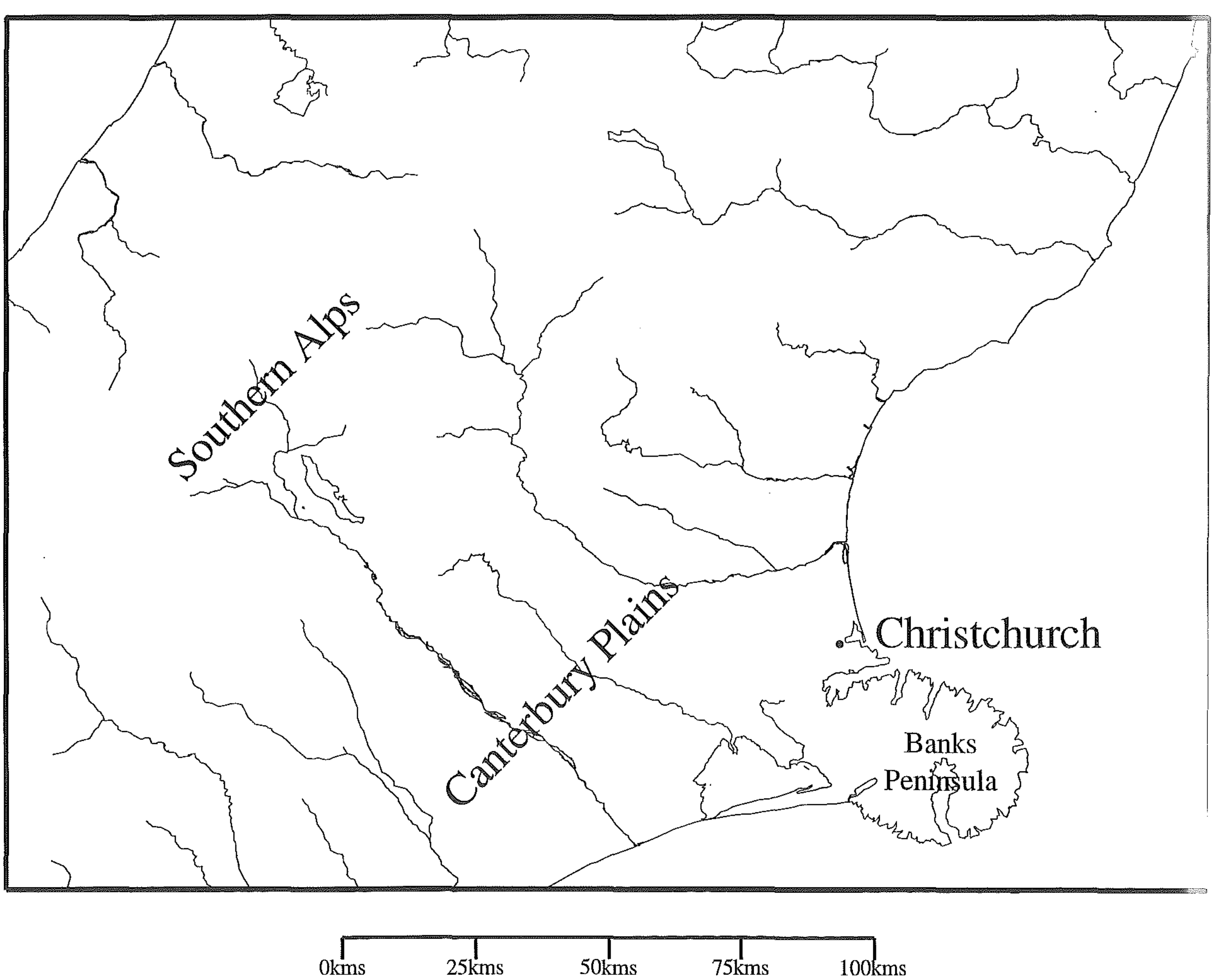
Four factors were considered in choosing a study area. The first of these was that the area should have a suitable range of landscapes so that the generality of the landscape process can be tested. This requires that the study area vary significantly in landform, vegetation, naturalness, and the influence of water. It also helps if the study area is well known to the researcher so that the outcome can be easily compared with reality. If this was not the case then each output would have to be systematically compared with representations of reality such as hardcopy topographic maps and photographs. It may also be necessary to conduct field visits. When the study area is known, then the outputs can be more quickly assessed, and spurious output spotted. Another consideration for determining a study area was that the necessary digital data can be obtained. It is also beneficial if the area has already been classified manually.

The area chosen is a cross section of the middle of the South Island of New Zealand (Figure 3.5). It consists in total of approximately 3.7 million hectares, however, only 2.8 million hectares of this are land. When divided into pixels of 500m cell size, a matrix of 328 (rows) by 453 (columns) cells result. The study area consists of a large

Figure 3.5 Study Area



Hydrology layer is shown



variety of landscapes. On the east coast there is the extensive Canterbury Plains, dividing the east and the west are the Southern Alps with mountains up to 2500m high, and on the west coast there is a relatively narrow strip of flat and hilly landforms. Banks Peninsula on the east coast is an extinct volcano with hilly to mountainous topography. The vegetation over the study area also varies. There is expansive pasture on Canterbury Plains and the adjacent foot hills, a mix of forest and tussock in the Southern Alps, and a mix of forest, scrub, and pasture on the west coast. A range of human modification also exists. Christchurch is an industrialized urban area of 300,000 people, while parts of the Southern Alps and west coast are relative wilderness. Several large rivers and lakes are present, as well as a range of coastlines. An assessment of the landscape in this area can be obtained from DOSLI's NZMS 262 (1:250,000) topographical maps - sheet numbers 10-13, and from Landcare's vegetation map. This study area is well known to the author who has travelled extensively throughout this region, both through work and a passion for exploration and outdoor recreation. Most of the important databases for this area were also obtainable after some negotiation. A disadvantage with the study area is that the landscape has not previously been manually classified using the attributes landform, vegetation, naturalness, and water. There is only one area in New Zealand that has been classified using these attributes, and that is the Auckland region (ARA, 1984). The Auckland region would not have been appropriate for this study because of the lack of landscape contrast there, and because it is unfamiliar to the author. The landscape in the study area has, however, been classified using different attributes. The most notable of these is the classification developed for the survey of natural resources (Ministry of Works and Development, 1983). This used formal artistic criteria and tended to evaluate rather than classify character. The Canterbury Regional Council recently commissioned a landscape study, but the classification resulting from this was an inventory of physical characteristics - similar to the Land Resource Inventory. It includes information, such as soil and genetic geomorphology, that is not directly relevant to a perceptual landscape classification.

3.5.2 Validity

To develop a classification process, there must be some way of assessing the worth of the output. Without some form of assessment of validity one cannot say whether a classification is useful or know whether it needs improving. Validating a computer generated landscape classification is particularly difficult because of the complex nature of landscapes. It is not possible to develop a landscape classification and then compare this with the real world because landscape classes are human constructs that only exist in the mind. The components of landscapes can be assessed in the field, but landscapes are a generalisation and composition of these components. Validation of landscape classifications has not been seriously discussed in the literature, and is in itself a theoretical issue.

In science, it appears that classifications are validated by further research. Classifications can be seen as representations of knowledge. As knowledge of a particular field increases with research, the validity of the existing classification can be assessed. It is in this way that classifications have evolved. By using a landscape classification as a frame of reference in applied or theoretical research, the usefulness of that classification will become apparent. If inconsistencies result between different areas for the same class, then the classification has perhaps not captured the essence of the landscape character. For example, if a public preference survey shows that the quality of a class in one area is high, and in another area the same class is low, then the important characteristics of landscape may not have been totally incorporated in the classification. Unfortunately, in this study it is not possible to validate a landscape classification in this way because of the time and resources required to do further research. It will therefore not be possible to say whether the resulting classification is valid in this sense. The classification will, however, be assessed using the criteria put forward in sections 2.8 and 2.9.

Two approaches for validating a process are: (1) to examine the outputs by comparing with a desired output, or (2) to examine the process itself (which includes input). If the process appears sound then the output can be assumed to be valid. With landscape classification it is difficult to validate an output because there is no correct output with which to compare it. Validation therefore needs to be predominantly process based.

There is some documentation on manual methods for classifying landscapes and this has been discussed in section 2.7. A comparison will be made between these manual approaches and a GIS approach, based on the general and specific criteria, to determine whether there has been an improvement. It is only possible to compare classification processes since it is not possible to physically compare outcomes. Even comparing processes is difficult because manual methods are often intuitive.

Errors are another issue that will need to be considered for assessing validity. Errors include database errors, computational errors, and logical errors. Because GIS is particularly powerful with spatial information, errors can be easily propagated (Goodchild, 1993). It is therefore necessary to confirm that this has not been the case to ensure the classification is valid. Error will be discussed in section 6.4.3, which addresses the validity of the process developed in this thesis.

3.6 Past research

The use of computers for landscape assessment has been mostly limited to programmes that give perspective views (such as VIEWIT), photomontage, and also overlays of grids and polygons (Brown, 1981). Past research in automated landscape classification is extremely limited. There does not appear to be any automated process developed for classifying landscapes as a whole, although, there has been some work on the use of GIS for identifying some attributes of the landscape. Barbanente et al. (1992) used GIS and digital databases to automatically identify three landscapes: cliff, ravine, and system of farms in regular grid. Lesslie et al. (1988) used GIS to map wilderness areas in Tasmania, and Kliskey and Kearsley (1993) did a similar study in New Zealand. As mentioned previously, Duffield and Coppock (1975) used a primitive GIS that had focal neighbourhood function capabilities for delineating recreational landscapes. These researchers have worked independently of the landscape theory previously discussed. They have only examined one part of the landscape, and have therefore not addressed what are the important attributes needed for a total landscape classification. The components that they have identified are not necessarily relevant for a landscape classification. Many of them cannot be easily incorporated into a total landscape classification because they are too detailed.

Jackson (1990) discussed the application of GIS for identifying landscape features in New Zealand using digital terrain models (DTM). This is one of the earliest published applications of this type of research in New Zealand. Jackson demonstrated how information on slope, aspects, contours, and views could be obtained. This was at a time when some of these functions were quite new for commercial GIS. Now, many more functions have been made available. There has also been considerable progress in the development of national digital databases, although some problems that Jackson mentioned regarding availability are still pertinent today. It is now possible to implement some of Jackson's ideas on a larger scale and identify more features of the landscape.

There are many other landscape character assessment studies that have used GIS, including Brooke (1994), and Canterbury Regional Council (1993). With these studies, GIS has been used mainly as a presentation tool, and the analysis has been done with non-GIS means. These studies are not very useful for this research because it is the automation of analysis within GIS that will be investigated. Although their criteria are not explicitly stated, they appear to be based on different criteria to those established in this study.

There has been some work by geomorphologists and hydrologists using GIS to automatically extract terrain information from digital databases, which is of considerable relevance to this research. Lay (1991), Cowen (1993), Dikau (1989), Tang (1992) and Weibel and DeLotto (1988) have all discussed different aspects of this type of research. These works are of interest because terrain information is important for characterising the landform attribute of the landscape classification, and because some of the techniques can be applied in this study.

Past research will be reviewed in detail in chapters 4 and 5 in specific reference to the different landscape attributes.

3.7 Summary

There is an increasing range of NDDB that contain information useful for landscape classification. Most of these databases contain low level conceptual information and can be classed as primary or secondary NDDB. It is important for decision makers that tertiary NDDB, which address a single theme such as landscape, are derived from them. A landscape classification requires high level conceptual information. Operational definitions based on four abstractions - classification, generalisation, association, and aggregation, provide a framework for deriving this information from the low level information available in NDDB. A difficult challenge in using GIS to classify landscapes automatically, is expressing association and generalisation. Focal neighbourhood functions can be used to express association and it has been argued that they are more appropriate than buffer functions, nearest neighbourhood functions, and visibility functions. Generalisation can be achieved using attribute, and spatial information along with a range of different conditional rules. The role of these procedures will be demonstrated in the following chapters, which deal systematically with each of the four attributes (vegetation, naturalness, water, and landform) to be used in the landscape classification.

CHAPTER 4

AUTOMATED CLASSIFICATION OF LANDCOVER

4.1 Introduction

This chapter outlines the development of a process for automatically classifying landcover for the use in a landscape classification. Landcover for the purpose of this study consists of vegetation, naturalness and water. The classification of these three attributes is addressed separately. Although they are related, for instance exotic vegetation affects the naturalness of an area, this separation is necessary to simplify the task. Most of the components that contribute to landcover are already conceptualised in the digital databases available in New Zealand, while this is not so with landform components. For this reason the classification of landcover is relatively simple compared to the classification of landform and is therefore presented first. All figures and tables are placed at the end of the chapter.

4.2 Vegetation

4.2.1 Past research

Virtually all manual landscape classifications, which have used the physical landscape components, have used vegetation as an attribute. The vegetation classes have been based on major differences in vegetation form. Classes such as grassland, scrub, and forest, and classes that are compositions of form have been commonly used (eg.

Linton, 1970, and Auckland Regional Authority, 1984). These classes are similar to those used in Raunkiaer's life form classification (cit. in Tansley, 1946). The plant taxonomy used by Botanists and Ecologists has not been used because it is based on plant evolution rather than outward appearance.

The use of GIS for incorporating vegetation information within landscape studies is becoming increasingly common. Lesslie et al. (1988) and Kliskey and Kearsley (1993) both incorporated vegetation within their wilderness identification processes, but only used the distinction between exotic and natural vegetation. Bird et al. (1994) used GIS to monitor landscape change and included many different vegetation classes. The classes were manually derived from aerial photos and GIS was used for analysing change. In New Zealand there has not been any attempt to derive suitable vegetation classes for a landscape classification automatically from existing databases.

The automatic classification of vegetation using remote sensing has been widely researched (Leckie, 1990), and the results appear promising for use within landscape studies. DOSLI has completed a pilot project that successfully mapped broad vegetation classes using Landsat images (Dept. of Survey and Land Information, 1994). The use of remote sensing techniques will not be investigated in this study because remote sensing is concerned with creating NDDB, while this study is concerned with using NDDB.

4.2.2 Suitable databases

Landcare has produced a digital vegetation map of New Zealand (Newsome, 1995), and vegetation information is also included in their LRI. The Ministry of Forestry has produced a coverage of indigenous and exotic forests, and DOSLI, as mentioned previously, has experimented with the use of Landsat images to produce a landcover map of the central North Island. Landcare's vegetation database is currently the most suitable to be used in a landscape classification. It has nationwide coverage with 49 different vegetation classes. It was derived from the LRI and from field work, but is slightly dated since it was based on field work from 1981-1987. Newsome (1995) notes that the accuracy of this database is acceptable at the scale of mapping, which was 1:1,000,000, but cautions that the exotic forests and pasture-scrubland classes have changed since it was published. In contrast, DOSLI's landcover database, which is being derived from Landsat images, is using a base map at a scale of 1:50,000. It, however, only has 20 different classes and has only been completed for the central North Island. The Ministry of Forestry data sets are nationwide, were developed from base maps at a scale of 1:250,000, but only records the presence of two classes, exotic forest and indigenous forests.

For this study Landcare's vegetation database (Newsome, 1987) was updated using the Ministry of Forestry's exotic forest database. If Landcare's database was not recording exotic forest in a certain area and the Ministry of Forestry's was, then Landcare's database was changed to exotic forest, otherwise it did not change. These two databases were used because they were available for the study area, and also because they identify classes that are necessary for a landscape classification. The age of Landcare's database was not considered a serious drawback in this study, where the primary concern is the classification process. Once a process has been developed, it can be easily applied to current databases when they become available.

4.2.3 Classification process

Landcare's vegetation database contains 47 classes, which are listed in Figure 4.1. Newsome (1987) describes precisely what these classes are. This database is provided in vector format and one of the first tasks was to convert it to a raster format with a cell size of 500m. Very little spatial accuracy is lost during this conversion because the minimum size polygon of the vector coverage is 500 ha which is considerably larger than the cell size used. To preserve the vegetation class information in the grid coverage, this attribute had to be represented by integers before being converted.

Some of Landcare's classes are too detailed to be used in a landscape classification and so were generalised. It is doubtful whether the general public perceive the difference between a lowland podocarp-broadleaved forest and a highland podocarp-broadleaved forest from a distance. To most people these would be just indigenous forests. If they could, it is still doubtful whether this distinction would be significant in determining landscape quality. Twelve groups were created from Landcare's original classes. These are listed below, along with Landcare's classes that constitute each group.

1 Horticulture	(C1, C2)
2 Pasture	(G1, G2, GS1, GS2, GS3, GS6, GF1, GF2, GF3, GF4)
3 Tussock grassland	(G3, G4, G5, G6, GS4, GS5, GS7, GS8, GF5, GF6)
4 Lowland indigenous scrub	(GS1, GS2, GS3, S1, S2, FS1, FS2, FS3, FS4, FS5, FS6, FS8, M4)
5 Exotic scrub	(GS6, S4, FS8)
6 Alpine scrub	(GS4, GS5, GS7, GS8, S3, FS7, M4)
7 Indigenous forest	(GF1, GF2, GF3, GF5, GF6, FS1, FS2, FS3, FS4, FS5, FS6, FS7, F1, F2, F3, F4, F5, F6, F7, F8)
8 Exotic forest	(GF4, FS8, F9)
9 Alpine herbfields, rock, and ice	(M1)

10 Wetland (M2)

11 Sanddune (M3)

12 Vegetation not significant (Urban areas, lakes, and large rivers)

This level of generalisation was selected to ensure that important vegetation groups were included. These groups form the basis of the twelve components of the vegetation attribute; they distinguish major changes in the form or colour of the vegetation, and whether it is native to New Zealand. The groups are based predominantly on the author's knowledge of New Zealand's vegetation, along with information from Newsome (1987). Newsome (1987) also groups the classes, as shown in Figure 4.1. However, these groups were not used because they do not distinguish between exotic and native vegetation, nor between tussock and lowland pasture. If the original vegetation database had contained information on the form, colour, and naturalness of each class then this generalisation could have been implemented automatically. It will become apparent that when different compositions of these groups are considered and these are combined with other landscape attributes, the classification becomes quite detailed.

The list above shows that some of Landcare's classes have been included in more than one group. For example, pasture exotic forest (GF4) is included in group 2 (pasture) and group 8 (exotic forest). It is necessary to do this to establish the presence or absence of each group, and it does not matter whether these groups spatially overlap. A separate grid coverage was made for each group, and these are illustrated in Figure 4.2 and Figure 4.3.

Altitude information was used to allocate occurrences of some Landcare classes to the groupings. M4 (Pakihi heathland communities) can exist in a (sub) alpine environment and a lowland environment (Newsome, 1987). To know what group to assign different areas of this class, it was necessary to use altitude information, which can be easily implemented with GIS. A threshold of 500m was used to assign this class to either lowland scrub or (sub) alpine scrub. Landcare's databases also contain a class that consists of ice, snow, scree, and sand. Sand dunes were distinguished from this class using an altitude threshold of 200m.

Once the 12 single theme vegetation component grids had been derived, vegetation compositions were then determined. This required a series of steps. Considering each vegetation component grid separately, the value 100 was assigned to the cells where the component was present, and the value zero to the cells where it was absent. A focal mean function with a neighbourhood analysis window (NAW) radius of 3000m was passed over each grid. The resulting mean values give the percentage of the NAW that contains the vegetation component. This in effect describes the spatial influence of each vegetation component for each cell. The rationale for a 3000m NAW will be discussed in section 4.2.4 along with the effects of other NAW radii. The results were classified (based on critical thresholds) into four levels of spatial influence and are shown in Figure 4.4 and Figure 4.5. The threshold levels 0%, 20%, and 50% were used since 0% indicates presence/absence, 20% seems an appropriate minimum presence (this is discussed in section 4.2.5), and 50% indicates a majority. It was not necessary to determine the spatial influence of the 12th class because urban areas and water are classified later.

These 11 spatial influence grids were then overlaid to produce a grid that contained the unique combinations of these grids. The last map in Figure 4.5 shows a vector representation of this combined grid. There were 360 unique combinations identified in the study area. If all the possible combinations had been present in the study area this would have totalled 4,194,304. If the spatial influence grids had been classified into more than four levels, then the process would have the potential to identify even more combinations. For example, if each coverage had contained five levels then there is the potential for 48,828,125 unique combinations to be identified. It was necessary to keep the number of levels to around four since the software appears to have a limit of about 10,000,000 potentially unique combinations.

This combined grid still contains information on the spatial influence of the 11 vegetation components. Therefore, vegetation compositions can be identified by querying this. For example, an indigenous forest-tussock composition is common in New Zealand. This class can be identified by doing a query for areas where there is a spatial influence of both indigenous forest and tussock. Forty six different vegetation classes (not all are compositions) were identified in this way. These are listed in Table 4.1 under level 1. The definitions used for identifying each class are given in Appendix 1, and were implemented using ARCPLOT. These definitions generalise the large number of possible compositions to a manageable size by using a relatively complex set of rules based on many different attribute values. Not all these vegetation classes existed in the study area. The resulting vegetation classification is shown in Figure 4.6.

Many different classes could have been identified using the spatial influence information. The classes identified in Table 4.1 under level 1 have been chosen because they reflect major differences in appearance (form and colour), naturalness, and contentiousness. Some classes, such as wetlands, are contentious in landuse planning. What is important about this process is not the actual classes identified but the fact that it demonstrates that vegetation compositions (associations) can be expressed. As our understanding of landscapes improves and a substantive rationale develops for the importance of different classes then the above classes can be revised using other explicit definitions.

By using ARCPLOT to select compositions based on a set of definitions, it is possible to list the compositions that have not been accounted for. Depending on the remaining compositions, the definitions were either altered so that the compositions were included or a definition for a new class was developed if this was considered appropriate. It was also possible to check that the definitions were mutually exclusive by counting the number of compositions selected for each class. If the total number selected was greater than the number of compositions available, then some compositions were selected twice and therefore the definitions overlapped and alterations were needed. A check was also made to ensure all areas were selected.

In theory it is not necessary to combine the grids, and then query the attribute table of the combined grid. It is possible to use a series of conditional statements that query each individual component grid, however, this would be a slower process, and it would have been difficult to check that definitions were mutually exclusive. With ARC/INFO the quickest method is to use ARCPLOT to query the attribute table of the combined grid.

This vegetation classification can be further generalised by grouping different classes. This was done to produce six different levels of generalisation. The way the different classes were grouped is shown in Table 4.1. Figure 4.7 shows graphically the effect of different levels of generalisation. No keys are provided with this figure to avoid cramming, but the colours are the same as used in Figure 4.6 and the keys can be ascertained by using this and Table 4.1. Such generalisation is important for reasons discussed in section 2.9. This will become even more apparent when the different landscape attributes are combined to produce a landscape classification.

Why were six levels of generalisation developed and what rationale is there for the different classes within each level? The fact that there are six levels is not particularly important. What is important is that different levels of generalisation can be easily expressed. Perhaps it will become apparent if the classification is used in landscape research which of the levels are important. The classes used for levels 2-6 were chosen for similar reasons as the level 1 classes. They reflect important differences in appearance, naturalness, and contentiousness, but as the levels become more general these reasons need to be more apparent.

The generalisation process used to identify the six levels uses relatively simple conditional rules based on the existence of one attribute. This is kept simple so that a hierarchical structure is produced whereby the relationships between generalisation levels can be easily interpreted. Information on classes at the general levels can be applied to classes that are at more detailed levels because the classes feed into each other - many to one going from detailed to general. Complex conditional rules based on many different attribute values would not have been appropriate because the links between levels would have also been complex - many to many.

4.2.4 The neighbourhood analysis window (NAW)

The classification process described above used a 3000m radius NAW to determine the spatial influence of different vegetation components. This radius was selected after a careful investigation of the effects of a range of radii from 1000m to 5000m. Figure 4.8 compares the effect of three different search radii - 1000m, 3000m, and 5000m (the key is the same as for Figure 4.6), The amount of agreement (percentage of area with the same class) between 1000m and 5000m search radii is low - 61%. When the search radius is small, less compositions and many small discrete areas are identified. With a search radius of 5000m there is a lot more generalisation, a few large discrete areas are identified, and many of these are classified as compositions. It is difficult to know which search radius is more appropriate. If a large search radius is used then areas that are far away are being used to classify the focal area. It is not appropriate if this is too far.

A 3000m neighbourhood search radius is large enough to go beyond small hills, but it is not too large to require considerable amounts of processing when the resolution of the raster coverage is 500m. The search radius should be related to how people view and experience landscapes. However, sufficient cognitive research on this is not available. People can often see for more than 3000m but how much detail is perceived beyond this distance? Does the foreground of a view have more impact than the background? If so by how much? To address such questions, it would be useful to know more about how people experience landscape. It is probably highly variable, and is not only dependent on the person but on the situation. Discussion and empirical research are required. For the time being, different landscape classifications can be created using different search radii, and the variability in the results can be presented. Variability can also be represented using fuzzy set theory, and an application of this will be presented in section 5.4.2 with regard to landforms.

It should be noted that the distance of the search radius is measured using horizontal distance, and does not incorporate ground distance, which also has a vertical component. This results because a two dimensional grid is used to represent a three dimensional surface. The effect of this is that the neighbourhood extent is more in hilly and mountainous areas in terms of ground distance. This may not be appropriate because in such areas topography can reduce the amount of movement or exploration. However, one mountain top may be easily viewed from another mountain top even though the ground distance between the two may be considerable because of a deep valley in between.

Annuluses could also be used for determining the shape of the NAW. The annulus shape comprises of one smaller circle within a larger circle (donut shape). Cells that fall outside the radius of the smaller circle but inside the radius of the larger circle will be included in the processing of the neighbourhood. An annulus would enable the spatial influence of components to be specified for a range of distances. It would be possible to quantify the spatial influence of components for different degrees of proximity - close, medium distance, and far away. This information could then be used for developing complex definitions for different landscape classes. Not only can annuluses be used but also wedges can be specified that control the aspect of the NAW, eg. 0-90 degrees. This would provide even more opportunity for specifying the exact nature of the spatial influence of different components. The way different landscape components are composed could then be quantified. The NAW can also be weighted by using a kernel. This kernel could enable an appropriate distance decay functions to be specified for each landscape component, which could then be incorporated in the spatial influence calculations. The problem with using these GIS features, which are available with ARC/INFO, is that it is not known which complex compositions are important, or which distance decay functions should be used for the different components. Therefore, these features will not be used in this study.

4.2.5 What is a significant amount of spatial influence?

When determining definitions for vegetation compositions, it was necessary to specify what amount of spatial influence of a particular component is significant. For example, if 1% percent of a neighbourhood is grass and 99% is forest, should this be called "forest", "grassland-forest", or "forest with a small amount of grass"? In this circumstance it was considered that "forest" was appropriate, because the influence of grass was not significant enough. For most of the class definitions a 20% threshold was used for the important vegetation components. It is difficult to know whether this is appropriate as it depends on how people perceive landscapes, and there is no substantive research on this. Other thresholds as well as 20% were experimented with and Figure 4.9 shows the results. The difference in the outcome is significant. The amount of agreement between the 10 percent threshold and 30 percent threshold is only 63%. When a low threshold is used there is a high mix of vegetation components, and not many "pure classes" are identified, while with a high threshold there are more "pure classes".

Some components dominate over other components. For example, all things being equal, forest dominates landscapes more than grass. Therefore, the thresholds used in the definitions are related to the components being used. It will also be noticed from the definitions in Appendix 1 that the "or" statement is used. This is so that a range of combinations can be considered.

4.2.6 Sensitivity to cell size

The cell size greatly affects the processing speed. If the cell size is halved then the number of cells in the grid coverage increases by four, and therefore any operation that processes each cell will take much longer. However, there is more to it than that. When a neighbourhood function is used, the NAW of each cell needs to be analysed. If the NAW is set at a certain distance in metres, then the number of cells within the radius increases as the cell size is reduced. For example, with a NAW of 3000m and a cell size of 500m, there will be approximately 110 cells within that radius that will need to be processed. If the cell size was reduced to 200m then there would be approximately 700 cells that would need to be processed. Therefore, reducing the cell size not only increases the number of focal cells but also increases the number of neighbouring cells.

Cell size also has an effect on the spatial accuracy of the boundary of the classes. With larger cell sizes this will be less accurate. Also, when a vector coverage is converted to a grid then small objects may be lost if the cell size is too large.

To see how sensitive the classification is to cell size, a range of different cell sizes was experimented with. Figure 4.10 shows the results. There is not a significant difference between the use of 300, 500, and 700m cell sizes. The agreement between the use of 300 and 700m cell sizes is 95%. There is a difference in the coarseness of the boundaries of the classes. This is difficult to see at the scale used in Figure 4.10, but it is obviously coarser with the larger cell size. The variation resulting from different cell sizes is not great because the minimum size polygon in Landcare's vegetation database is 500 ha. The cell sizes are significantly less than this (9, 25, and 49 ha) so very little detail is lost during vector to raster conversion. It should be noted that the search radius was held constant. Different cell sizes affect the speed of the processing quite significantly. With a cell size of 500m the process can be completed in a couple of hours, but with a cell size of 300m approximately eight hours is required.

4.3 Naturalness

4.3.1 Introduction

Although people are familiar with the concept of naturalness, it is a difficult concept to define. In this thesis it relates to the degree of development or cultural influence in a landscape. It is chiefly concerned with the amount of cultural modification of the surface cover. Naturalness therefore spans a spectrum from very unnatural landscapes, such as urban environments, to untouched landscapes, such as wilderness areas. Whether a landscape is natural or not depends on how different aspects of human modification are perceived. Is an exotic forest perceived as natural and is this more natural than an agricultural landscape? What is natural depends very much on the individual and therefore requires public perception studies to ascertain this information scientifically.

For landscape classification it is not actually necessary to rank the naturalness of different areas, as a nominal classification is sufficient. Areas that are similar in naturalness need to be grouped together. This can be done by classifying naturalness character rather than ranking naturalness. Areas that are similar in terms of human modification need to be identified. It is possible to do this using a range of parameters as will be demonstrated in this chapter.

4.3.2 Past research

Naturalness has been a common attribute used in landscape studies, although a range of different approaches has been used to define it. Bennett (1985) ranked naturalness for different areal units using a score of one to five, and this tended to be an intuitive procedure using field observations, rather than using explicit guidelines. Linton (1970) incorporated naturalness with vegetation and landuse. He used fairly broad classes - urbanized and industrial, farmland, and wild landscapes. The Manchester study (cit.in Countryside Commission, 1988) incorporated a whole array of man made components, such as towns and villages, railways, roads, power lines, and buildings.

The application of GIS for classifying naturalness is not new, however, it appears that most initiatives have been focused at the wilderness end of the naturalness spectrum (Lesslie et al., 1988, Kliskey and Kearsley 1993). This study will attempt to identify the whole range of the naturalness spectrum, from urban and rural areas, to remote areas. There does not appear to have been any published research that investigates the use of GIS for identifying automatically this range of landscapes. The research on wilderness identification does provide a starting point from which a process can be developed.

Lesslie et al. (1988) used four indicators to identify wilderness. These were:

1) Remoteness from settlement,
2) Remoteness from access,
3) Aesthetic naturalness (free from structures), and
4) Biophysical naturalness.

Apart from biophysical naturalness, these indicators were obtained by using GIS to measure the nearest distance from each cell, in a raster representation, to various human made entities, such as settlements, roads, structures, and logging operations etc. These distance measurements were then used to derive the different indicators. These indicators were then classed and weighted before being combined to ascertain wilderness quality.

In this study, elements of biophysical naturalness have been classified under vegetation, and therefore it is not necessary to include this again in a naturalness classification. It is also not necessary for a landscape character classification to specify quality, but instead it should distinguish character that may explain differences in quality.

Kliskey and Kearsley (1993) used similar indicators as Lesslie et al. (structures, access, vegetation, and use levels). These indicators were mostly obtained using buffers around different unnatural entities in vector coverages. These buffer coverages were then overlaid and "wilderness purism" scores calculated.

As previously discussed in section 3.4.2, the use of neighbourhood mean functions is more appropriate than the nearest distance calculations used by Lesslie et al. and is also more appropriate than the buffer functions used by Kliskey and Kearsley (1993).

4.3.3 The automated process

The automated process developed in this study identifies 22 different classes of naturalness. These are listed in Table 4.2 under level 1. The major problem with classifying naturalness is that there is a lack of information on how people conceptualize naturalness. There is a certain amount of information at the wilderness end of the naturalness spectrum (Stankely and Schreyer, 1987, and Kliskey and Kearsley, 1993), but not at the other end. Although urban areas are already conceptualized in topographical maps and in NDDBs, this has not been based on how the public conceptualise urban areas. The intermediary classes between urban and wilderness have also not been explicitly conceptualised. Common language, such as "rural" "town", and "settlement", give some clues to how development in the countryside is conceptualised, however, it can be quite vague. The classes used by the ARA (1984) also help.

A clue that can be used for deciding upon different classes is the amount of contention that exists over different development initiatives in the countryside. In well settled areas such as the Canterbury Plains, people do not get too concerned, relatively, about the impact of new roads or buildings on the landscape, however, in undeveloped areas they do (for example the Bealey Hotel in Arthur's Pass). The quality of the landscape is more sensitive to subtle changes at the natural end than at the developed end of the naturalness spectrum. This implies that more classes are required at the more natural end, which is how the classes in this study have been organised.

The availability of information in digital databases also affects which naturalness class can be identified. The databases used in this study for identifying naturalness classes were:

262 (1:250,000) topographic database (DOSLI),
Digital Chart of the World (DCW) (ESRI), and
Supermap2 (Statistics New Zealand).

Access was not available to DOSLI's 1:50,000 topographical databases because of cost and also because it had not yet been fully developed for the study area. It is possible to speculate on the use of additional information from this source.

The main database used was the 262 topographical database. When one looks at a hardcopy of a 1:250,000 topographic map, it is possible to assess naturalness in different area based on the number of roads, structures, railways, pylons, urban areas, etc. The method used in this study attempts to simulate this assessment by using a method similar to that used for classifying vegetation. Here, 15 single component layers were obtained from the three vector databases listed above and converted to raster coverages (500m cellsize) with a value of 100 assigned to areas where the components are present, and zero where they are absent. The spatial influence of the components was then expressed using a focal mean function. Figures 4.11-13 show the spatial influence classes of the different components and the extent of the actual components themselves. The different neighbourhood search radii used for each component is stated in these figures, as well as the different class intervals used. It should be noted that the search radii and class intervals are not the same for each component. The reason for this will become apparent later. The spatial influence classes were then used as parameters for defining naturalness classes.

Urban areas were identified from the secondary roads labelled "urban" in the DOSLI topographic database. A focal mean function with a NAW of 3000m was used to determine the spatial influence of these roads. Urban areas were defined as areas where this spatial influence was greater than 10% (refer to Appendix 2). This was considered the best approach for identifying urban areas because it was consistent. The DOSLI topographic database does not contain a polygon coverage of urban areas. The DCW contains an urban area layer, and the Ministry of Forestry have also digitised an urban layer, but these were not considered as consistent as DOSLI's "urban roads" layer.

The towns, large settlements, and small settlements, were identified by integrating the settlement layer of the DCW with the population data from Supermap2. With Supermap2 it is possible to select groups of meshblocks confined together to constitute a town or city. Each group is assigned a place name, and because the meshblocks are reasonably confined together they can be used as point information without having the problem associated with generalising over a large area. It was possible to automatically relate the DCW's location of places with Supermap2's population of towns or cities using place names. Some modification to the place names of the DCW coverage was required because it is necessary that these be spelt exactly the same as place names in Supermap2 in order for this transfer of data to work. Also, places were added to the DCW that were distinguished by Supermap2 as a town or city but were not present in the original DCW. Small settlements were places identified in the updated DCW but were too small to be distinguished by Supermap2. Large settlements were identified by Supermap2 as having a population less than 500, and a town had a population greater than or equal to 500 but was not identified as an urban area described previously.

The 11 other component layers were themes represented in DOSLI's topographic database. The combined roads were derived from national and provincial highways, and sealed and unsealed secondary roads. The secondary roads (sealed and unsealed) excluded forestry and urban secondary roads.

The walking track layer of DOSLI's topographical database was not used because it was too inconsistent. In some places it was detailed and contained the same information as the hard copy maps, and in other places it did not.

Once all 15 spatial influence layers had been derived and classified, they were overlaid. This resulted in 2139 unique combinations for the study area. A vector representation of this coverage is shown in Figure 4.13. Twenty two naturalness classes were then derived by querying the attribute table of this combined coverage. The definitions are described in detail in Appendix 2, and the actual classes are listed in Table 4.2 under level one. Utility includes pylons and railways. Checks were made to ensure that the definitions were mutually exclusive and exhaustive as described in section 4.2.3. The result of this process for the study area is shown in Figure 4.14. It appears from this figure that a buffer function was used, however, this was not so. The classes identified have been chosen because of their importance in planning disputes. At the more natural end of the spectrum subtle changes in naturalness can be contentious therefore more classes are needed. The classes also reflect the information that was available which, as will be discussed in section 4.3.6, was deficient for identifying some classes.

This naturalness classification was then generalised by grouping classes. Table 4.2 shows how the classes were grouped for each of the six levels of generalisation developed. Figure 4.15 graphical illustrates the effect of this generalisation. No key is provided with this figure to avoid cramming, but the colours are the same as used in Figure 4.14 and the key can be ascertained by using this and Table 4.2. Like level 1, the classes in level 2-6 maintain more detail with the more natural classes because of the contentiousness at this end of the spectrum. However, at a very general level, even detail here is lost.

A few details in the naturalness classification process warrant further clarification. To identify reasonably developed areas in the countryside (which are classed as "developed rural" and "rural" in this study), the spatial influence of the combined road layer was used. This was considered the best indicator of this class, although these areas contain much more infrastructures than this. It is fairly safe to assume that where there is a lot of development then there will be a high density of roads. The more intense the farming, the more activity there will be, and therefore the more access that will be required. For the study area a 10,000m radius NAW was considered the most appropriate for identifying the classes "developed rural" and "rural" because they are very general. This was decided upon after examining the results from a range of different search radii, from 3000m to 20,000m. The density of buildings could have also been used for identifying these classes as it shows a similar pattern as the density of roads. However the structures layer is probably not as consistent as the road layers because of the difficulty in defining a structure. It was therefore considered less effective.

Not all the information that was available was used in the definitions. For example, although national highways and provincial highways were identified separately, they were grouped together in the definitions. After different options were considered, it was decided that deriving separate classes based on these components was not necessary. The difference in naturalness between national and provincial highways is too inconsistent to be of any use in a landscape study. For example, is there a significant difference in naturalness between the Lewis Pass road, which is a national highway, and the Arthur's Pass road, which is a provincial highway?

An outcome from this process, which is perhaps undesirable is the very small slithers distinguished at some levels. For example, along the road across Arthur's Pass there are many small areas identified as "utility" and "highway", rather than "highway with utility". This is because the highway, and pylon components are not occupying the same area. The pylons are often located a few hundred metres from the road. The spatial influence of these two components therefore differs, and it is not possible to allow for this in the definitions without adverse effects.

4.3.4 Cell size

This automated process for classifying naturalness is affected quite significantly by cell size. The reason for this is that the original vector coverages consists of lines and points. When these are converted to a raster image, each cell is generalised so that any cell that overlays a point or line will be assigned the attribute value of the line or point. When a vector coverage of roads is converted to a raster coverage with a cell size of 500m then the road is represented by 500m wide cells, although the road may in reality be only 20m wide. If a 20m cell size was used then the raster coverage would be closer to reality. One may think that using a 500m cell size will lead to major errors. However, since roads are likely to have a spatial influence of more than 500m this is not a serious problem for landscape classification. The difference in outcome between different cell sizes occurs when focal means are calculated. Consider an isolated straight road converted to a raster grid with a cell size of 500m, with the value 100 assigned to cells where the road is present and zero where it is absent, and the focal mean calculated from a surrounding (square) neighbourhood of 2500 hectares (10 X 10 cells). The focal mean of the cells where the road is present will be equal to the number of cells where roads are present in the neighbourhood, which will equal 10, times 100 (the value of these cells), divided by the total number of cells in the neighbourhood (100). This would be:

$$10 \times 100 / 100 = 10$$

Now if a cell size of 20m was used and the neighbourhood extent stayed the same, the number of cells in the neighbourhood would be 62500 (250 X 250 cells), and the number of cells where roads are present within the neighbourhood would equal 250. The focal mean for cells where roads are present would therefore be:

$$250 \times 100 / 62500 = 0.4$$

These focal means are used to define naturalness, and since these values change with cell size then the definition will also change with cell size. Thus, for a given NAW area, the naturalness definition is dependent on the cell size. This sensitivity to cell size could be reduced if the definitions were based on the actual area of the components (eg. hectares of highway), rather than the number of cells representing the components. However, this would require knowing the average area of a cell that a component occupies. This is difficult to ascertain.

4.3.5 The use of Supermap2

As mentioned in section 3.3, Supermap2 is a database of the census results, produced by Statistics New Zealand, and organised by meshblocks or areal units. Among other things, it is possible to use information in Supermap2 to get an impression of development in different areas. Supermap2 contains information on the number of dwellings and this can be subdivided by the type of dwelling - hotel, motel, house, etc. There is also information on the population of different meshblocks. Available as "clip-ons" to Supermap2 are the business directory databases, which contain information on the extent of different industries in different regions based on the number of employees. All this information could be used to develop an impression of the type of development that exists in different areas. It is possible to use Supermap2 to distinguish an area as very tourism orientated, based on the number of hotels or the number of employees working in hotels. However, Supermap2 has not been used significantly in this classification process because of problems relating to spatial accuracy.

A major problem with Supermap2 is that the information is organised by meshblocks whose sizes are related to population density. In towns and cities the meshblocks are quite small, while in rural areas these can be quite big. The spatial accuracy of the information is affected by the size of the meshblocks, and so in rural areas the spatial inaccuracy can be quite significant. In this study the emphasis is on rural areas therefore if Supermap2 data were used significant error would arise. Figure 4.16 demonstrates this for population and dwelling numbers in the Mackenzie district. These are mapped as density in order to correct for different meshblock areas. Because of the large meshblocks, significant generalisation occurs which can lead to misleading results, such as along the road between Twizel and the Mount Cook village, most of the dwellings and population are close to the main road, yet the meshblocks extend well beyond the road. It would be inappropriate to generalise the effect of these dwellings and populations over the whole extent of the meshblocks. Furthermore, if the meshblock boundaries were organised differently, it is highly probable that the statistics in this figure would change significantly. This is commonly known as the modifiable areal unit problem, which was reviewed recently by Wrigley (1995).

Meshblock data are appropriate for classifying towns or cities because in these areas the meshblocks are more confined so less generalisation occurs. The process does this at a very general level by determining the population of different urban areas.

To distinguish agricultural and forestry landscapes in rural areas, it is better to use other parameters than meshblock statistics. Agricultural landscapes can be identified by vegetation and by the number of roads (which indicates activity), while forestry can be accurately identified by the presence of exotic plantations.

4.3.6 Information deficiencies

The process described above identifies as many important classes of naturalness as possible from existing databases. There are, however, some important classes that have not been possible to identify. These are rural landscapes affected by tourism, mining, electricity generation, and other industries. If any of these industries are big enough at a particular location than perhaps they would be identified as a settlement or town, however, often they are not. As discussed in section 2.4, tourism development is a major contentious landscape issue, yet the extent of this industry is not mapped in rural areas. Some information may be ascertained from the 1:50,000 topographic database on the extent of ski fields, but not accommodation. The structures layer of the topographical databases only specifies that one or more buildings exist at a particular location, and it does not specify the type or size of the structure. The structure could be a small farm house or a hotel with 100 rooms.

Information on mining activities would be useful as this is also a contentious landscape issue in rural areas. Mining sites are mapped in topographical databases but the information is not detailed enough. It is only specified that at a particular point there is a mine. It is not possible to ascertain whether the mine is an underground mine, an opencast mine, or an old derelict mine overgrown with vegetation. It is therefore inappropriate to use this information for landscape classification as these different mines have significantly different impacts on the landscape. It would also be useful to have other major industrial sites mapped out and available in digital format. They would be relatively cheap to produce as they would only consist of point information, and there are not many major industries in rural areas. Such information could also be used for other planning issues, for example transportation. Statistics New Zealand does survey all industries, and it would not be too difficult for them to obtain grid references of each industry and map this information. However, Statistics New Zealand is restricted by law from making this information public. It appears from the directories of available databases that there has perhaps been more emphasis on mapping nature, for example animals, plants, and wetlands, and less emphasis on unnatural and potentially harmful things, for example industries, and hydro dams. Both types of information are needed to address environmental issues.

4.4 Influence of water

Classifying the influence of water, for the purposes of this study, requires classifying coastal areas, and identifying rivers and lakes. There is very little discussion about this in the literature regarding landscape classification, except for rivers (Mosley, 1989), and even less on doing this automatically within GIS.

4.4.1 Classification of coast

Early coastal classifications recognized two classes - emerging and submerging (Davis, 1902). Shepard (1937, and 1938) introduced the concepts of primary and secondary coasts. Primary coasts were based on the influence of land based processes, such as fluvial activity, glaciation, aeolian processes, or denudational processes. Secondary coasts were based on marine processes. Valentine (1952) developed a coastal classification that incorporates the above classifications.

Landscape classification is not concerned with genesis but only with the contemporary appearance. Therefore, Valentine's classification is inappropriate. For a coastal classification to be useful in a landscape classification, it also needs to be at a very generalised level. This is because landscape classification needs to incorporate a wide range of other attributes. Coastal morphology is more relevant for this study. Weerakkody (1993) used remote sensing to identify three features important to coastal morphology - the coastline indentedness, plan-curvature, and orientation. Indentedness was described as being formed by headlands, islets, spits, river mouths, lagoonal outfalls, beach rocks, rock outcrops, sea cliffs, coral reef, and engineering structures. Plan-curvature of the coastline uses notions such as concave, convex, or straight to describe the coastlines that are not particularly indented. Orientation was used because of the effects this has on marine activity, such as the refraction effect of waves, longshore drifting, and direction of littoral currents.

Existing digital databases in New Zealand contain very little information on the coast. Although beach rocks, rock outcrops, sea cliffs, and engineering structures appear on hard copy maps, this information has not been digitised. All that is available is an outline of the coast. From this it is possible to see what coasts are indented and which are not, however this information has not been conceptualised in the database. The coastline therefore needs to be spatially analysed to distinguish these classes.

It was decided that the coastal classification for this study would have four classes at its most detailed level - indented, very indented, non indented, and non-coastal.

Indentedness is considered important because of the prevalence of common language that describe this feature of the landscape, for example bay, headland, inlet, fiord, and sounds.

4.4.2 The coastal classification process

Indented coastlines can be identified by using expand and shrink functions. These functions actually expand or shrink a specified class by a specified number of cells. Figure 4.17 shows the different steps in the process for the Banks Peninsula region. The process starts with a grid that has one value for land and another value for sea. The land is then expanded 2500m (five cells) into the sea, using an expand function. This output is then shrunk by the same amount. The net effect of this expand/shrink sequence is that indented sea becomes land and the only sea is open. The semi-enclosed sea can then be identified by comparing the original land coverage with this open sea coverage. The coast can then be easily classified as indented or non-indented depending on the percentage of the neighbourhood that is indented or non-indented. If there was an indented coast within a 5500m radius then the coast was classified as indented. A very indented coast was defined as an indented coast that was further than 9500m from the open sea. This was identified using a buffer function on the open sea grid. Three coastline types were therefore classified. The spatial influence of these was set at 3000m inland using a buffer function which acted equally on all three classes.

The original land grid was obtained by converting DOSLI's 1:250,000 topographic polygon coverage of coast to a raster coverage with 500m cell size. The DCW or the Ministry of Forestry's database could have also been used, but they were less detailed, especially the DCW. The actual coastline can be identified by converting the arcs of the polygons to a grid coverage. However, this did not overlay precisely with the land grid because the generalisation effects of converting lines to grids are different from converting polygons to grids. Instead, the process shrinks the land grid by one cell, and then the difference between this and the original land grid gives the coastline.

This process expands and shrinks land by 2500m to identify indentedness. The effect of using 2500m is that inlets or bays that are less than 5000m across are identified as indented. One may question why 5000m should be used as a threshold and not 1000m or 10,000m. If a large number is used, reasonably straight coastlines that are also slightly concave become identified as indented. A range of thresholds was experimented with, and it was decided that 5000m was the most appropriate. The other thresholds that needed to be specified were also determined through experimentation. This includes the 9500m from the open sea used to distinguish between indented and very indented, and the 3000m radius NAW used to determine the spatial influence of the indented sea. What these thresholds should be is fairly arbitrary. What is important, however, is that these figures are explicitly stated.

4.4.3 Classification of rivers and lakes

Mosley (1989) characterised a range of different "riverscapes" based on an extensive perception study. Many of these characteristics are covered in this study by the other landscape attributes. Also, some characteristics are too detailed to be included in a total landscape classification, such as river straightness, and eroded banks. It was decided that it was appropriate to only include rivers and lakes over a certain size. This is consistent with Linton (1970) and the need to generalise.

DOSLI's topographical databases, DCW, and Ministry of Forestry databases contain river and lake layers. The DCW contains both rivers and lakes in one hydro layer, but these cannot be distinguished using attribute information. The Ministry of Forestry databases could be used, but it was initially considered that DOSLI's topographical database was better because it had information on the size of the rivers. It was also considered more consistent and accurate.

Large lakes were distinguished from smaller lakes by size using a threshold of 500 hectares. The spatial influence of these large lakes was then ascertained using a focal mean function with a 3000m search radius. If a cell had a large lake present within a 3000m radius then it was classified as lake.

It is possible to identify regions in which there are many small lakes. This can be easily done by first identifying the small lakes and then using a focal mean function. This would identify another class, but was not done because of the need to generalise.

Larger rivers can be identified by their hierarchy level given in DOSLI's 1:250,000 database. This hierarchy appears to be based on the Strahler method (cit.in Petts and Foster, 1985), except for level 7 (the highest level), which is braided rivers. Level 7 is difficult to classify according to size, because it includes the large, braided rivers, such as the Rangitata, and many small, braided reaches of small rivers. It was not possible to automatically distinguish large braided rivers from small braided rivers. Since the inclusion of sections of small braided rivers was not appropriate for this study, a database developed by Landcare was used instead. This database was derived from DOSLI's 1:250,000 database by isolating the levels 5, 6, and 7, and manually deleting the small braided reaches. It is not possible to say exactly what specifications were used. Once a suitable database of rivers had been obtained, the spatial influence of rivers was determined, as for lakes, with a 3000m NAW.

4.4.4 Water classification

Once the coast, lakes, and rivers had been classified, and their spatial influences determined, a water classification was produced by overlaying this spatial influence information with a process similar to that used for classifying vegetation, and naturalness. Figure 4.18 illustrates this process. Eleven unique combinations were identified out of a possible 16. Since there were a low number of possible combinations, all of these were used for level 1 of the water classification. Figure 4.19 shows this classification. The names of the classes indicate how they were defined, however, precise definitions are given in Appendix 3. Like the other main landscape attributes, this classification was generalised down to six different levels by hierarchically grouping classes together. Table 4.3 shows how this was done, and Figure 4.20 illustrates graphically the results. Again, no keys are provided with this figure to avoid cramming, but the colours are the same as used in Figure 4.19 and the keys can be ascertained by using this and Table 4.3. The classes used at each level of generalisation reflect important differences in appearance and contentiousness. The distinction of a coastal class was maintained throughout the generalisation because these areas are given special consideration in the Resource Management Act.

The process for classifying water is sensitive to cell size because rivers and coasts are linear features that become misrepresented with vector to raster conversion, as discussed with regard to the naturalness classification (refer to section 4.3.4).

4.5 Summary

The processes used to classify vegetation, naturalness, and water, follow a common sequence of steps. Once the appropriate NDDBs have been decided upon, the important landscape components are identified by grouping or generalising different objects in the NDDBs, and generating single theme (or binary) coverages. Except indented coastlines, all the components are conceptualized in existing NDDBs, thus making this step relatively simple. It was possible to conceptualise indented coastlines by using a combination of expand and shrink functions. The spatial influence of each component is then calculated with a focal mean function, and this information is then used to define landscape attribute classes using overlay composites.

The following decisions were required to classify vegetation, naturalness, and water:

- the NDDBs used,
- the generalisation of these NDDBs,
- the determination of spatial influence, which included
    - the size of the neighbourhood analysis window (NAW), and
    - the spatial influence thresholds,

. the definition of the attribute classes, and

. the cell size.

With the classification of coastlines it was also necessary to decide what constitutes an indented coastline, and what distance from the open sea makes an indented coastline very indented.

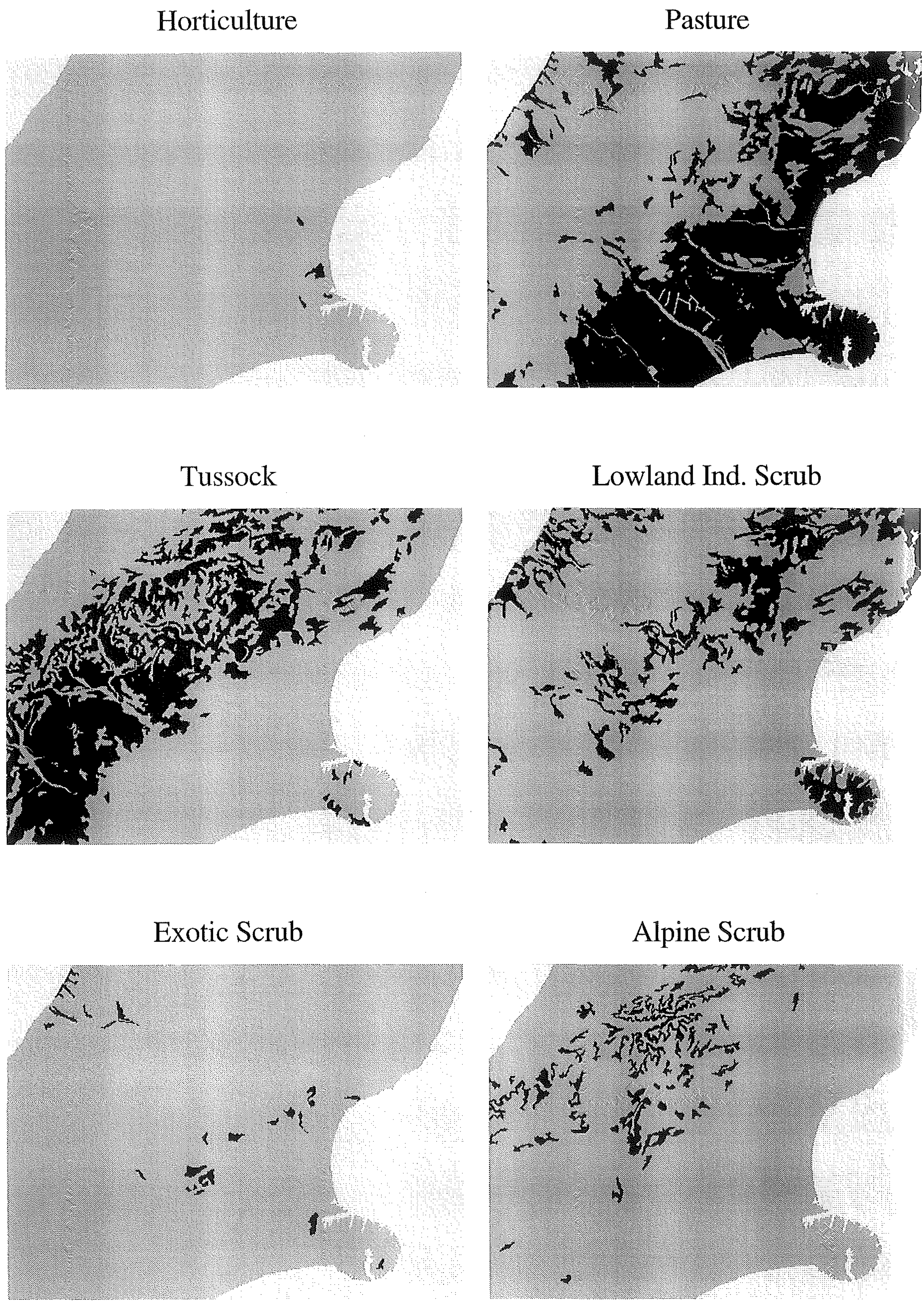
The sensitivity of the classification to the size of the NAW, and spatial influence thresholds was investigated and found to be substantial. The effects of different cell sizes depended on whether the components were originally represented by lines or polygons. Objects originally represented by lines were distorted significantly by vector to raster conversion and this subsequently affects the class definitions. This distortion depends on cell size and needs to be built into the definitions. To avoid having different sets of definitions for different cell sizes, it is necessary to decide on an appropriate cell size. With a cell size of 500m the spatial detail of NDDBs was not unduly lost, the necessary components can be represented, and the processing speed is acceptable.

Figure 4.1 Vegetation classes used in Landcare's vegetation database

	CROPLAND
C1	Orchards or vineyards and pasture
C2	Horticultural crops and pasture
	GRASSLAND
G1	Improved pasture
G2	Unimproved pasture
G3	Short tussock grassland
G4	Snow tussock grassland
G5	Short tussock—snow tussock grassland
G6	Red tussock grassland
	GRASSLAND-SCRUB
GS1	Grassland and mixed indigenous scrub
GS2	Grassland and *Leptospermum* scrub or fern
GS3	Grassland and *Cassinia* scrub
GS4	Tussock grassland and sub-alpine scrub
GS5	Grassland and *Dracophyllum* scrub
GS6	Grassland and gorse scrub
GS7	Grassland and matagouri
GS8	Grassland with sweet brier or sweet brier and matagouri
	SCRUB
S1	Mixed indigenous scrub
S2	*Leptospermum* scrub or fern
S3	Sub-alpine scrub
S4	Gorse scrub
	GRASSLAND-FOREST
GF1	Pasture and podocarp-broadleaved forest
GF2	Pasture and broadleaved forest
GF3	Pasture and beech or podocarp forest
GF4	Pasture and exotic forest
	GRASSLAND-FOREST (cntd)
GF5	Tussock grassland and beech forest
GF6	Tussock grassland and podocarp-broadleaved-beech forest
	FOREST-SCRUB
FS1	Kauri and *Leptospermum* or mixed indigenous scrub
FS2	Podocarp-broadleaved forest and scrub
FS3	Podocarp-broadleaved-beech forest and scrub
FS4	Beech forest and scrub
FS5	Beech-broadleaved forest and scrub
FS6	Broadleaved forest and scrub
FS7	Sub-alpine scrub and indigenous forest
FS8	Exotic forest and scrub
	FOREST
F1	Podocarp forest
F2	Lowland podocarp-broadleaved forest
F3	Highland podocarp-broadleaved forest
F4	Lowland podocarp-broadleaved-beech forest
F5	Highland podocarp-broadleaved-beech forest
F6	Beech forest
F7	Beech-broadleaved forest
F8	Broadleaved forest
F9	Exotic forest
	MISCELLANEOUS
M1	Sub-alpine or alpine herbfield
M2	Wetland communities
M3	Sand-dune communities
M4	Pakihi heathland communities

Figure 4.2 The Extent of the Different Vegetation Classes

Grey scale, with dark as present and bright as absent



Data source: Newsome and MOF	Cell size: 500m	Map ID. vegcant1

Figure 4.3 The Extent of the Different Vegetation Classes (cont.)

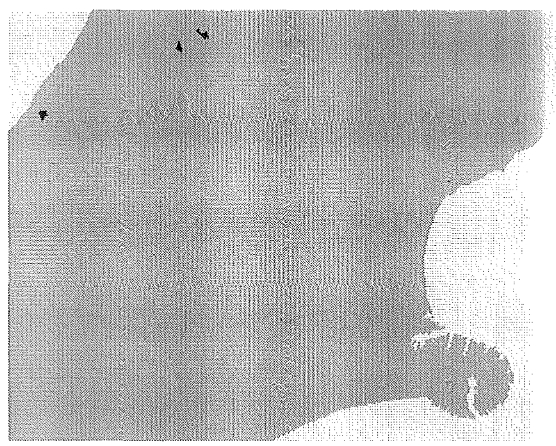
Grey scale, with dark as present and bright as absent

Indigenous Forest

Exotic Forest

Alpine Herbfields, Rock, or Ice

Wetland



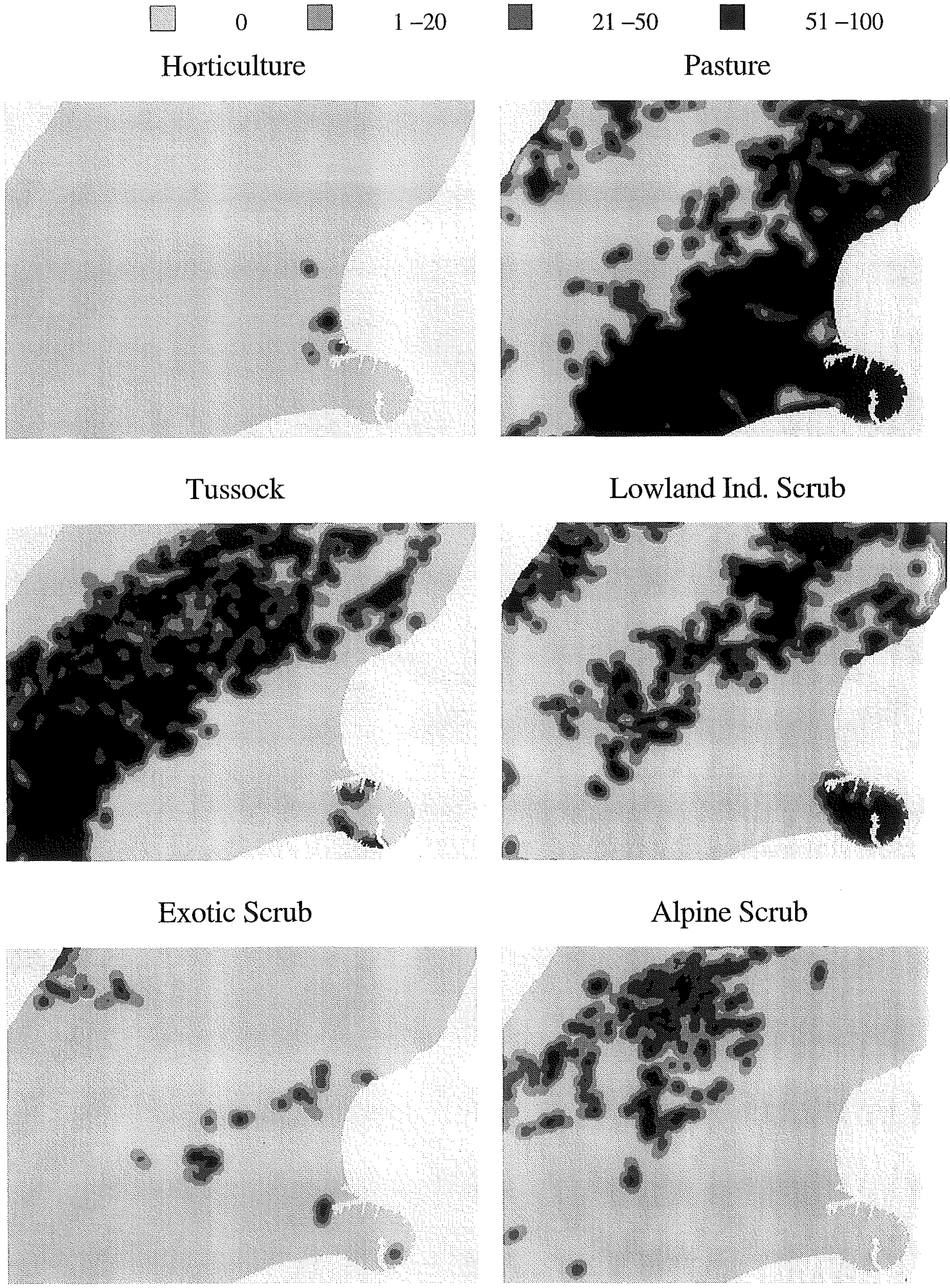
Sanddune

Urban, River, or Lake

Data source: Newsome and MOF	Cell size: 500m	Map ID. vegcant2

Figure 4.4 The Spatial Influence of the Different Vegetation Classes

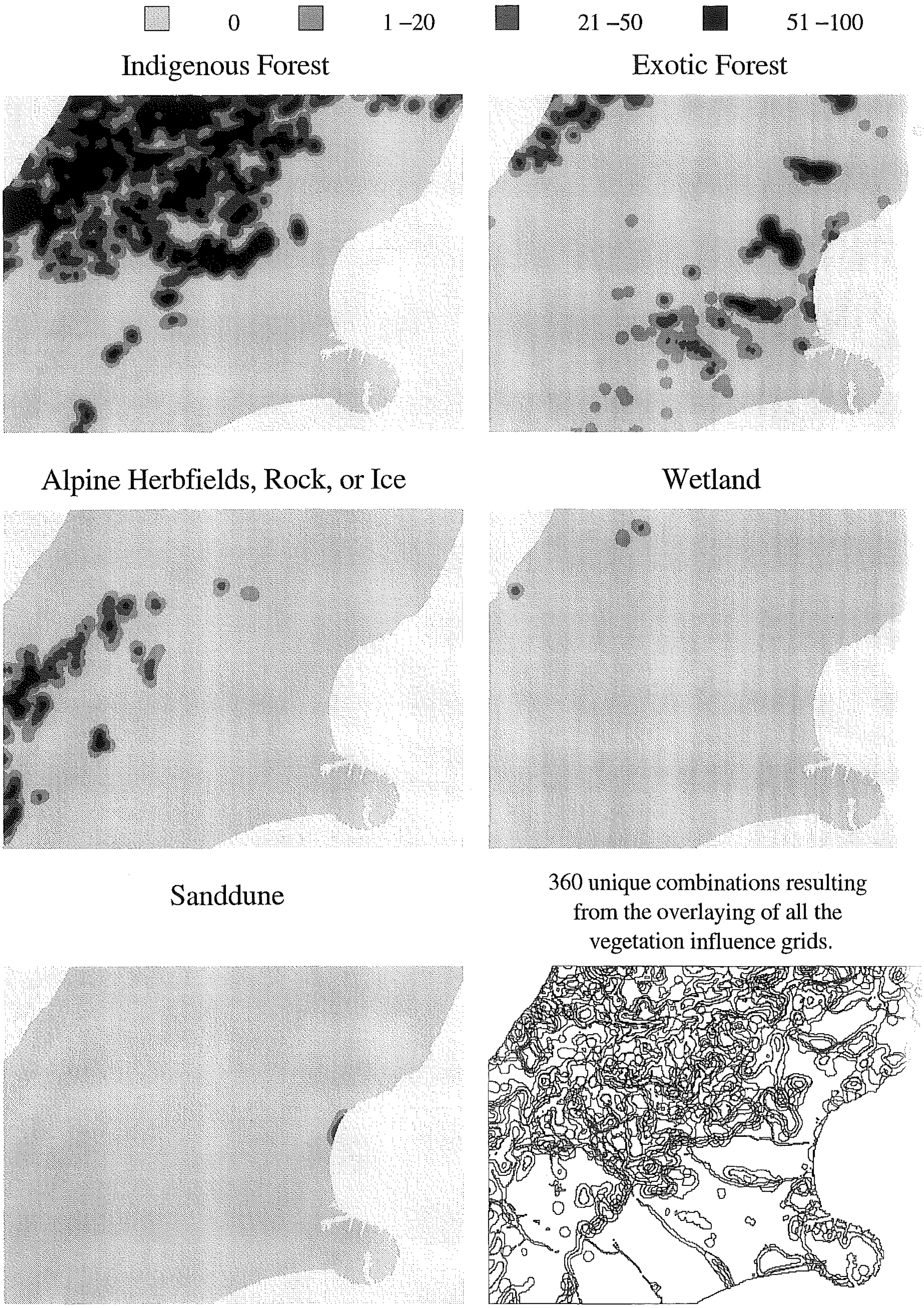
(The percentage of cells within the search radius that contain the specified vegetation class)



Data source: Newsome and MOF	Cell size: 500m	NAW radius: 3000m	Map ID. vegcantme[illegible]

Figure 4.5 The Spatial Influence of the Different Vegetation Classes (co

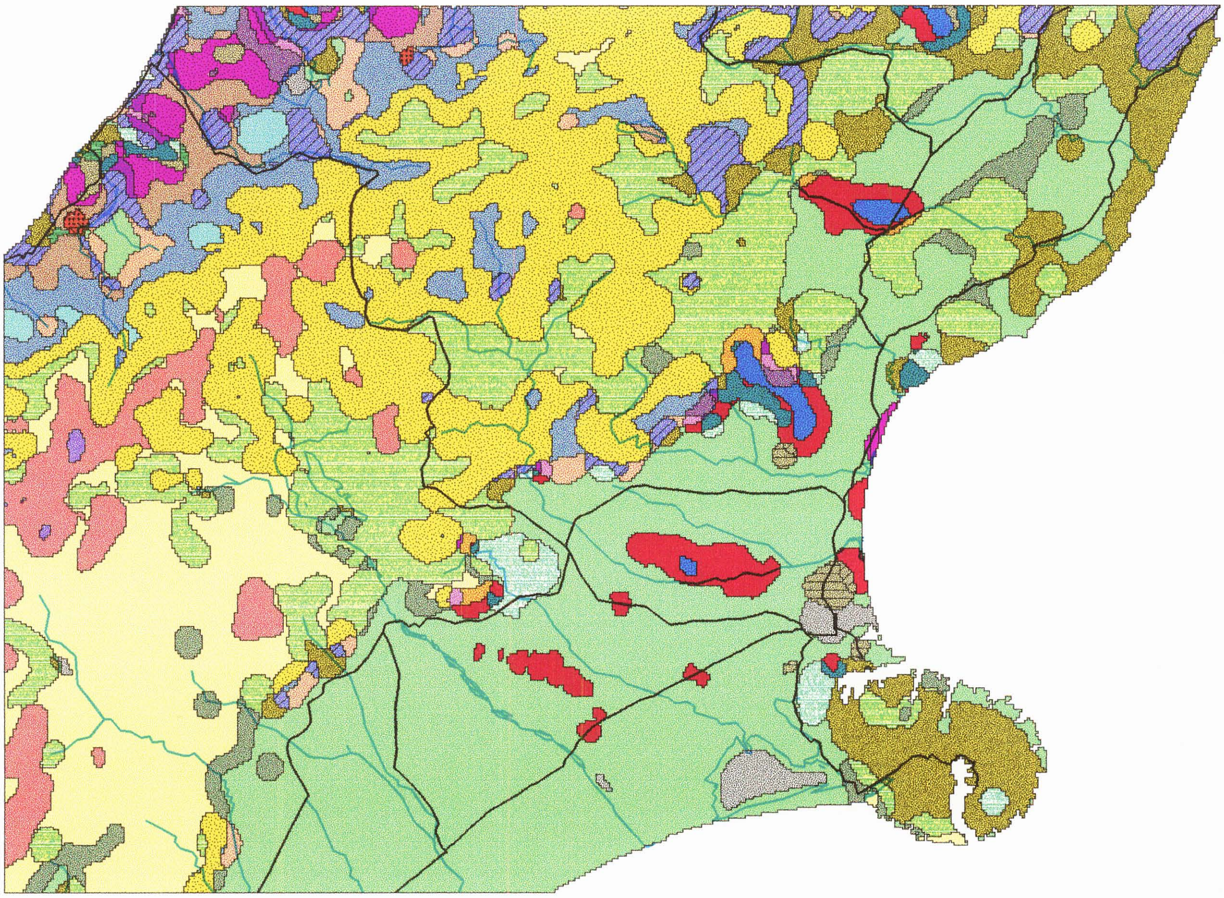
(The percentage of cells within the search radius that contain the specified vegetation class)



Data Source: Newsome and MOF	Cell size: 500m	NAW radius: 3000m	Map ID. vegcantme

Figure 4.6 Vegetation Level 1

Main road and hydrology layers are added for geographical reference

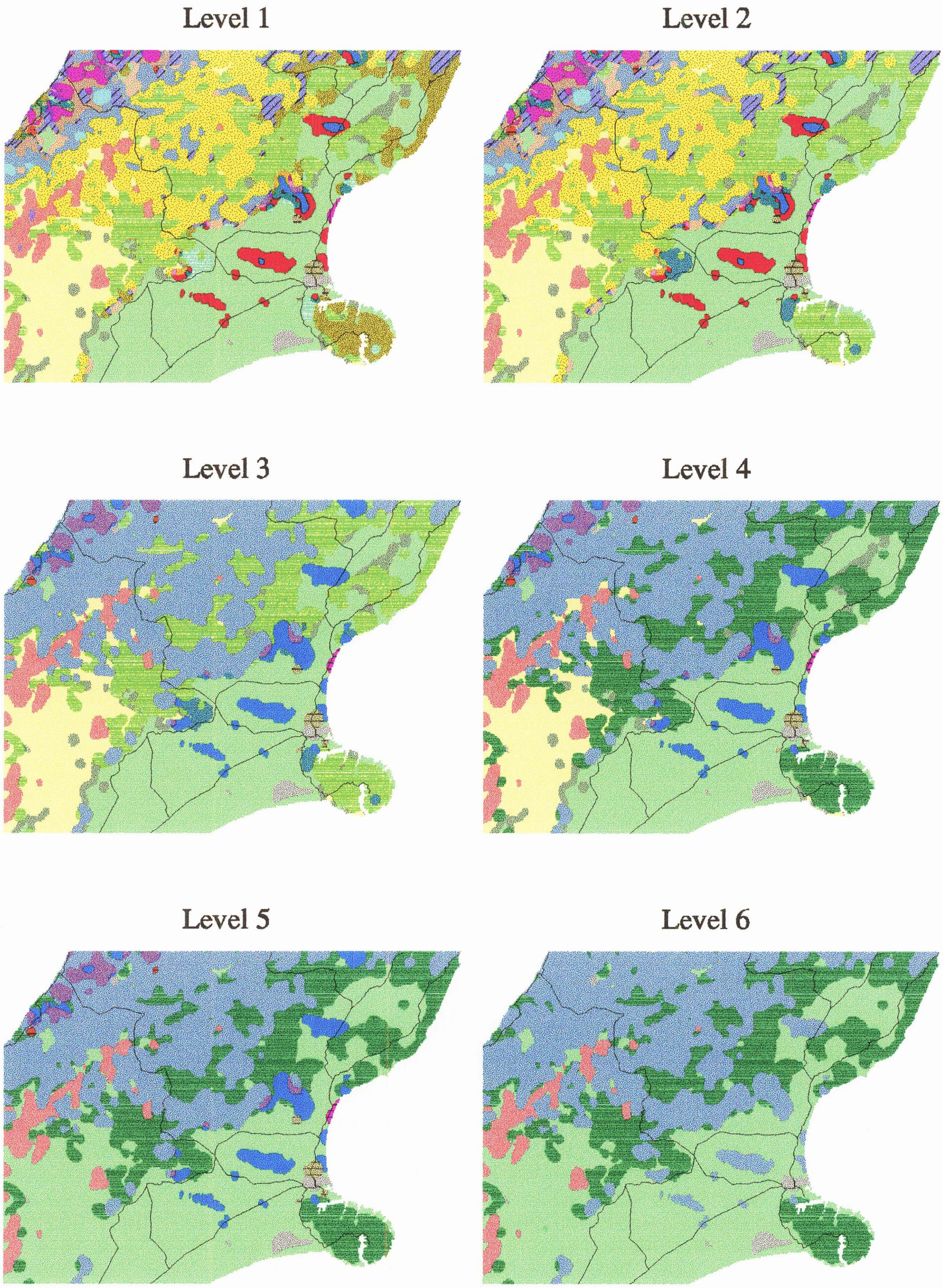


- Hort. pasture
- Hort. scrub
- Hort. forest
- Pasture
- Tussock
- Tussock pasture
- Pasture ind. scrub
- Pasture ex. scrub
- Tussock grassland scrub
- Lowland ind. scrub
- Pasture ind. forest
- Pasture ex. forest
- Tussock ind. forest (scrub)
- Tussock ex.forest (scrub)
- Ind.forest ind. scrub
- Ind.forest alpine scrub
- Ex.forest scrub
- Ind. forest scrub pasture
- Ex.forest scrub pasture
- Ind.forest
- Ex.forest
- Tussock Scrub Alpine
- Alpine herbfields, ice, and rock
- Mixed forest
- Mixed forest grassland
- Mixed forest scrub
- Mixed forest grassland scrub
- Wetland ind. scrub
- Wetland ind. forest
- Sanddune grassland
- Urban, waterbody, or not classified

Data source: Newsome and MOF	Cell size: 500m	NAW radius: 3000m	Map ID. veglev1

Figure 4.7 The Effect of Generalisation on Vegetation

Main roads are added for geographical reference

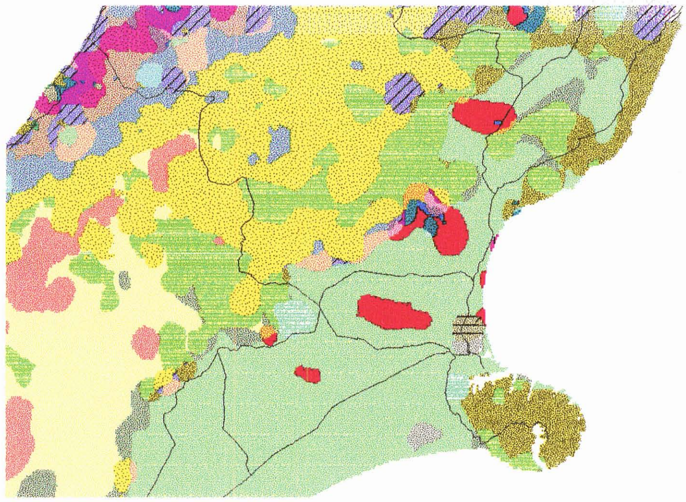


Data source: Newsome and MOF	Cell size: 500m	NAW radius: 3000m	Map ID. veglevl - 6

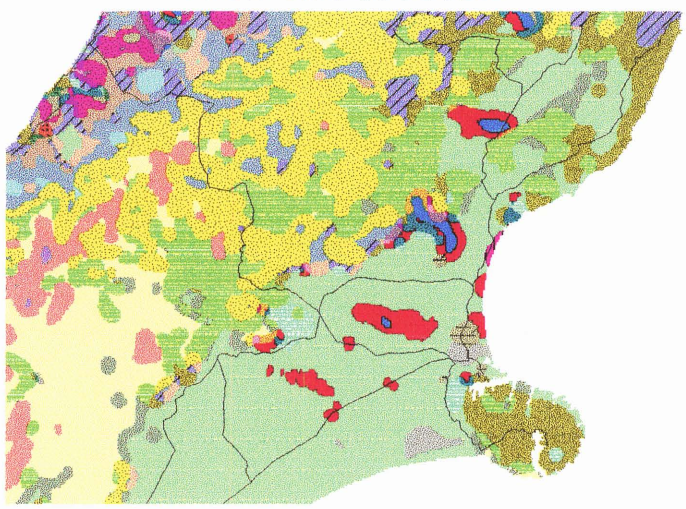
Figure 4.8 The Effects of Different NAW Radii

Main roads are added for geographical reference.

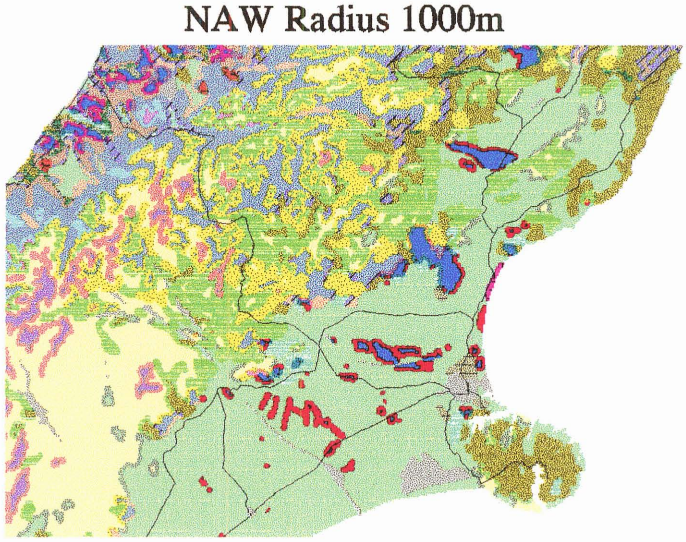
NAW Radius 5000m



NAW Radius 3000m



NAW Radius 1000m

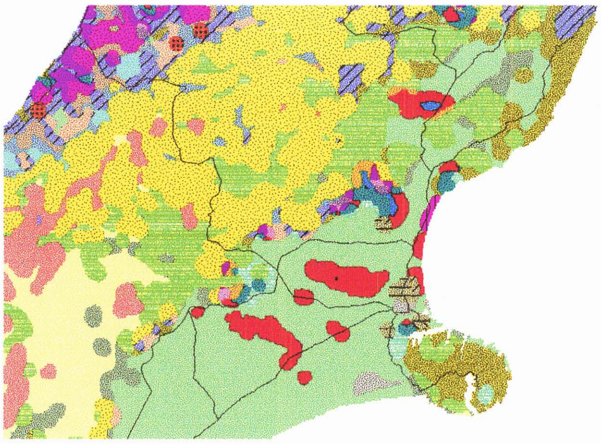


Data source: Newsome and MOF	Cell size: 500m	Map ID: vegnaw

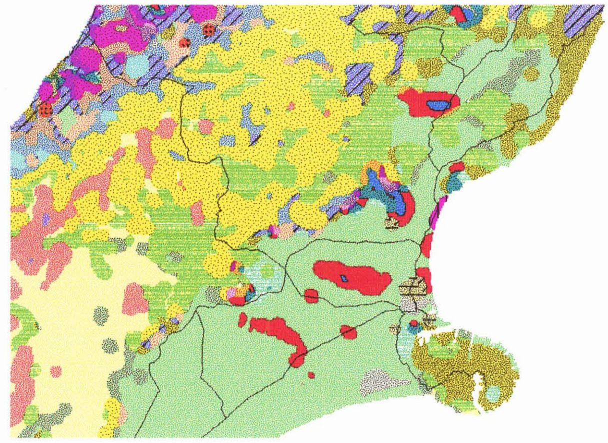
Figure 4.9 The Effects of Different Spatial Influence Thresholds

Main roads are added for geographical reference.

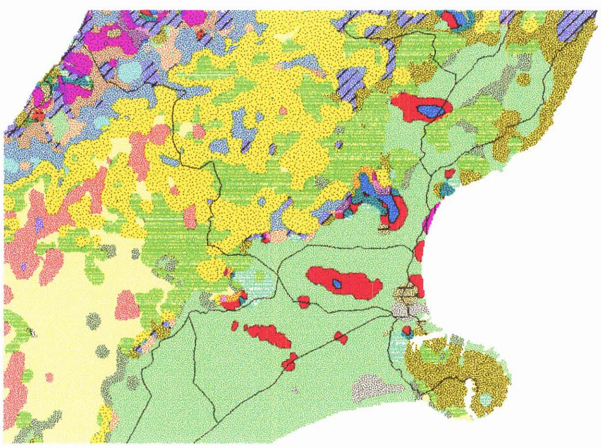
10 Percent Threshold



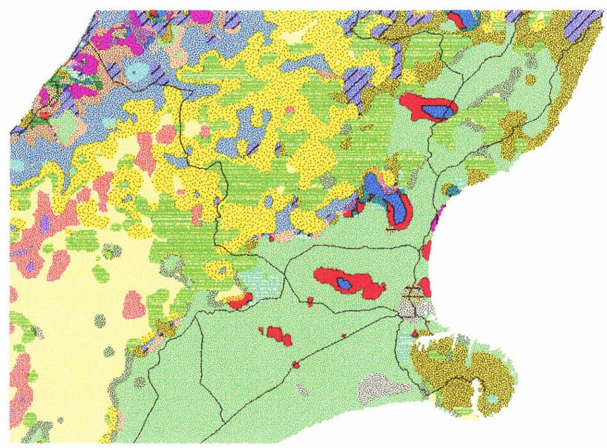
15 Percent Threshold



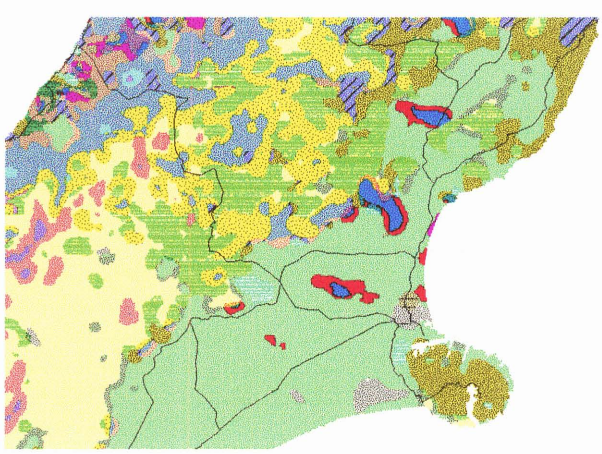
20 Percent Threshold



25 Percent Threshold



30 Percent Threshold

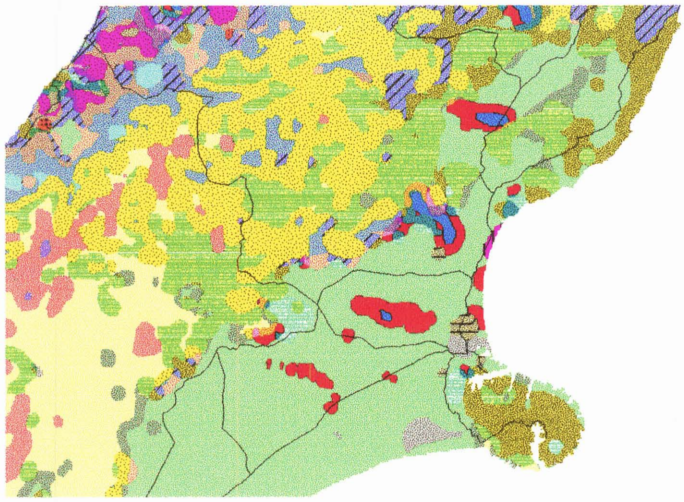


Data source: Newsome and MOF	Cell size: 500m	NAW radius: 3000m	Map ID: vegthreshold

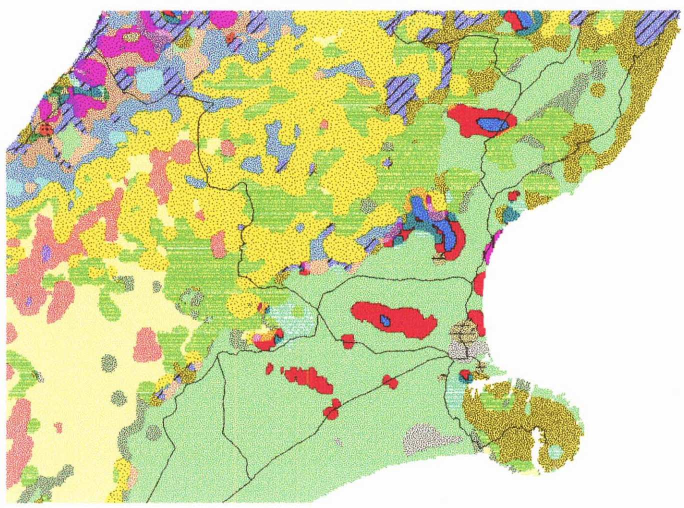
Figure 4.10 The Effects of Different Cellsize

Main roads are added for geographical reference.

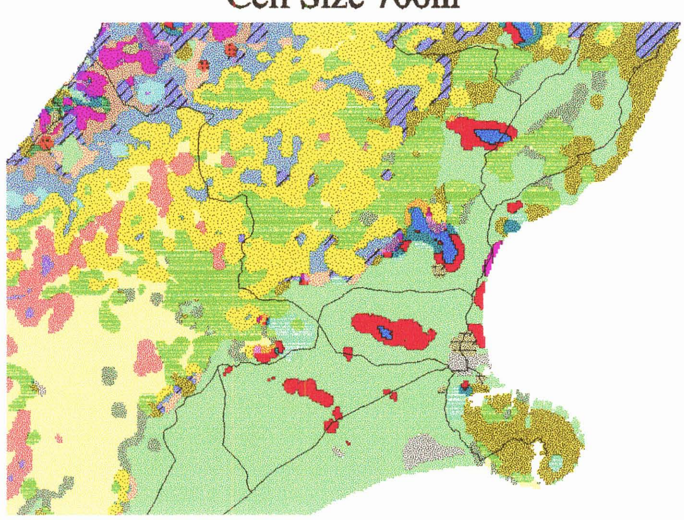
Cell Size 300m



Cell Size 500m



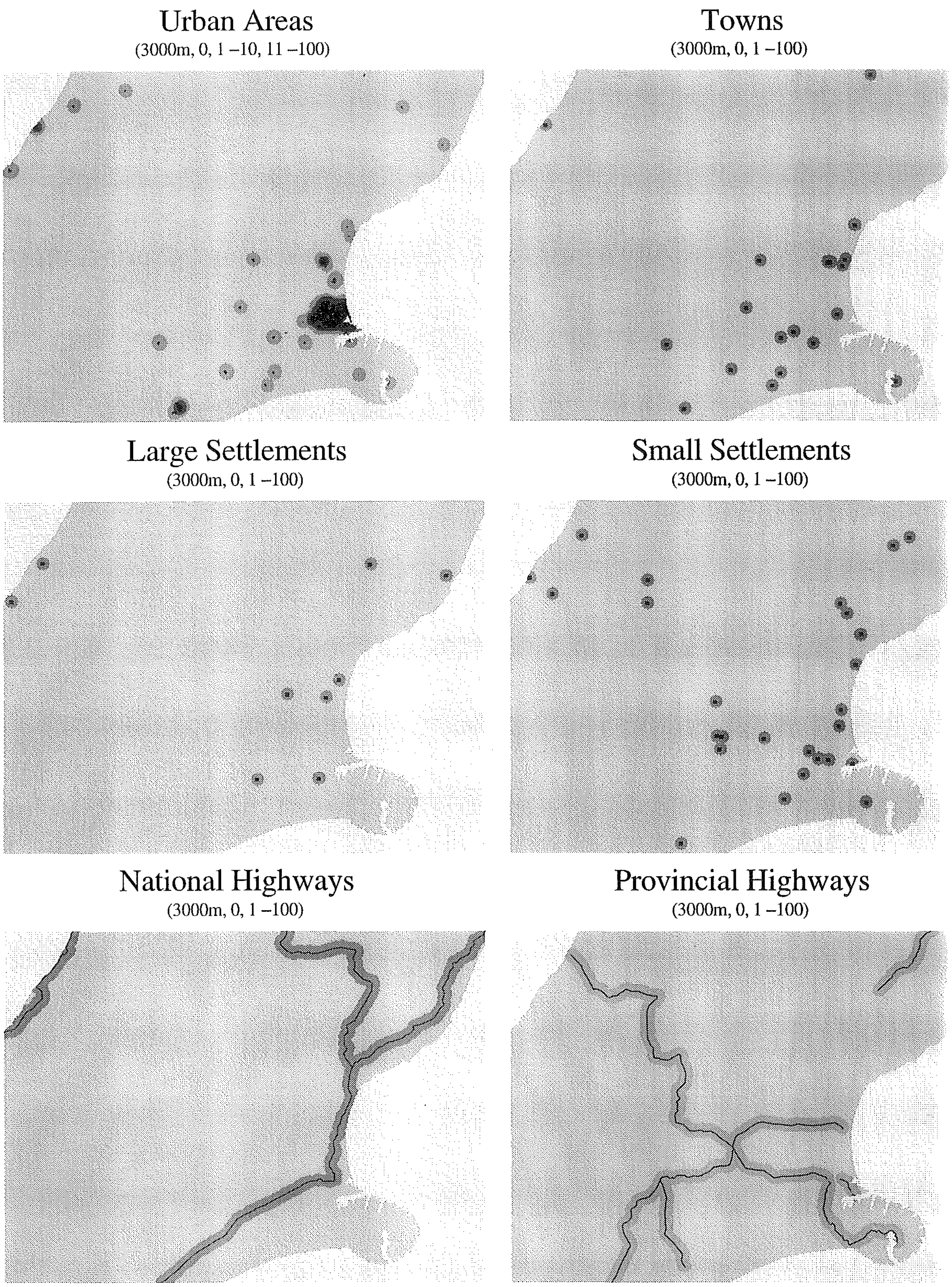
Cell Size 700m



Data source: Newsome and MOF	NAW radius: 3000m	Map ID: vegcellsize

Figure 4.11 Spatial Influence of Different Infrastructure

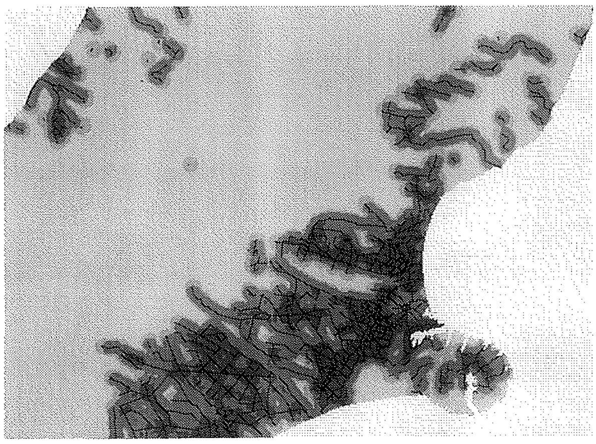
The percentage of cells within the specified NAW that contain the given infrastructure.
The figures in brackets are the NAW radii, followed by the class intervals that were used.
The number of class intervals varies. The lighter shadings represent less influence.
The actual infrastructure is also represented.



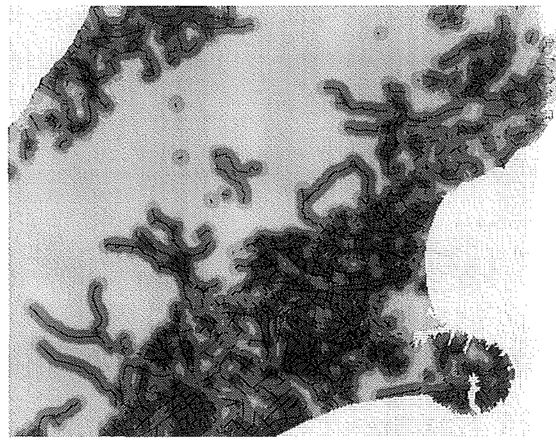
Data source: DOSLI 1:250,000, DCW, and Supermap2 | Cell size: 500m | Map ID. natcomponents

Figure 4.12 Spatial Influence of Different Infrastructure (cont.)

Sealed Secondary Roads
(3000m, 0, 1 –5, 6 –20, 21 –100)

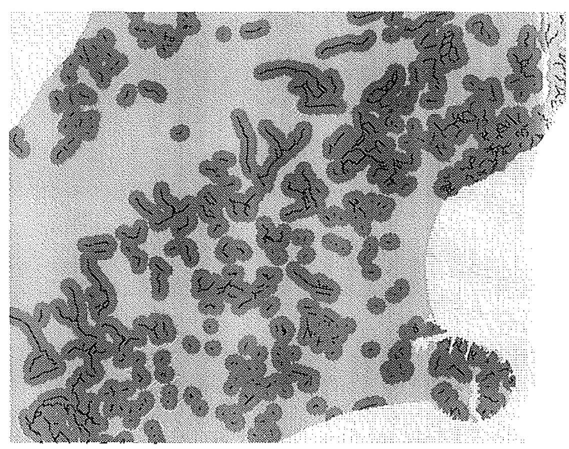


Unsealed Secondary Roads
(3000m, 0, 1 –5, 6 –20, 21 –100)

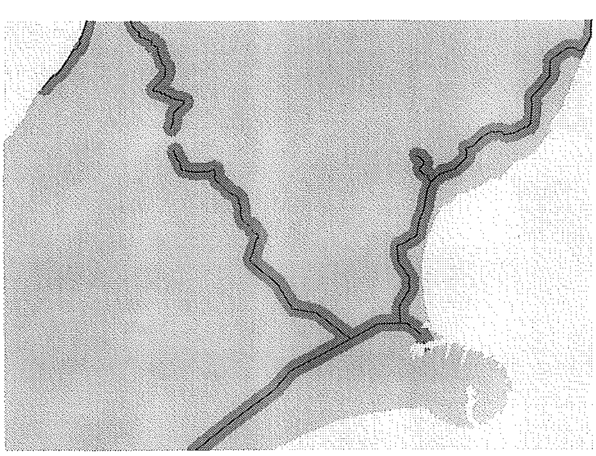


All Roads Combined
(10000m, 0 –10, 11 –20, 21 –50, 51 –100)

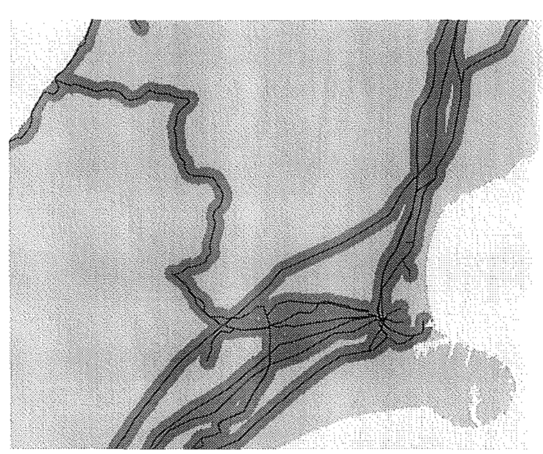
4 Wheel Drive Tracks
(3000m, 0, 1 –100)



Railways
(3000m, 0, 1 –100)



Pylons
(3000m, 0, 1 –100)

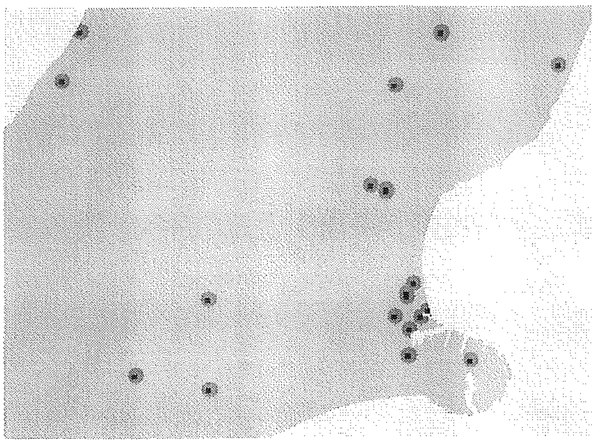


Data source: DOSLI 1:250,000, DCW, and Supermap2	Cell size: 500m	Map ID. natcomponents

Figure 4.13 Spatial Influence of Different Infrastructure (cont.)

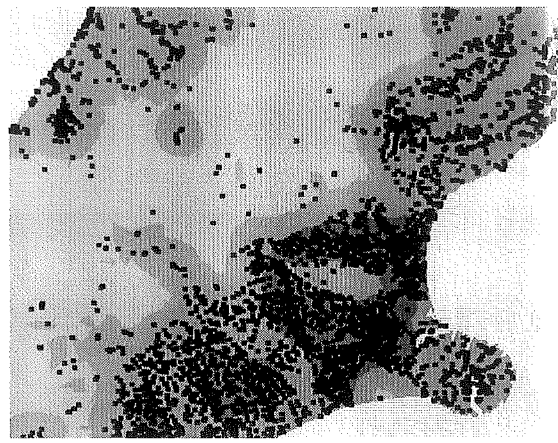
Radio or TV. Masts

(3000m, 0, 1 –100)



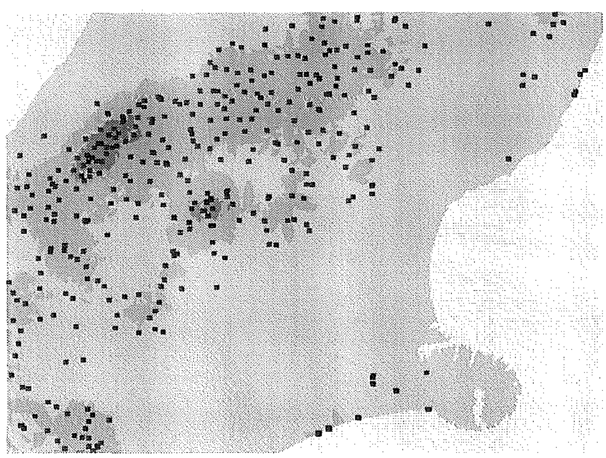
Buildings

(10000m, 0, 1 –7, 8 –20, 21 –100)

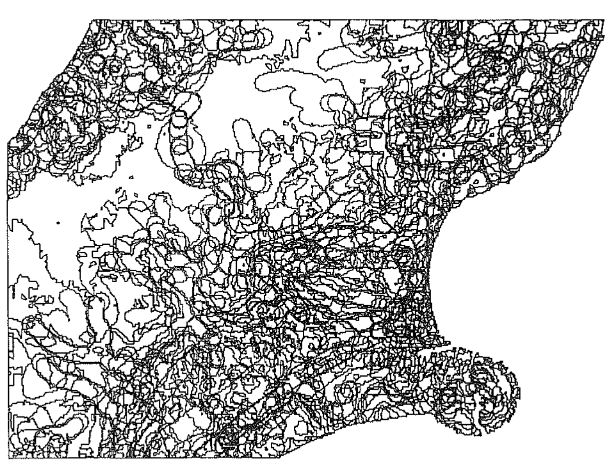


Huts

(10000m, 0, 1 –2, 3 –5, 5 –100)



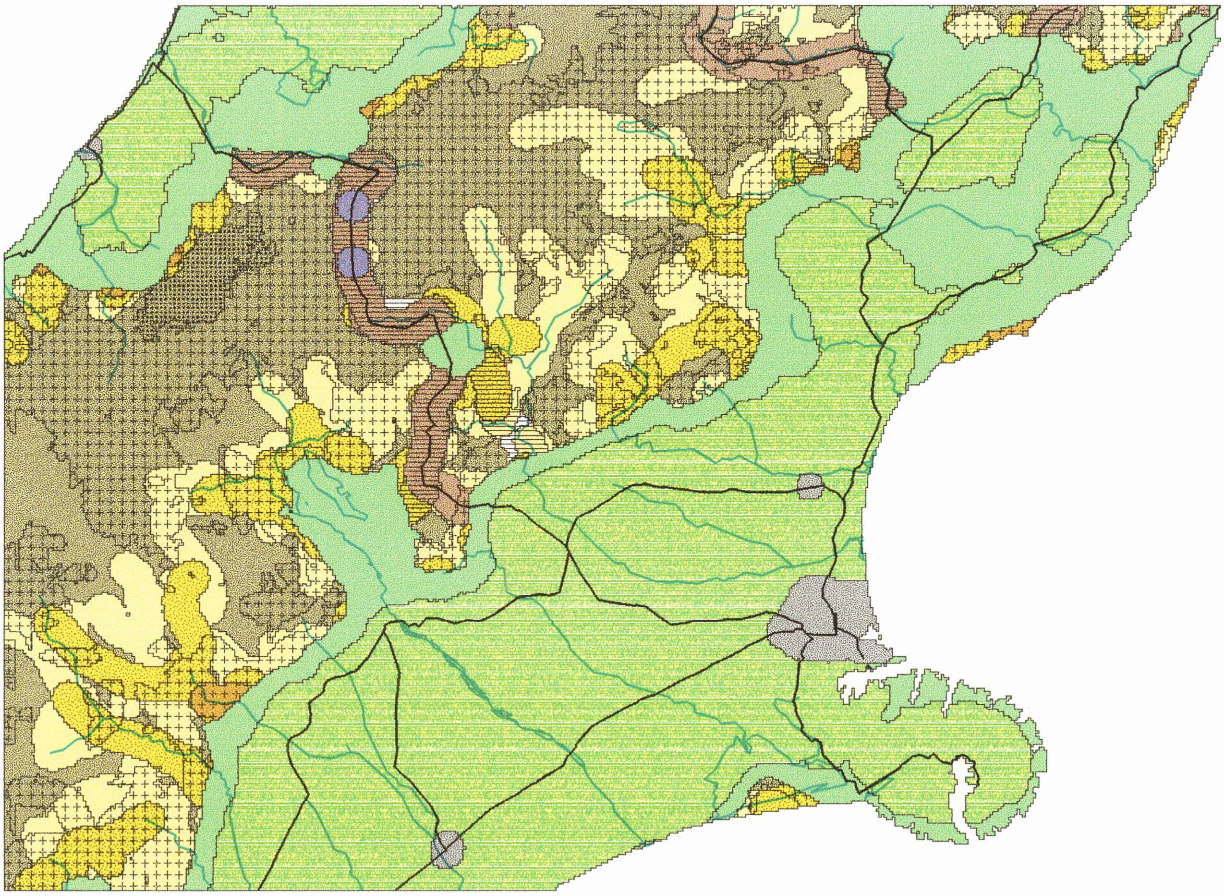
2139 unique combinations resulting from the overlaying of all the different infrastructure influence coverages



Data source: DOSLI 1:250,000, DCW, and Supermap2	Cell size: 500m	Map ID. natcomponents

Figure 4.14 Naturalness Level 1

Main road and hydrology layers are added for geographical reference

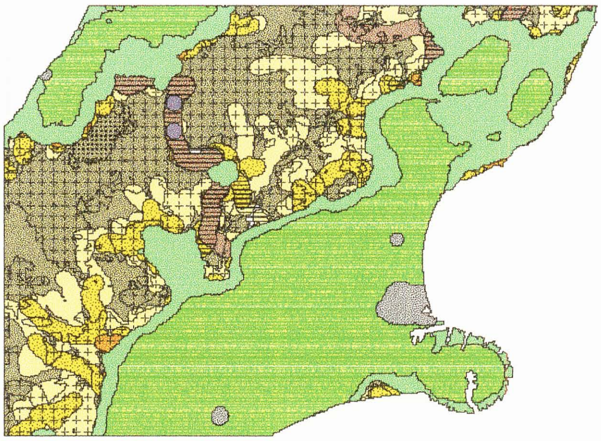


- Urban
- Developed Rural
- Rural
- Isolated Town
- Isolated Large Settlement
- Isolated Small Settlement
- Highway with Utilities
- Highway with Buildings
- Highway
- Utility with Sealed Secondary Road
- Utility with Unsealed Secondary Road
- Utility with 4WD. Track
- Utility
- Sealed Secondary Road with Buildings
- Unsealed Secondary Road with Buildings
- 4WD. Track with Buildings
- Sealed Secondary Road
- Unsealed Secondary Road
- 4WD. Track
- Backcountry Many Buildings
- Backcountry Few Buildings
- Remote

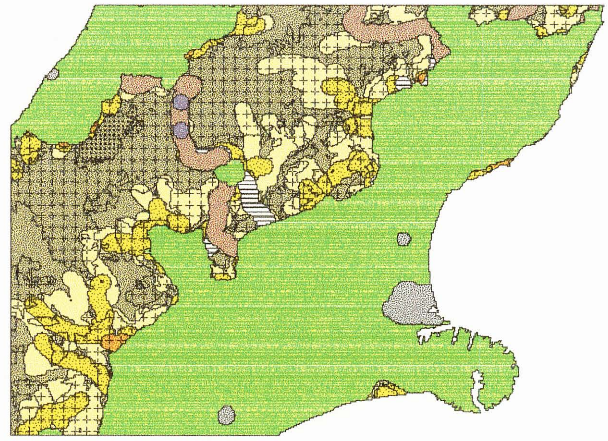
Data source: DOSLI 1:250,000, DCW, and Supermap2 Cell size: 500m NAW radius: 3000 - 10000m Map ID. natlev1

Figure 4.15 The Effects of Generalisation on Naturalness

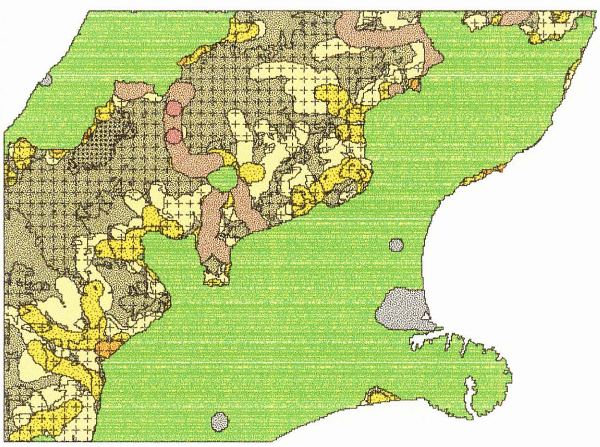
Level 1



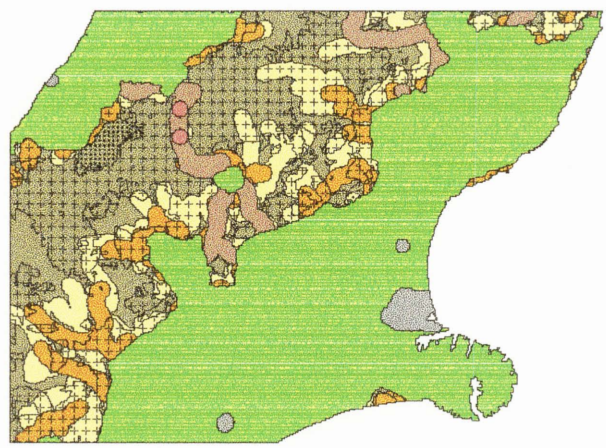
Level 2



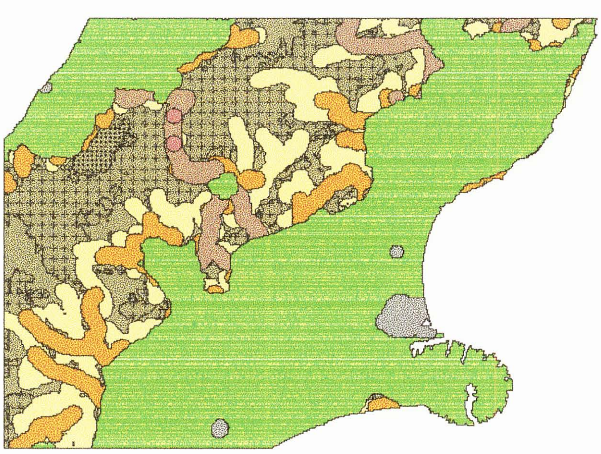
Level 3



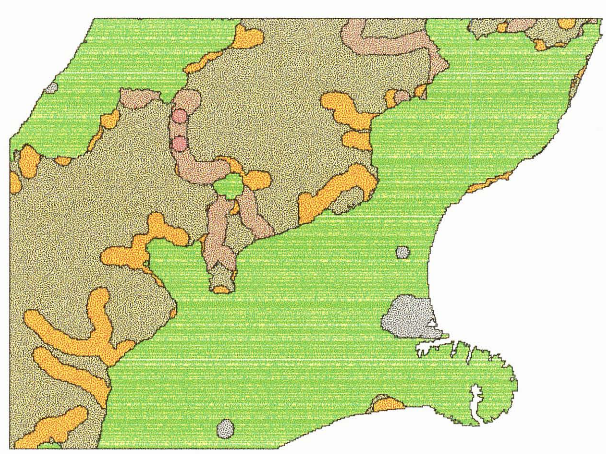
Level 4



Level 5



Level 6



Data source: DOSLI 1:250,000, DCW, and Supermap2. Cell size: 500m NAW radius: 3000 - 10000m Map ID. natlev1 - 6

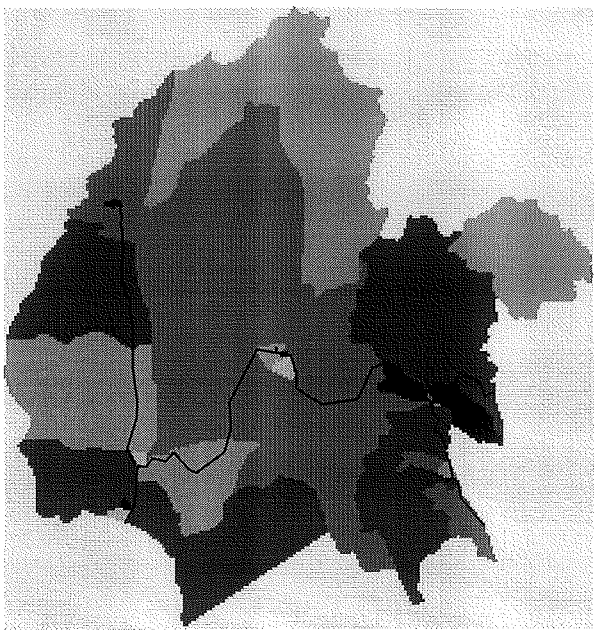
Figure 4.16 Supermap2 Population and Dwelling Data

Mackenzie District (South Canterbury)

Main roads are added for geographical reference

Population

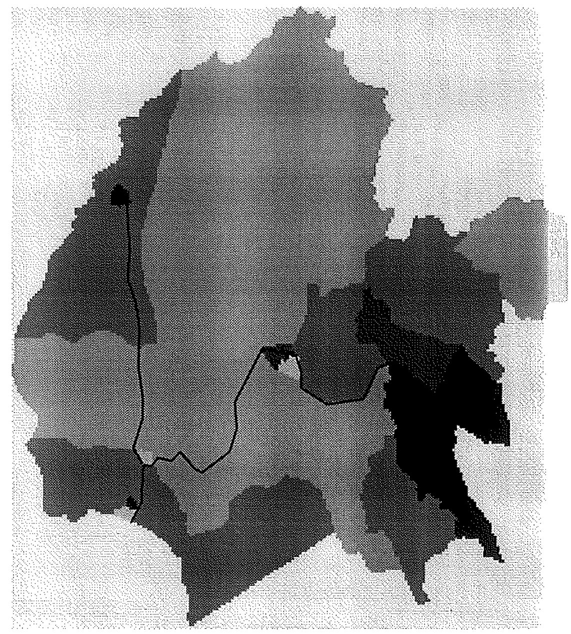
0	21 – 50	100+
1 – 20	51 – 100	



Population Density

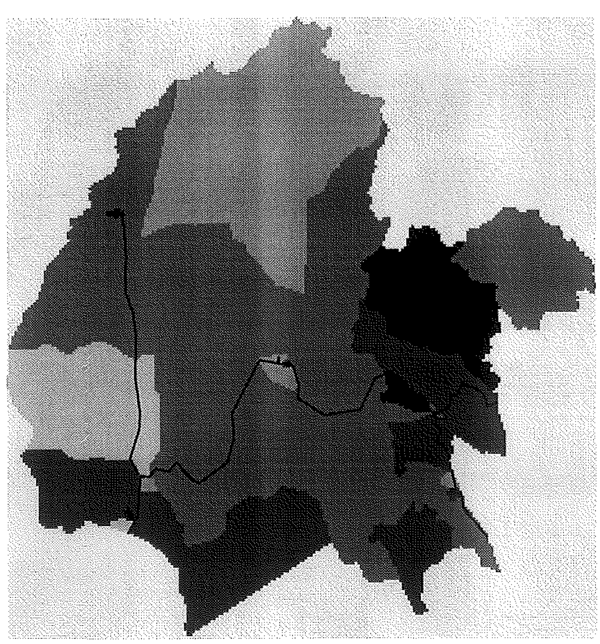
(Population / 10,000 hectares)

0	1 – 5	10+
0 – 1	5 – 10	



Number of Dwellings

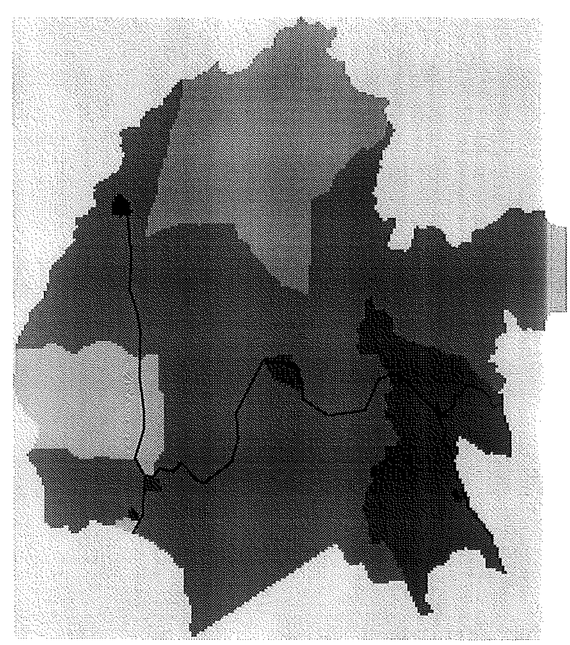
0	6 – 15	30+
1 – 5	16 – 30	



Dwelling Density

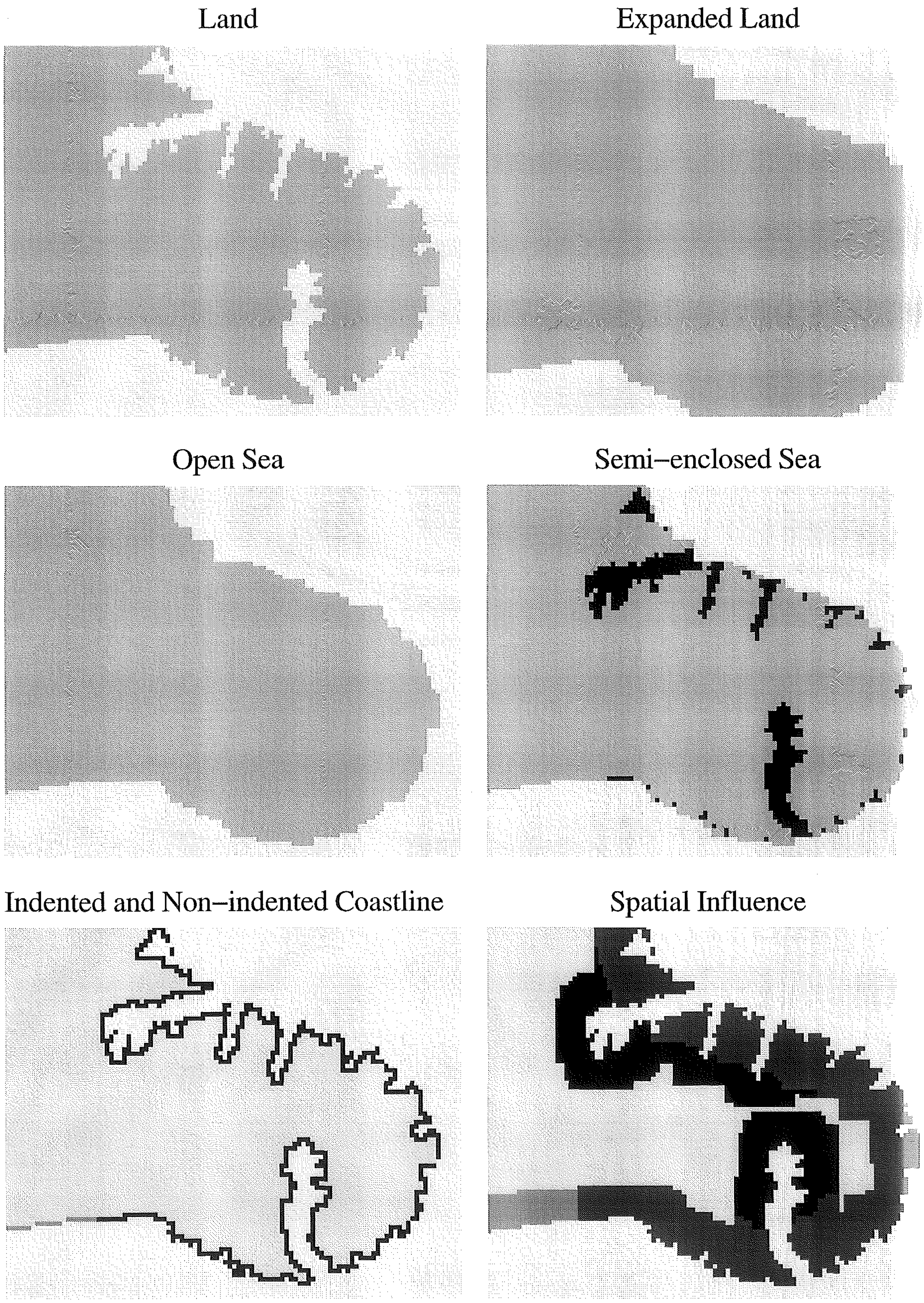
(Dwellings / 10,000 hectares)

0	0.1 – 1	10+
0 – 0.1	1 – 10	



Data Source: Supermap (Statistics N.Z.)	Cell Size: 500m	Map ID. mackpopdwel

Figure 4.17 Coastal Classification Process.

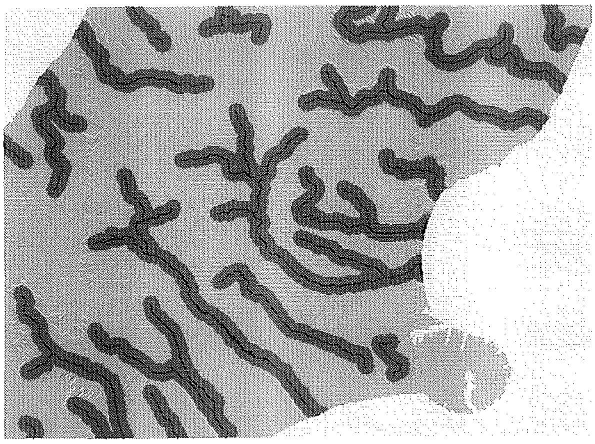


Data source: DOSLI 1:250,000 Cell size: 500m Map ID. watprocess

Figure 4.18 The Spatial Influence Of Different Water Components

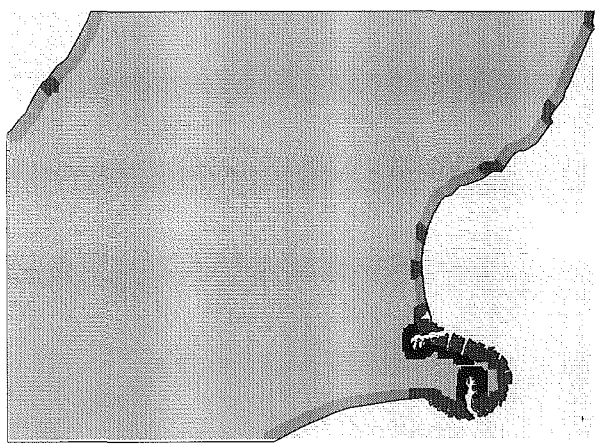
The water components that were used are outlined.

Rivers

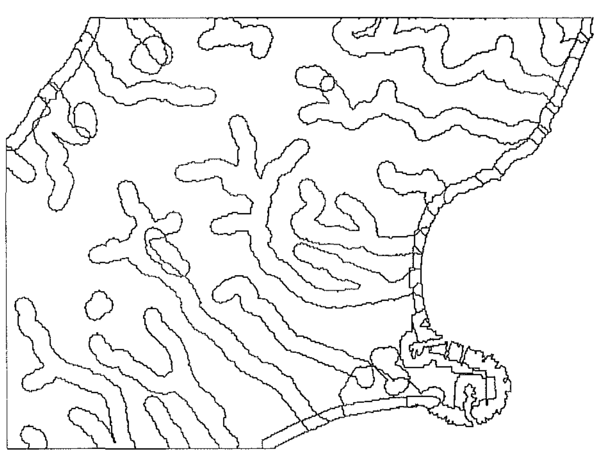


Large Lakes

Coast



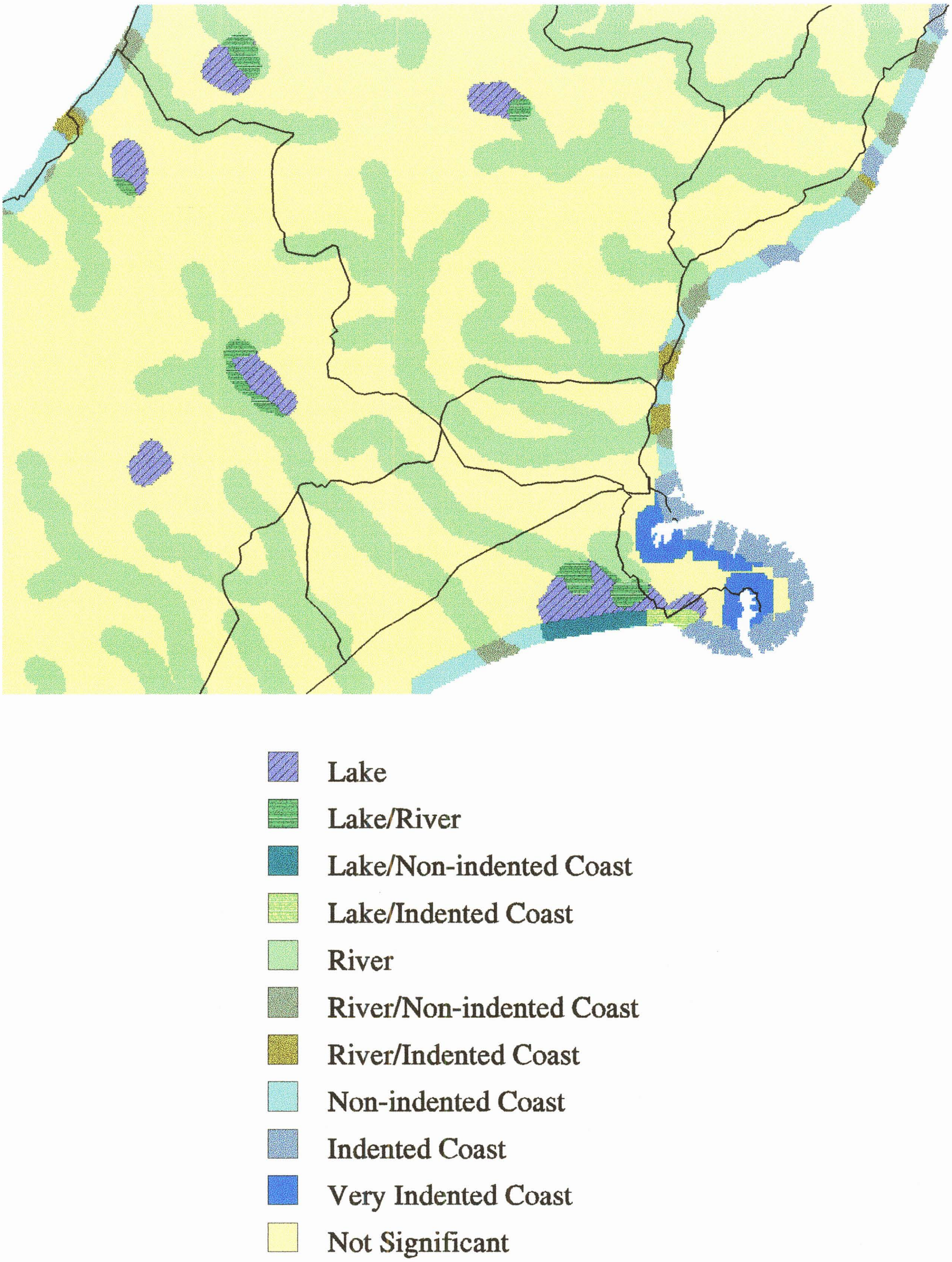
Unique Combinations Resulting From Overlaying
The Rivers, Lakes, and Coast Coverages.
(11 Unique Combinations)



Data source: DOSLI 1:250,000	Cell size: 500m	NAW radius: 3000m	Map ID. watcompco

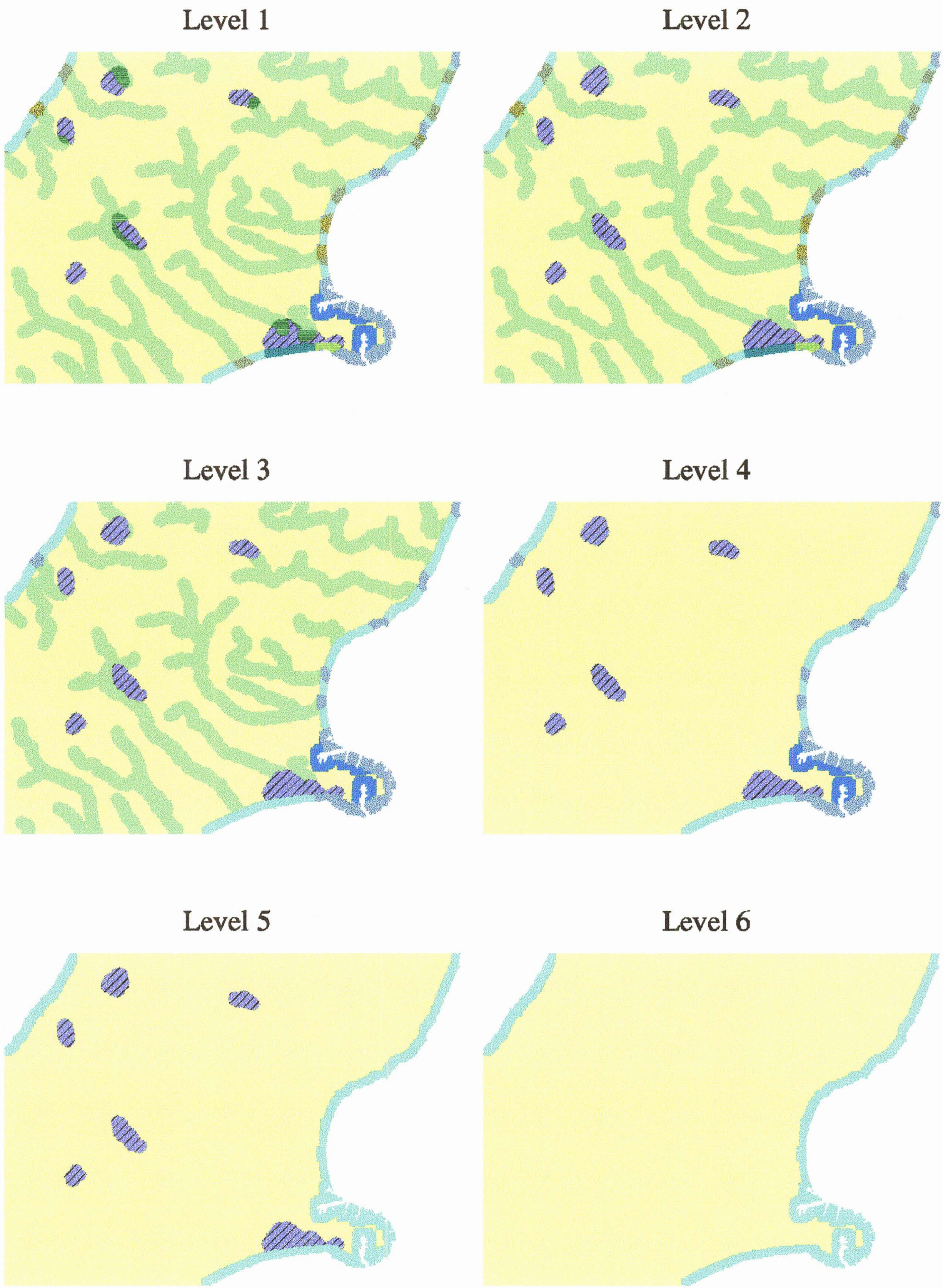
Figure 4.19 Influence of Water - Level 1

Main roads are added for geographical reference.



Data source: DOSLI 1:250,000	**Cell size: 500m**	**NAW radius: 3000 - 5000m**	**Map ID. watlev1**

Figure 4.20 The Effects of Generalisation on the Influence of Water



Data source: DOSLI 1:250,000	Cell size: 500m	NAW radius: 3000 - 5000m	Map ID. watlev1 - 6

Table 4.1 Generalisation of vegetation classes

<table>
<tr><th>LEVEL 1 (46 Classes)</th><th>LEVEL 2 (26 Classes)</th><th>LEVEL 3 (14)</th><th>LEVEL 4 (12)</th><th>LEVEL 5 (10)</th><th>LEVEL 6 (7)</th></tr>
<tr><td>Horticulture pasture</td><td rowspan="6">Horticulture</td><td rowspan="6">Horti-culture</td><td rowspan="6">Horti-culture</td><td rowspan="6">Horti-culture</td><td rowspan="9">Grass-land</td></tr>
<tr><td>Horticulture tussock</td></tr>
<tr><td>Horticulture scrub</td></tr>
<tr><td>Horticulture forest</td></tr>
<tr><td>Horticulture wetland</td></tr>
<tr><td>Horticulture sand dune</td></tr>
<tr><td>Pasture</td><td>Pasture</td><td>Pasture</td><td>Pasture</td><td rowspan="3">Grass-land</td></tr>
<tr><td>Tussock</td><td>Tussock</td><td>Tussock</td><td>Tussock</td></tr>
<tr><td>Tussock pasture</td><td>Tus.past.</td><td>Tus.past.</td><td>Tus.past.</td></tr>
<tr><td>Pasture ind.scrub</td><td rowspan="2">Grassland/ scrub</td><td rowspan="2">Grassland scrub</td><td rowspan="6">Scrub</td><td rowspan="6">Scrub</td><td rowspan="6">Scrub</td></tr>
<tr><td>Tussock grassland scrub</td></tr>
<tr><td>Lowland ind.scrub</td><td rowspan="2">Ind.scrub</td><td rowspan="2">Ind. scrub</td></tr>
<tr><td>Alpine scrub</td></tr>
<tr><td>Exotic scrub</td><td rowspan="2">Ex.scrub</td><td rowspan="2">Ex.scrub</td></tr>
<tr><td>Exotic scrub pasture</td></tr>
<tr><td>Ind.forest pasture</td><td>Ind.for./ past.</td><td rowspan="7">Ind.forest</td><td rowspan="7">Ind.forest</td><td rowspan="7">Ind. forest</td><td rowspan="16">Forest</td></tr>
<tr><td>Ind.forest tussock (scrub)</td><td>Ind.for./ tus.(scrub)</td></tr>
<tr><td>Ind.forest ind.scrub</td><td>Ind.for./ ind.scrub</td></tr>
<tr><td>Ind.forest ex.scrub</td><td>Ind.for./ ex.scrub</td></tr>
<tr><td>Ind.forest alpine scrub</td><td>Ind.for./ alpine scrub</td></tr>
<tr><td>Ind.forest scrub pasture</td><td>Ind.for./ scrub past.</td></tr>
<tr><td>Indigenous forest</td><td>Ind.forest</td></tr>
<tr><td>Ex.forest pasture</td><td>Ex.for. past.</td><td rowspan="5">Exotic forest</td><td rowspan="5">Exotic forest</td><td rowspan="5">Exotic forest</td></tr>
<tr><td>Ex.forest tussock (scrub)</td><td>Ex.for./ tus.(scrub)</td></tr>
<tr><td>Ex.forest scrub</td><td>Ex.for./scrub</td></tr>
<tr><td>Ex.forest scrub past.</td><td>Ex.for./scrub/ past.</td></tr>
<tr><td>Exotic forest</td><td>Exotic forest</td></tr>
<tr><td>Mixed forest</td><td>Mixed forest</td><td rowspan="4">Mixed forest</td><td rowspan="4">Mixed forest</td><td rowspan="4">Mixed forest</td></tr>
<tr><td>Mixed forest grass</td><td>Mixed for./ grass.</td></tr>
<tr><td>Mixed forest scrub</td><td rowspan="2">Mixed for./ scrub</td></tr>
<tr><td>Mixed forest grass scrub</td></tr>
<tr><td>Alpine (herb fields, rock, & ice)</td><td rowspan="2">Alpine</td><td rowspan="2">Alpine</td><td rowspan="2">Alpine</td><td rowspan="2">Alpine</td><td rowspan="2">Alpine</td></tr>
<tr><td>Tussock alpine scrub</td></tr>
<tr><td>Wetland</td><td rowspan="6">Wetland</td><td rowspan="6">Wetland</td><td rowspan="6">Wetland</td><td rowspan="6">Wetland</td><td rowspan="6">Wetland</td></tr>
<tr><td>Wetland grassland</td></tr>
<tr><td>Wetland Ind.scrub</td></tr>
<tr><td>Wetland Ex.scrub</td></tr>
<tr><td>Wetland Ex.forest</td></tr>
<tr><td>Wetland Ind.forest</td></tr>
<tr><td>Sand dune</td><td rowspan="6">Sand dune</td><td rowspan="6">Sand dune</td><td rowspan="6">Sand dune</td><td rowspan="6">Sand dune</td><td rowspan="6">Sand dune</td></tr>
<tr><td>Sand dune grassland</td></tr>
<tr><td>Sand dune Ind.scrub</td></tr>
<tr><td>Sand dune Ex.scrub</td></tr>
<tr><td>Sand dune Ex.forest</td></tr>
<tr><td>Sand dune Ind.forest</td></tr>
<tr><td>Not significant (Urban areas, rivers, & lakes)</td><td>Not significant</td><td>Not sig.</td><td>Not sig.</td><td>Not sig.</td><td>Not sig.</td></tr>
</table>

Level 1 (22 classes)	Level 2 (15 classes)	Level 3 (13 classes)	Level 4 (11 classes)	Level 5 (9 classes)	Level 6 (6 classes)
Urban	Urban	Urban	Urban	Urban	Urban
Developed rural	Rural	Rural	Rural	Rural	Rural
Rural					
Town					
Large settlement	Large settlement	Settlement	Settlement	Settlement	Settlement
Small settlement	Small settlement				
Highway/utilities	Highway	Highway/utility	Highway/utility	Highway/utility	Highway/utility
Highway/buildings					
Highway					
Utility/sealed secondary road	Utility				
Utility/unsealed secondary road					
Utility/4WD.track					
Utility					
Sealed secondary road/buildings	Sealed secondary road/buildings	Sealed secondary road/buildings	Secondary road/buildings	Secondary road	Secondary road
Unsealed secondary road/buildings	Unsealed secondary road/buildings	Unsealed secondary road/buildings			
Sealed secondary road	Sealed sec. road	Sealed sec. road	Secondary road		
Unsealed secondary road	Unsealed sec. road	Unsealed sec. road			
4WD.track/buildings	4WD.track/buildings	4WD.track/buildings	4WD.track/buildings	4WD.track	Remote
4WD.track	4WD.track	4WD.track	4WD.track		
Backcountry/ many buildings	Backcountry/ many buildings	Backcountry/ many buildings	Backcountry/many buildings	Back./many buildings	
Backcountry/ few buildings	Backcountry/ few buildings	Backcountry/ few buildings	Backcountry/few buildings	Back./few buildings	
Remote	Remote	Remote	Remote	Remote	

Table 4.2 Generalisation of naturalness classes

<table>
<tr><th>Level 1
(16 classes)</th><th>Level 2
(12 classes)</th><th>Level 3
(6 classes)</th><th>Level 4
(5 classes)</th><th>Level 5
(3 classes)</th><th>Level 6
(2 classes)</th></tr>
<tr><td>Non-indented coast</td><td>Non-indented coast</td><td rowspan="4">Non-indented coast</td><td rowspan="4">Non-indented coast</td><td rowspan="12">Coastal</td><td rowspan="12">Coastal</td></tr>
<tr><td>Non-indented coast/lake</td><td rowspan="2">Non-indented coast/lake</td></tr>
<tr><td>Non-indented coast/lake/river</td></tr>
<tr><td>Non-indented coast/river</td><td>Non-indented coast/river</td></tr>
<tr><td>Indented coast</td><td>Indented coast</td><td rowspan="4">Indented coast</td><td rowspan="4">Indented coast</td></tr>
<tr><td>Indented coast/lake</td><td rowspan="2">Indented coast/lake</td></tr>
<tr><td>Indented coast/lake/river</td></tr>
<tr><td>Indented coast/river</td><td>Indented coast/river</td></tr>
<tr><td>Very indented coast</td><td>Very indented coast</td><td rowspan="4">Very indented coast</td><td rowspan="4">Very indented coast</td></tr>
<tr><td>Very indented coast/lake</td><td rowspan="2">Very indented coast/lake</td></tr>
<tr><td>Very indented coast/lake/river</td></tr>
<tr><td>Very indented coast/river</td><td>Very indented coast/river</td></tr>
<tr><td>Lake</td><td rowspan="2">Lake</td><td rowspan="2">Lake</td><td rowspan="2">Lake</td><td rowspan="2">Lake</td><td rowspan="4">Not significant</td></tr>
<tr><td>Lake/river</td></tr>
<tr><td>River</td><td>River</td><td>River</td><td rowspan="2">Not significant</td><td rowspan="2">Not significant</td></tr>
<tr><td>Not significant</td><td>Not significant</td><td>Not significant</td></tr>
</table>

Table 4.3 Generalisation of Water Classes

CHAPTER 5

AUTOMATED CLASSIFICATION OF LANDFORM

5.1 Introduction

This chapter investigates the automatic classification of macro morphological landforms using GIS and digital elevation models (DEM). In the past, manual methods have been used for classifying macro morphological landforms from contour maps. Hammond's (1954 and 1964) procedure has to a certain extent become the de facto standard. A process developed by Dikau et al. (1991), which automates Hammond's manual procedures using GIS, is applied to the study area. Although this produces a classification that has good resemblance to the landforms in the area, it has some problems. A new process is presented that partly solves these problems. Landform classification is very sensitive to the operational definition used and this will be demonstrated. An application of fuzzy set theory that uses the notion of entropy is used to present this sensitivity.

For landscape classification, landform should be classified by morphology rather than rock type, structure, age or origin. It is usually the morphology that gives the greatest visual impression to the general public. Usually the rock type or structure is not even seen from a reasonable distance as the land may be covered by trees or buildings. Landscape assessment is concerned with the present character rather than the genesis. Genetic concepts are useful for understanding the processes forming the landforms but do not necessarily describe the appearance of a landform. The aims of a visual landscape classification are different from those of a genetic geomorphological classification, and therefore a different approach is required.

Within the fields of geomorphology and hydrology, the automatic mapping of morphological landforms has been of interest, for instance in modelling erosion (Dikau et al., 1991), providing watershed information (Band, 1986), and mapping land components (Dymond et al., 1995). A morphological landform classification has long been of interest to climatologist for developing climate models - topoclimatology (Geiger, 1971). Although these disciplines have a different purpose for landform information compared to landscape research, the ideas and methods initiated are very useful. In general, geomorphological classifications are based at the meso-relief, micro-relief and nano-relief levels, while landscape classification needs to incorporate macro-relief, and some elements of meso-relief (Linton, 1970). Dikau (1989) defines the macro landform scale to be landform greater than 10 square km and less than 1000 square km in area.

5.2 Manual classification

Hammond (1954 and 1964) has developed a macro morphological landform classification that was applied to the whole of North and South America. Wallace (1955) used Hammond's classification, with a few modifications, to classify New Zealand's landforms. Hammond's classification is very quantitative with clear, explicit definitions that can be easily applied by other researchers. It is perhaps this quality that explains why Hammond's classification has been so widely applied. The classification scheme used by Hammond is presented in Figure 5.1. A combination of three important parameters was used to identify different landforms. These were relative (local) relief, slope, and profile type. Relative relief is the maximum difference in height over a certain area. Hammond used a square grid measuring 9.65km (6 miles) across to determine the search area. After experimenting with different grid sizes, Hammond (1964) found that this size was

> "neither too small as to cut individual slopes in two and thus distort the determination of local relief, nor so large as to include areas of excessive diversity" (p.17).

Gentle slope is used to distinguish areas of relief and non relief. He chose 8 percent inclination as the upper limit of gentle slope, justifying this value by saying that it,

> "falls within the range of inclination in which the difficulty of machine cultivation increases rapidly, erosion of cultivated fields becomes troublesome, easy movement of vehicles becomes impeded, and in general one becomes highly conscious that he [sic] has a sloping surface to deal with" (p.17).

He also noted that the Soil Conservation Service in the U.S had used this threshold. However, the method used to identify this critical gradient is not explained by Hammond. As discussed in section 5.4.4, this is an elusive parameter to define. Profile type is explained in more detail in section 5.3.1. It is a means for expressing whether flat areas are above or below the surrounding terrain and so is used for identifying tablelands.

Subsequent to Hammond's work other landform classification schemes have been developed. Many are an adaptation of Hammond's work and Table 5.1 summarizes three of these.

Wallace (1955) has produced the only morphological classification of landforms for the entire of New Zealand (refer to Figure 5.2). A 1:1,000,000 base map was used and this was completed nearly forty years ago. As previously mentioned, Wallace used a method based on Hammond's scheme. Wallace (1955) remarked regarding future developments that he

> "earnestly hoped that others with more advanced concepts and better databases will work on a larger scale and reveal the inadequacies of this early effort" (p. 27).

Wallace did not explicitly calculate slopes because this would have been too laborious. Today, such slope information is easily available because automatic extraction of information from digital databases has advanced considerably. These data would have probably been beyond Wallace's most wild hopes. Despite these advances, which will be discussed and demonstrated in this chapter, there has been very little further development in New Zealand with this type of morphological classification since his attempt. This study will try to fulfil Wallace's hope.

The only other real initiative or discussion on morphological landform classification in New Zealand since Wallace's effort has been in response to the Protected Natural Areas (PNA) programme (Myers et al., 1987). The PNA programme was instigated to satisfy the requirements of the Reserves Act (1977) which established provisions for

> "...the preservation of representable samples of all classes of ecosystems and landscape...".

A discussion on landform classification resulting from this produced two papers: "Terrain evaluation for rapid ecological survey" (Crozier and Owen, 1983); and "A landform classification for PNA surveys in Southern Alps" (Whitehouse, Basher, and Tonkin, 1990). It appears that the main emphasis of the PNA survey was the protection of ecosystems and, in particular, significant representations of natural flora. As a result, there was no deliberation over visual landscape assessment theory. Crozier and Owen's classification scheme is based on the work of Wallace, which in turn can be traced back to the work of Hammond. The classification scheme devised by Whitehouse (et al.) appears to have been the adopted scheme used in the PNA program for the Southern Alps. This was genetically based which means that landform data collected for the PNA program is not the most appropriate for a visual landscape classification. Landform data from the PNA program is also difficult to use because most of it is not in digital format, and also the definitions of the different landform classes are not precise enough. For example, "valley floor" is defined as, "the comparatively broad, flat bottom of a valley". How broad is broad? With several different field teams, there could be inconsistency between different areas.

There have been many publications that describe New Zealand's landforms from a genetic perspective. A recent notable example is Soons and Selby (1982) but this does not help much for the development of a landform classification that needs to be morphological.

5.3 Automated classification

Computers have been used for extracting terrain parameters from DEMs for at least the last twenty years. Collins (1975) discussed different algorithms that could be used for identifying features such as tops of hills, bottoms of depressions, watershed or depression boundaries and areas, storage potential of watersheds, slope, and aspect. With the development of commercial GIS and national digital databases (NDDB) in the mid 1980s, there has been a resurgence of interest in this field (Dikau, 1989, Weibel, 1988, Weibel and Heller, 1991, Dikau et al., 1991, and Moore et al., 1993). Significant advances have been made, and many processes for identifying these parameters are now becoming standard functions within a GIS. Functions have been developed for generalising extensive terrain surfaces using triangulated irregular networks (TIN) (Midtbo, 1992). TIN and other algorithms have been used for generating DEMs from contours (Weibel and Heller, 1993), and slope can be obtained easily from either a TIN or a DEM. It is not the intention of this thesis to discuss in detail the mechanics of these functions as many general GIS books do this (eg. Aronoff, 1991). What is of interest in this thesis is how these parameters can be used to identify different landforms.

Regarding landscape research, there have only been a few published works on automatic landform classification. Barbanente et al.(1992) developed routines for identifying ravines and cliffs automatically. These are not features that can be justifiably included in a landscape classification because of the need to generalise. Jackson (1990) used GIS to identify certain terrain parameters using what are now fairly well known GIS functions. It is necessary now to determine more complex parameters and how these parameters can be used for identifying landforms.

The identification of parameters (parameterization) is an important first step in identifying landforms. These parameters are then used to develop parametric signatures of different landforms (described as formalisation). Dikau (1989) used this approach to identify plateaux, convex scarps, straight front slopes, concave foot-slopes, scarp forelands, cuesta scarps, valleys and small drainage ways, and crests.

Many of these landform features are, however, at the nano-meso scale, which is too detailed for a landscape classification that requires macro scale landforms.

Dikau, Brabb, and Mark (1991), in a very obscure publication, developed automated routines that do identify macro landforms. The process they developed automates Hammond's manual process nearly exactly and produces a similar result, which they demonstrated on the landforms of the entire state of New Mexico in the United States. Given that Hammond's classification has, to a certain extent, become the standard approach for a morphological landform classification, this is a significant development. In any classification, standardisation is important. The automated process developed by Dikau et al. is therefore of particular relevance to this thesis and will be discussed in detail.

5.3.1 Automating Hammond's classification scheme

Table 5.2 compares Hammond's scheme with the automated scheme developed by Dikau et al. The main difference between the two approaches is the number of classes identified and the method of generalization. The combination of parameter classes that Hammond's classification identifies could provide as many as 96 landform units, but it only identifies the more common landform units, which totalled 45. Perhaps this was required for practical reasons. The automated approach identifies all 96 landform units. Hammond's process also merged areas smaller than 2072 square kilometres into adjacent areas, so that the information could be generalized on to a 1:5,000,000 scale map. The automated approach does not do this.

Another difference concerns the use of spatial averaging windows. While a similar size square window was used by Dikau et al. (9.8 km sides compared to Hammond's 9.65 km), the averaging procedure was different. Hammond's approach moves the window along in 9.65km steps. This means that all the area within the window is generalised to one landform type. With the automated approach a neighbourhood function is used, as described in section 3.2.1.1, and its window moves in 200m steps, where 200m is the raster cell length. For each step, a generalization of the window was calculated and this information was assigned to the focal cell (the cell in the centre of the window). With Hammond's scheme, areas near the edge of the window boundary could be easily generalised wrongly as information outside the window boundary could be important to these areas but would not have been considered. This problem is partly solved with the automated approach using a neighbourhood focal function.

The basic procedures used in the automated approach developed by Dikau et al. are described in Table 5.3. It identifies the three components required - slope, relative relief, and profile type. Slope was calculated using a three by three moving window on a DEM, and from each placement of the window, the nine adjacent elevation points were used. Relative relief was calculated using a 49 by 49 moving window on a DEM (200m cell size). For each window placement, the difference between maximum and minimum elevation was used as the measure of relative relief. Figure 5.3 illustrates how the profile type was identified. As mentioned previously, profile type is used to determine whether the flat areas are above or below the surrounding terrain and is used principally for identifying tablelands. Three classes are distinguished: lowland gentle sloping, upland gentle sloping, and not gentle sloping. Upland and lowland profiles are identified by first calculating the maximum elevation within the moving window. The height of the central grid cell is subtracted from this. If this is less than half of the relative relief within the moving window, then the central cell is identified as upland. Otherwise, the central cell is lowland. The resulting upland and lowland coverage is then overlaid with a slope coverage to identify upland and lowland gentle sloping areas. The percentage of gentle sloping areas that are in lowland profiles is then calculated using a focal neighbourhood function.

Once these three components have been identified and classified, unique combinations are found by overlaying them. These are listed in Table 5.4, where the codes are the same as used in Hammond's scheme (refer to Figure 5.1). The subclasses are labelled using a capital letter, a number, and a small letter. These represent the different components used for identifying the subclasses. The capital letters from *A* to *D* represent different slope classes, the numbers from *1* to *6* represent different relative relief classes, and the small letters from *a* to *d* represent the different profile classes. The combinations of the different classes identify the 96 different subclasses. Once the subclasses are identified, the landform classes and types are determined by grouping the subclasses as shown in Table 5.4.

The database used by Dikau et al. for classifying the landforms of New Mexico was a 100m grid DEM. This was used to generate a 200m grid DEM. The software they used was a grid modelling system, an image processing system, and ARC/INFO. The hardware they used was a Sun Sparc 2, Vax 4000, Microvax II, and Prime.

5.3.2 Automated classification of New Zealand's landforms

Given that Hammond's landform classification scheme is reasonably well recognised and accepted, and also given that this scheme has been previously automated, it was decided that an automated process based on Hammond's scheme should be investigated for classifying New Zealand's landforms. ARC/INFO, a Sun Sparc 10 workstation, and a 100m contour database with spot heights were used. The contour database was converted to a 200m grid DEM using ARC/INFO's TIN, and TIN to grid functions. The process was thereafter similar to that developed by Dikau et al. (1991). A range of neighbourhood functions, as discussed in section 3.2.1.1 were used, as well as, a slope function within the GRID module of ARC/INFO, and a classify function (CLASS). The same class intervals, codes and labels were used as in Dikau et al. (1991). Figure 5.4 shows the different stages of the process for the Banks Peninsula region. First a DEM was produced. From this, slope can be calculated, which was then classed as less than or greater than (and equal to) 8 percent. The "mean slope" was calculated by assigning the value 100 to areas that were gentle sloping (< 8%) and the value zero where it was not. A focal mean function with a NAW of 5600m was then used to calculate the percentage of the neighbouring area that was gentle sloping. These percentages, classed into intervals, define the "mean slope" component. Relative relief was calculated from the DEM using a focal range function and a NAW of 5600m. A circular pattern results because of the influence of high points that affect the whole of the circular NAW. The relative relief values were classed into six intervals. Profile was calculated from the DEM by using a focal maximum function, and relative relief to identify upland and lowland profiles. This was then combined with the slope classes to identify the three profile classes. The profile component is represented by "profile percent" classes, which describe the percentage of gentle sloping areas that are in lowland profiles. The spatial averaging procedure used to accomplish this was as follows. A focal sum function counts the number of cells in the neighbourhood that were gentle sloping, and also the number of cells classed as lowland gentle sloping. From these values, the percentage of gentle slope areas that are lowland can be calculated. Figure 5.5 shows the resulting landform classes for the study area. The processing time was about two hours.

One difference between the process developed in this study and that developed by Dikau et al. was the shape of the NAW. Dikau et al. used a square window, while the process developed in this study uses a circle. A circle seems more appropriate than a square, for the obvious reason that the extent of the boundary of a circle will always be the same distance from the focal point, unlike a square. With the latest GIS technology it is easy to use a circle as a moving window. Perhaps it was not a viable option when Dikau et al. were developing their process. The radius used for the search window in this study was calculated to be 5529m in order for the area of the window to be the same as that used by Dikau et al. and Hammond. This radius is rounded to a multiple of the cell size, which with a 200m cell size becomes 5600m.

The automated process produces a classification (Figure 5.5) that has resemblance to the landforms of this area and is similar to Wallace's classification of the same area. It is difficult to quantitatively compare these two classifications since Wallace's (1955) classification is not available digitally. Wallace classifies virtually all of Banks Peninsula's landform as "low mountains". The automated approach identifies a significant proportion of Banks Peninsula as "low mountains" as well, but it also recognises that large parts of Banks Peninsula have flat areas, either as broad spurs on the far eastern parts of Banks Peninsula, or as valley floors. These flat areas have affected the classification and have resulted in a proportion of Banks Peninsula being identified as "open low mountains". The automated approach has also integrated plains and hills to generate a class that is a composition of these classes. As identified in the criteria given in section 2.9, composition is important for landscape classification.

The automated process, however, does have some problems. The first of these is the large regular shaped block in the Canterbury plains identified as "flat or nearly flat plains" in Figure 5.5. In reality there is no significant visual difference in landform between this area and the neighbouring areas on the Canterbury Plains. This area is the result of difficulties in producing an accurate TIN when the contours are far apart. Subsequently, this affects the slope calculation, which is important for distinguishing classes. This problem could be resolved if more contours or spot heights were added.

A second problem with the automated approach is the way classes change as the distance away from the areas of relief increases. For example, in Figure 5.5 the area between the Canterbury Plains and Banks Peninsula has a series of classes going from "plains" to "plains with hills" to "plains with high hills" to "plains with low mountains" to "low mountains". This reflects a progressive change in relative relief towards Banks Peninsula and is not a particularly desirable result. It is not how you would expect people to conceptualize the landforms in this area. As discussed above, it is desirable to have a composition class that incorporates the change from plains to mountains but this should not be done with progressive zonation.

A third problem with this automated approach is that some areas that are quite different in appearance are being classified the same. This is particularly the case with areas classified as "open" Some areas are "open" because they are at the interface between the plains and the mountains, while other areas are also "open" because they are in a broad valley, or on flat spurs. The process cannot distinguish between these different landforms. On the north eastern side of Banks Peninsula an area is classified as "open low mountains" and as previously noted this was because of the large flat spurs in this region. It does not seem appropriate that this area should be classified the same as areas that are at the interface between mountains and plains. The operational definition is unable to distinguish some objects that are of micro or meso scale, such as flat spurs, from objects that are of macro scale, such as plains. It is also for this reason that some areas are classified as "tablelands" when they are just ordinary hills.

Related to this scale issue is slope. Slope is very dependent on the scale at which it is measured, a matter that will become more apparent in section 5.4.3 when the effects of cell size are examined. This process uses the same slope criteria as Hammond (8 percent), but measures slope at a different scale, thereby, in effect, adopting a different slope criterion. It is necessary to determine whether this new slope criterion is appropriate. This issue regarding slope is discussed further in section 5.4.4.

If it was thought to be appropriate that conical volcanoes should be identified in the classification then this could in theory be included in an automated process. Dikau (1989) shows how concave and convex surfaces (in any direction) can be identified by using aspect and slope. It seems viable that conical shapes could be identified by their convex surfaces in the horizontal direction, and, possibly, concave surfaces in the vertical direction to develop a parametric signature of conical shaped volcanos. However, the issue is whether it is appropriate that volcanos are included in a landscape classification.

Although this automated classification has problems, it nevertheless has important advantages over manual processes. These are that it is totally explicit and that it can also be applied to large areas to produce results relatively quickly. This automated approach can also be viewed as just the start of a process that can evolve as better techniques develop. Because the process is explicit, one can analyse and improve on it.

5.4 Sensitivity to operational definition

The automated approach developed by Dikau et al. (1991) and then subsequently implemented in New Zealand is very dependent on critical thresholds specified for different parameters. For example, an eight percent slope threshold is used, and particular bounds are chosen for the component class intervals. The process also uses a neighbourhood analysis window that is defined by its radius. It would be interesting to know the effect of changing these values. With GIS and the use of macro programmes, it is possible to structure the process so that different thresholds can be easily changed. The macro used to run the landform classification process developed in this study contains variables for all parametric thresholds. These variables were then defined at the beginning by a separate sub-macro. As the processing time was only two hours it was possible to produce many different classifications that were the result of different parameter settings. Figure 5.6, Figure 5.7, and Figure 5.8 show, respectively, the effect of different slope thresholds, relative relief class intervals, and NAWs on the resulting landform classification (the relative relief class intervals are altered by dividing or multiplying the class bounds by the factors shown in Figure 5.7). The amount of agreement (ie. percentage of cells with the same class) between the classification that uses 2 percent slope and the classification that uses 14 percent slope is 21% for the Banks Peninsula area. The agreement between classifications with relative relief decreased by a factor of 4 and increased by a factor of 4 is 91%, and between a NAW of 1,000m and 10,000m radius is 43%. These figures show that the resulting classification is very dependent on how these parameters, especially slope and the NAW, are defined. However, the sensitivity to these parameters will depend on location.

The sensitivity analysis does not produce surprising results. The way the process is structured it is not surprising that if you change the definition of gentle sloping from being less than 2 percent slope to less than 14 percent slope, then there will be more "open mountains". By definition, in this classification process, for an area to be classified "open" it must contain a certain proportion of flat areas. By using 14 percent, then more areas will be identified as gentle sloping, and therefore more area will be identified as "open". The changes in relative relief levels have not affected the classification outcome substantially for the Banks Peninsula region, but it is easy to conceive that changes in relative relief classes could affect the outcome in certain locations where the topography is close to being either a mountain or a hill.

The effect of different NAW radii on the classification process is more complicated. It needs to be remembered that NAWs were used at many different stages of the process. It is used to calculate the percentage of area that is gentle sloping, the relative relief, and three times when calculating profile. The same size NAW was used for all these operations. The radius of the NAW will affect the boundary between areas of relief and no relief, subsequently the distinction between the classes "plains", and "plains with hills or mountains" changes with different radii. With relative relief, the larger the NAW then the more likely that the difference between the highest point and the lowest point will be greater. The size of the NAW also affects the amount of generalisation. When the NAW radius is only 1000m, the classification is more detailed than when the NAW radius is 10,000m. With a 1000m radius, micro relief is being identified, such as flat spots on the eastern spurs that have been identified as tablelands. As discussed previously, with landscape classification the identification of macro landforms rather than micro landforms is important. Small flat areas on spurs are not macro relief.

Figure 5.6, Figure 5.7, and Figure 5.8 show 21 different landform classifications of the same area. For each figure only one parameter has been altered and the others have been held constant. If the combinational effect of changing several parameters simultaneously was investigated, then virtually hundreds of different classifications would be produced.

5.4.1 A definitive classification

When Hammond produced his landform classification, it would not have been practical to investigate the effects of different operational definitions. It would have been important that the definitions of different landforms be chosen and only these are implemented, as this task would have been laborious enough. Now with GIS technology, one can see that it is possible to investigate different parameter thresholds. But it is still difficult to choose which operational definitions are appropriate as it depends on whose conceptual model is being considered. For example, a Dutch person will probably have a different definition of a mountain than a Nepalese. When viewing landforms, some people may focus on small areas, while others may view more widely and get an overall impression. As demonstrated, it is now possible to produce many different conceptual models of landforms, but having hundreds of classifications is of little use to research that needs a single frame of reference. A single classification needs to be decided upon.

One way of choosing an appropriate classification is to use the class that occurs most frequently (majority), for a given cell, from a wide range of different classifications that represent many different conceptual model. This can be easily implemented with GIS. The more advanced GIS software can do this with one command. Although hundreds of different conceptual models can be created, it seems that with ARC/INFO (version 6.2) only 47 coverages could be incorporated in the majority function. Figure 5.9 is the majority of 45 different classifications. The following parameter settings were used:

Five slope settings - 4, 6, 8, 10, and 12 percent;

Three relative relief settings - Hammond's,
Hammond's divided by 2, and
Hammond's multiplied by 2; and

Three NAW radii - 2400m, 5600m, and 8400m.

The combination of all these settings produces 45 different classification. It should be noted that when the majority function is used in ARC/INFO and there is no clear majority (ie. when two or more classes share the highest frequency) for a particular cell, then no value is assigned to that cell. For Banks Peninsula there were a few cells where this was the case, but where this happened the cell value from Hammond's parameter settings was used instead. It should also be noted that a cell size of 400m was used because of the amount of processing involved.

A majority classification could be used as a definitive classification because it incorporates a wide range of conceptual models. However, a majority classification is sensitive to the range of conceptual models chosen, and perhaps a different range is more desirable. With GIS this majority calculation is very quick, so different ranges of parameter setting could easily be experimented with. On the other hand it could also be argued that Hammond's classification should be the definitive classification as it has been in use since 1954 and has become a de facto standard.

5.4.2 An application of fuzzy set theory

As discussed in section 2.9, landscapes are fuzzy entities, as they are based on human conceptualization and this varies between different people. Fuzzy set theory provides a means of presenting this fuzziness by providing information that shows the degree of membership of different classes that exist for each cell. Using the example presented in the previous section, membership is calculated by comparing all the 45 different outcomes. For each class, a coverage is created that shows the degree of membership (frequency of occurrence) that exists for different cells. The membership of each class was calculated by first generating grid coverages that consisted of only the value for that class, for example a grid coverage that consisted only of 2 (2 corresponded to "tablelands"). An "equal to" function was used to count for each cell how many of the 45 different classifications equalled this blank coverage value. This provided information on the membership of that class. This process was repeated for all the classes. Figure 5.10 shows the results for the landform types. In this case there are only five possible classes so this information can be easily presented. When there are hundreds of different classes, which will be the case with a landscape classification that consists of the unique combination of four different attributes, then this information will not be easy to present and would in fact be too much for anyone to assimilate.

One way of presenting this membership information for easier assimilation is to use the notion of entropy (Wilson, 1970, Ashby, 1994). Entropy provides information on the distribution of the membership of the different classes for a given area (in this case a cell). It is implemented by first calculating for each class the proportion of the 45 outcomes that are assigned to that class. Thus if a particular cell is assigned to class A in 15 outcomes, the coverage for class A will show a value P of 0.33 for that cell, while coverages for the other classes will show P values totalling 0.67. The entropy coverage is then created by combining these P values with the formula for entropy (Eqn. 5.1). If the membership of one class is very high and the membership of all the other classes is low then entropy will be low. If the memberships of all the classes are fairly even and there is no class that stands out, then entropy will be high. Low entropy indicates a high degree of consensus between classifications, and a high entropy value means there is very little consensus between classifications.

The equation for entropy of a cell is:

$$Entropy = -\sum_{i=1}^{n} P_i . \ln(P_i)$$

$$where \sum_{i=1}^{n} P_i = 1 \qquad (5.1)$$

n = the number of different classes

P = the membership of each class

ln *is the natural* log

(*Wilson,* 1970)

The entropy calculated from the 45 different landform classifications generated for the Banks Peninsula area is shown in Figure 5.11.

The entropy values show that when the classes are general there is more agreement, but as the classes become more specific there is less agreement. It is interesting to speculate whether this reflects consensus in society. Are people more likely to agree that a particular landform is a mountain but less likely to agree whether the mountain is high or low?

Entropy appears useful for evaluating landscape classifications and their application. For instance, one use for a landscape classification is a frame of reference for psychophysical landscape assessment, as discussed in section 2.5.1. It would be appropriate if the photos for the public preference surveys were taken of areas where there is agreement over its classification. Entropy provides this information.

The entropy values calculated in Figure 5.11 are not specific to any one classification. They provide general information about a particular area. However, it is possible to provide consensus information that is specific to one classification. If a definitive classification is agreed upon (and perhaps this will be a majority classification) then it will be appropriate that consensus information is obtained that is specific to that classification. This can be done by again using the "equal to" function to count how many of the 45 classifications equal a suggested definition for each cell. If the majority classification, as shown in Figure 5.9, is accepted as the definitive classification then the amount of agreement between this and the 45 different landform classifications can be calculated. The result is shown in Figure 5.12. It can be argued that this approach (which will be now referred to as the agreement model) is better than the use of entropy. The agreement model is easier to understand and to implement within GIS. On the other hand, entropy does provide additional information about all the other possible classes that could be classified for a given area.

This application of fuzzy set theory is simpler than that used by Burrough (1989) and Burrough et al. (1992) for soil classification. Nevertheless, it is still an effective application. Burrough's et al. (1992) approach is more complex because it considers the probability of the different parameter settings that produce the possible outcomes, whereas in this study, the probability of the different parameter settings is assumed to be equal. This assumption is necessary because it is not known what the probability of the different settings should be. Perhaps some settings, such as 14 percent slope, are unlikely to agree with public perception, and this should be incorporated in the process by assigning this parameter setting a low probability. This application is simpler also because it uses simulation to determine membership rather than complex mathematical calculations. It should be remembered that the results from these fuzzy set theory applications, presented previously, do not express the statistical probability of a class. The results can only be used as a relative indication of the probability of different classes.

5.4.3 The effects of cell size on the classification process

The effects of using different cell sizes on the process were also investigated, and produced some interesting results. Figure 5.13 shows that different cell sizes have a significant effect on the resulting landform classification. Over the whole study area, the agreement between 200m and 500m cell size for the landform classes was 90%, although for Banks Peninsula it was only 61%. The reason for this effect of cell size was investigated by visualizing, for each cell size, the individual stages of the process. Figure 5.14 and Figure 5.15 show the process for 100m and 1000m cell sizes respectively. It is apparent that it is the variation in the slope classes that are causing most of the variation in the output. Figure 5.16, and Figure 5.17 show the effect of cell size on slope classes (70% agreement between 100m and 1000m cell size for Banks Peninsula), and "mean slope" (54% agreement between 100m and 1000m cell size for Banks Peninsula) respectively. The reason for this variation in slope with different cell sizes becomes apparent when the cells are examined in relation to the contours and TIN lines (Figure 5.18). With this automated process the DEM is produced from the TIN coverage. The DEM is then used to determine slope by using a neighbourhood function that compares the heights of the neighbouring cells and then calculates slope. From Figure 5.18, it is clear that as the cell size is increased the detail in the topography is being lost. With a 100m cell size, non macro topography is being identified, such as flat spots on spurs and ridge tops, and small steep sections. With the larger cell sizes, such topography is being lost and it even appears that detail at the macro scale is being lost as well. This difference is thus affecting the "mean slope" (Figure 5.17). This effect depends on the presence or absence of different scales of topography, and whether this topography consists of flat objects or steep objects. It illustrates the scale dependency of slope that Dymond and Harmsworth (1994), and Moore et al. (1993) have also illustrated.

5.4.4 Slope - the elusive parameter

Slope is a critical parameter for identifying landforms and is used in manual methods as well as in automated methods. Yet slope is difficult to objectively measure. To measure slope objectively using manual techniques in the field, usually requires that a scale be specified by choosing a particular slope length. Calculating the mean slope using a slope length of one metre will give a different result to using a slope length of one kilometre. It is also necessary to specify where these slope lengths begin and finish. For practical reasons, manual methods for calculating the mean slope of an area have not been explicit, and so it is difficult to automate these using GIS.

A comparison was made between GIS generated slope measurements and manual slope measurements for the whole of the study area. The LRI contains manually measured slope information classed into intervals for areal units. The LRI slope information was reclassed as flat if it contained a slope interval less than 12 percent, otherwise it was reclassed as non-flat. It was then stored as a 200m resolution GIS layer. For comparison, a GIS generated slope coverage was produced from a 200m cell size DEM. From this, a range of flat/non flat coverages were produced based the following thresholds: 1, 2, 4, 6, 8, and 12 percent. These were then compared with the classified LRI slope coverage, by calculating the amount of agreement (number of cells classified the same). The agreements for the different slope thresholds were as follows:

Slope	Percentage agreement
1	87
2	88
4	88
6	88

8	87
12	84

These agreement figures appear to be quite high but they actually reflect quite significant differences between manual and GIS slope measurements. The analysis was done on very general slope classes (just two classes) and these classes have a dramatic effect on the classification outcome. If two classifications were derived for the study area and they both used a 12 percent threshold but one was based on the GIS slope measurements and the other on the LRI data, then only 84% of the area in the classifications would be in agreement (ie. 16% would be different). This analysis shows that it is unwise to take slope thresholds based on manual measurement and use them in classifications based on GIS measurement. The GIS slope measurements used in this study and Dikau et al. (1991) are not flawed, they are just obtained differently.

If slope information from the LRI is used in the process then the "mean slope" is relatively stable with different cell sizes as shown in Figure 5.19. There is 98% agreement between 100m and 1000m cell size for Banks Peninsula. It is apparent from a comparison of Figure 5.17 and Figure 5.19 that using LRI slope information provides a more stable result in relation to cell size than using the DEM derived slope information. The slope information in the LRI is obtained from field measurements that are determined at a macro scale. This information is stored in a polygon coverage. Because these polygons are large, detail is not lost when these polygons are converted to grids, even with large cell sizes. The problem with using the LRI is that the slope information for each areal unit is given as an interval. If the terrain within the areal unit is variable then this slope interval may be large. There can also be more than one slope interval given for an areal unit. It can therefore be difficult to determine if the slope of an areal unit is above or below the slope criteria. With the LRI data it was assumed that an areal unit was "not flat" if it contained a slope interval that extended above the critical slope threshold of 8%, and because slope information is stored in intervals this resulted in a 12% threshold being used. It should be noted that the LRI may be inconsistent because of the difficulties in determining a totally explicit field method for calculating slope, and that not all countries have access to such databases.

As demonstrated in the previous section, the "mean slope" determined from DEMs changes considerably when the cell size is changed. How do we know what is the best cell size to use? Also, is it desirable to have a process that is dependent on a particular cell size? What happens if an accurate DEM with 200m cell size is not available? Alternative methods for automatically calculating slope were therefore investigated.

Instead of calculating slope from a DEM it is possible to derive slope from a TIN (based on the slope of the triangle facets), and then convert this slope information directly to a grid coverage. Figure 5.20 shows the effect of different cell sizes on slope obtained directly from a TIN. There is 53% agreement in slope classes between 100m and 1000m cell size for Banks Peninsula. There are some obvious differences with this figure compared to Figure 5.16 where slope is obtained directly from a DEM, especially with larger cell sizes. The slope calculated directly from TIN is still very sensitive to cell size because of the effects of micro topography. The TIN identifies micro relief objects but these are generalised when converted to a grid coverage. The degree of generalisation depends on what cell size used. The use of TIN therefore does not solve the problem.

Another alternative method for determining "mean slope" that reduces the effect of micro relief and is less sensitive to changes in cell size is to first remove small flat areas from the slope class grid before the "mean slope" is calculated (slope can be calculated from either a DEM or directly from a TIN). Small flat areas can easily be identified by their size. From the definition for macro landform size given by Dikau (1989), this threshold size should be 10 square kilometres. Once identified, these flat areas can be converted to non-flat areas. This approach is implemented in the following section.

5.5 A new automated landform classification process

As previously mentioned, Dikau et al.'s (1991) classification process has certain problems. These being that it produces a progressive zonation when landform changes from plains to relief, it does not distinguish open valleys from a plains-mountain interface, and it is affected by micro relief. A new process was therefore developed that partly solves these problems. This process was developed using a 500m cell size to ensure the processing time was not too great. It will be demonstrated that the outcome is not severely affected by cell size.

Figure 5.21 and Figure 5.22 show the different steps in the first phase of the process, which in summary produces three classifications of landform:

1) a set of six relief types,
2) a division of "flat" types into open valley and plain, and
3) identification of a special class of tableland within the "plain" type.

Starting with a DEM, a slope grid was derived just like Dikau et al.'s (1991) process, and this was classified according to slope. However, a 4 percent threshold was used instead of an 8% threshold to distinguish the low gradient cells. The reason for this is discussed later. Any small flat areas that were less than 10 square kilometres in size were then converted to non-flat areas to produce a "macro slope classes" grid. The next three steps identified open valleys. An open valley is a large flat area that has relief on opposite sides. This pattern was identified using an expand and shrink sequence (as used for identifying indented coastlines in the previous chapter). Areas identified as non-flat were expanded by 3000 metres (with a 500m cell size this corresponds to six cells), and then shrunk by 3000m. The effect of these two steps was that flat enclosed and semi-enclosed areas (open valleys) became non-flat. Open valleys were then identified by using a conditional statement on the "macro-slope classes" grid and the "shrunken" grid. That is, if a cell was flat in the "macro-slope classes" grid and was not in the "shrunken" grid then it was class as an open valley. For an area to remain classified as an open valley, it also had to be more than 10 square kilometres in size. A conditional statement was used for this.

Relative relief was determined by Dikau et al.'s (1991) process by using a focal range function. For areas that were previously identified as non-flat, the relative relief was classified into five classes to produce a relief type grid. The relief classes were:

0-150m - Low hills
150-600m - Hills
600-900m - High hills
900-1500m - Mountains
Above 1500m - High mountains

These relative relief classes are slightly different to those used by Dikau et al. They are intended to reflect how New Zealanders conceptualise terrain in New Zealand, although there is no substantive evidence to suggest how this is. The Banks Peninsula region is classified as high hills by Glasson (1991) in a visual assessment study. A relative relief interval of 600-900m achieves this. Two mountain classes are recognised, distinguishing the grander mountains, which often have permanent snow and bare rock, from the others. It should be noted that flat cells defined by gradient were maintained as flat areas even though some had high relative relief neighbourhoods.

Tablelands were identified from upland and lowland profiles and these profiles were identified in a similar way to Dikau et al.'s process. However, the actual identification of Tablelands was simpler than Dikau et al.'s because "profile percent" classes were not used. Instead, if an area was upland and flat in the macro-slope coverage, then it was identified as a tableland. No tablelands were identified in the whole region using this process.

A coverage that has the potential to identify eight morphological landform classes (five relief types, plains, open valley, and tableland) was then produced by overlaying the maps of relief types, open valleys, and tablelands. Figure 5.23 shows this for the whole study area. This landform components map cannot be used in a landscape classification in this form because it does not contain composition classes, but instead identifies the sharp boundaries between different landform types (eg. plains and mountains). However, it could be used for other purposes (eg. climate and hydrology modelling).

Once the landforms had been conceptualised, the second phase of the landform classification could commence. Landform compositions were identified in a similar way to that used for the landcover attributes. Each of the eight landform components were singled out into individual grids, with the value 100 assigned to cells where the particular component is present, and the value zero where it is not. A focal mean function, with a 3000m radius NAW, was then applied to each component grid, and these mean values were placed into one of four class intervals (the results are shown in Figure 5.24 and Figure 5.25). These eight spatial influence grids were then overlaid to produce a new grid that contained unique combinations of them (a vector representation is shown in Figure 5.25). Since eight grids were combined and each had the possibility of four different classes, then the combined grid had the possibility of 65,536 unique classes. However, there were only 613 unique combinations in the study area. Twenty two landform classes were then identified by querying this combined coverage. The classes are listed in Table 5.5 under level 1, and the definitions used to identify them are described in Appendix 4. The classes have been chosen because of their distinctiveness in form, and to a certain extent reflect the classes used by past classifications. Checks were made to ensure that the definitions were mutually exclusive and exhaustive as described in section 4.2.3. Not all these landforms existed in the study area. The resulting level 1 classification is shown in Figure 5.26.

In deriving a landform component map, several parameter thresholds had to be determined - 4 percent slope, a 6000m maximum valley width criteria, and as already discussed the various relative relief classes. A slope of 4 percent was used for distinguishing flat and non-flat areas. This differs from Hammond's 8 percent, which was also adopted by Dikau et al. (1991). As discussed in section 5.3.2, using DEMs to derive slope produces a different result compared to using field measurements. Therefore it is likely that a different slope threshold is needed with automation compared to Hammond's method. The effects of different slope thresholds were investigated by implementing the process with different slope thresholds (Figure 5.27). The amount of agreement between the use of a 1% slope threshold and an 8% threshold is 67%. With 8 percent, 7,528 more cells were classed as plains or open valleys than with 1 percent. The opposite occurred for the classes containing relief. Low hills and hills are virtually absent with 8 percent, and the non relief classes extend well into areas that can be regarded as relief.

A comparison was made between the resulting slope classes and the slope information in the LRI (similar to that shown in section 5.4.4 but this time using a 500m cell size). As previously discussed, the LRI slope information is based on areal units, slope is given in class intervals, and occasionally more than one interval is given to an areal unit. Despite this, it still provides the best available representation of slope for which a comparison can be made. A slope interval of 0-7 degrees (based on LRI intervals of 1-3 and 4-7) was used to represent flat areas. The 4 percent threshold produced a slope class grid that had the highest agreement with the LRI (91%). The slope threshold of 1 percent and 8 percent both had agreements of only 88%. Four percent therefore seems an appropriate threshold. Even when 4 percent was compared with the LRI slope interval of 1-3 degrees, the agreement was still high (90%). Although hills are not very well represented with a 4 percent threshold, it appears more suitably for identifying the extent of open valleys.

A 6000m maximum valley width threshold was decided upon by assessing the effects of different width criteria. Valley widths vary considerably and topographic maps show that these can be 5000m in the Rangitata catchment. To be sure all such valleys were identified, 6000m was decided upon (this was achieved by using an expand and shrink of 3000m). If the maximum valley width criterion is set too high then some large basins become identified as open valleys.

The landform classification can be easily generalised by grouping different classes. This was done to produce six different levels of generalisation. The way the different classes were grouped is shown in Table 5.5. Figure 5.28 shows graphically the effect of different levels of generalisation. No keys are provided with this figure to avoid cramming, but the colours are the same as used in Figure 5.26 and the keys can be ascertained by using this and Table 5.5. Like the rationale for the level 1 classes, the classes in levels 2-6 have been chosen because of their distinctiveness in form. At the more general levels this distinctiveness needs to be more apparent.

This new process produces a landform classification that does not have the same problems as that developed by Dikau et al. (1991). The interface between relief and plains is not identified as a progressive zonation, valley floors are distinguished, and micro relief does not alter significantly the outcome. Cell size, however, still affects the classification. There is 89% agreement between level 1 classifications based on 200m and 500m cell sizes. This is similar to the 90% found for Dikau et al.'s landform classes. However, for a comparison between this new process and Dikau et al.'s to be valid, it needs to be done at a similar level of generalisation. For level 3, which has a similar number of classes as Dikau et al.'s landform types, there is 93% agreement between 200 and 500m cell size. Cell size is still affecting the calculation of slope classes with this new process, despite the removal of small flat areas. Slope classes particularly affect the boundaries of large open valleys that gradually get steeper and therefore do not have a distinct boundary.

What this classification identifies as open valleys perhaps does not agree with how most people conceptualize valleys. The definition of an open valley as a large flat area that has non-flat areas on opposite sides, is perhaps too simple. People often associate rivers with valleys, so perhaps a river must be in the vicinity. This could be incorporated in the classification process. Another issue is that where there is an isolated hill surrounded by flat areas, the flat area between the hill and a nearby non-flat area becomes identified as a valley. This can be seen in 5.6 on the edge of the

Canterbury Plains. This is a problem with the process. One may also think that the maximum width of a valley should be determine by how high the surrounding relief is. For example, in the head of the Rangitata catchment the relief is very high, so although the flat areas are very wide (5 km), one still gets an impression of being in a valley. If the surrounding relief had been only low hills then this area perhaps would not be conceptualised as a valley. This problem could be solved with context dependent definitions that take the relative relief into account, but this makes the process more complicated.

As with the components discussed in chapter 4, the use of a 3000m search radius for determining the spatial influence of different components can also be questioned. There has been no cognitive research that can be used for determining what spatial influence different components of the landscape have on people's conceptualisation of the landscape. One could argue that this figure should not be constant for landforms. Some components, such as high mountains, have more spatial influence than other components, such as low hills. The use of context dependent search radii could also be incorporated into the process.

5.6 Summary

Automating landform classification is an interesting challenge. It produces classifications that have a good resemblance to manual methods, and because definitions are explicit they can be easily identified, questioned, and improved. This has been demonstrated with Dikau et al.'s (1991) process. Several problems were encountered when applying it to the study area: it produced a progressive zonation when landform changes from plains to mountains; it did not distinguish open valleys from a plains-mountain interface; and it was affected by micro relief. Also, the same slope threshold was used as Hammond's even though slope was measured differently. Although automating existing quantitative manual processes are important steps in the evolution of automation, definitions may need to be calibrated. This is the case with slope measurements. The effects of scale and generalisation also need special attention.

Dikau et al.'s (1991) process can be improved by adopting a 4% slope threshold, removing non macro relief, identifying open valleys using an expand/shrink sequence, using different relative relief classes, and by using spatial influence information of each component to identify landform compositions. A new process has been developed that adopts these improvements. There are opportunities for improving the process further with the use of more context dependent definitions, and the identification of particular distinctive landforms such as conical volcanos.

Figure 5.1 Hammond's classification scheme
Source: Dikau et al. (1991)

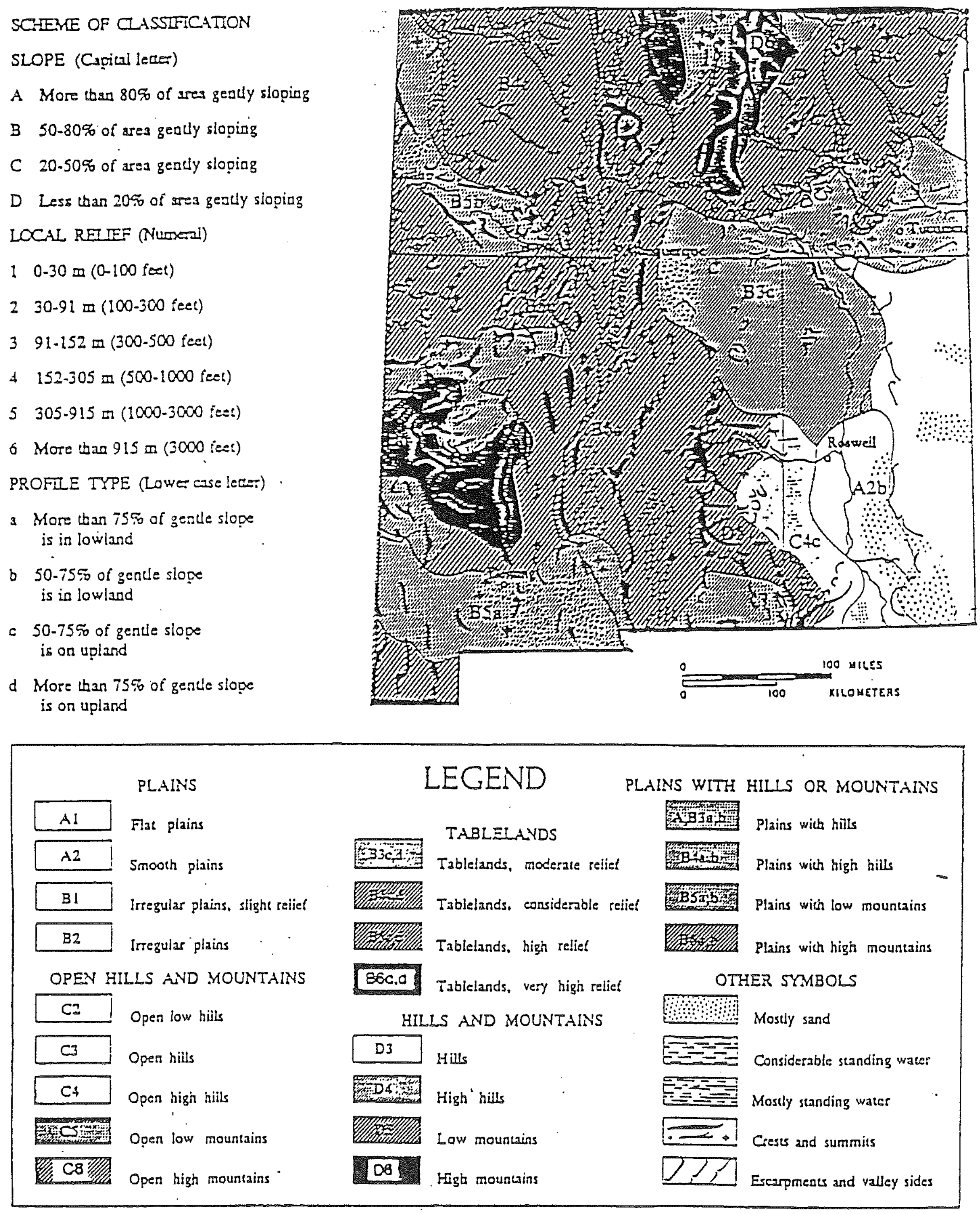


Figure 5.2 Wallace's landform classification of New Zealand
Source: Wallace (1955)

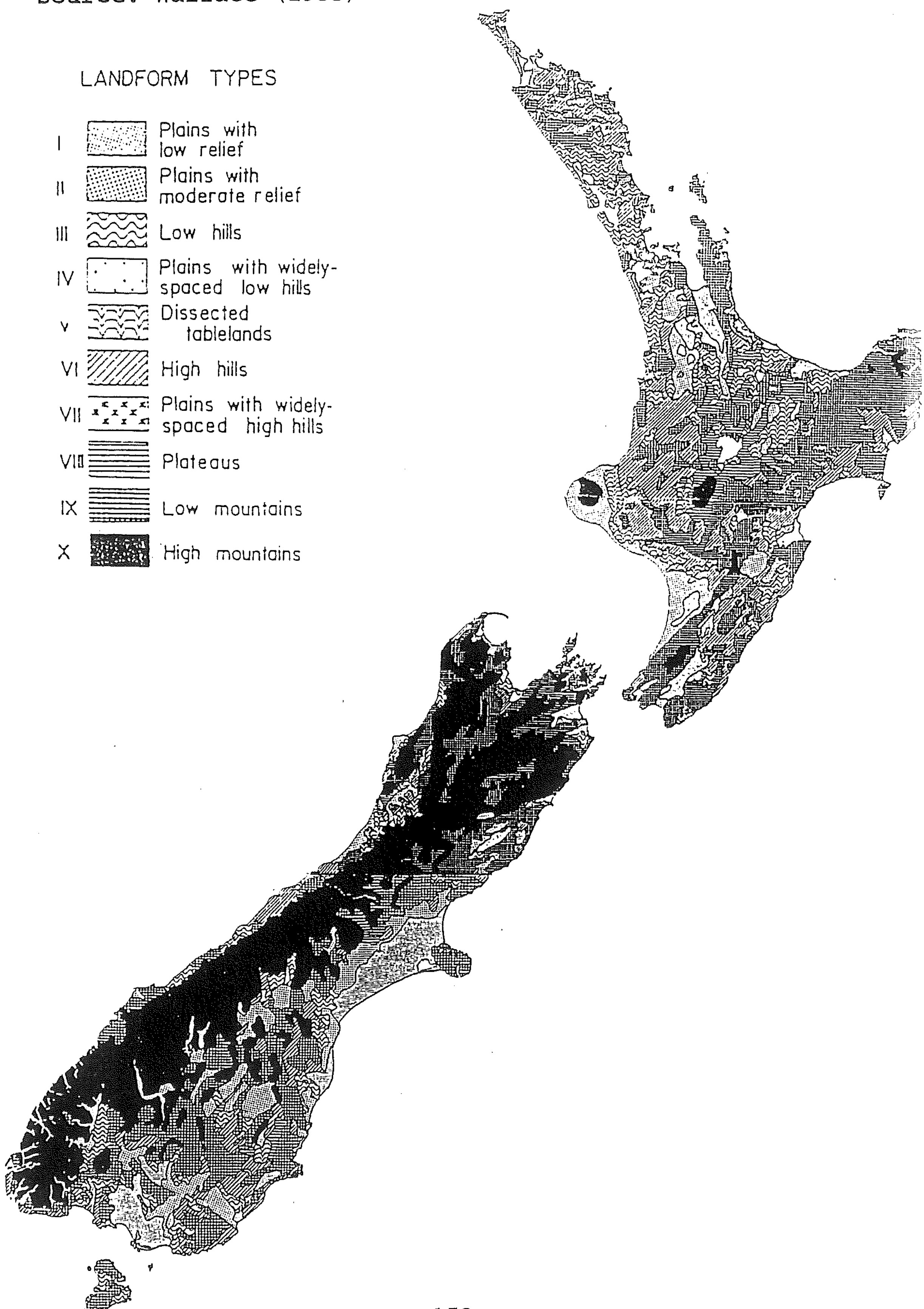
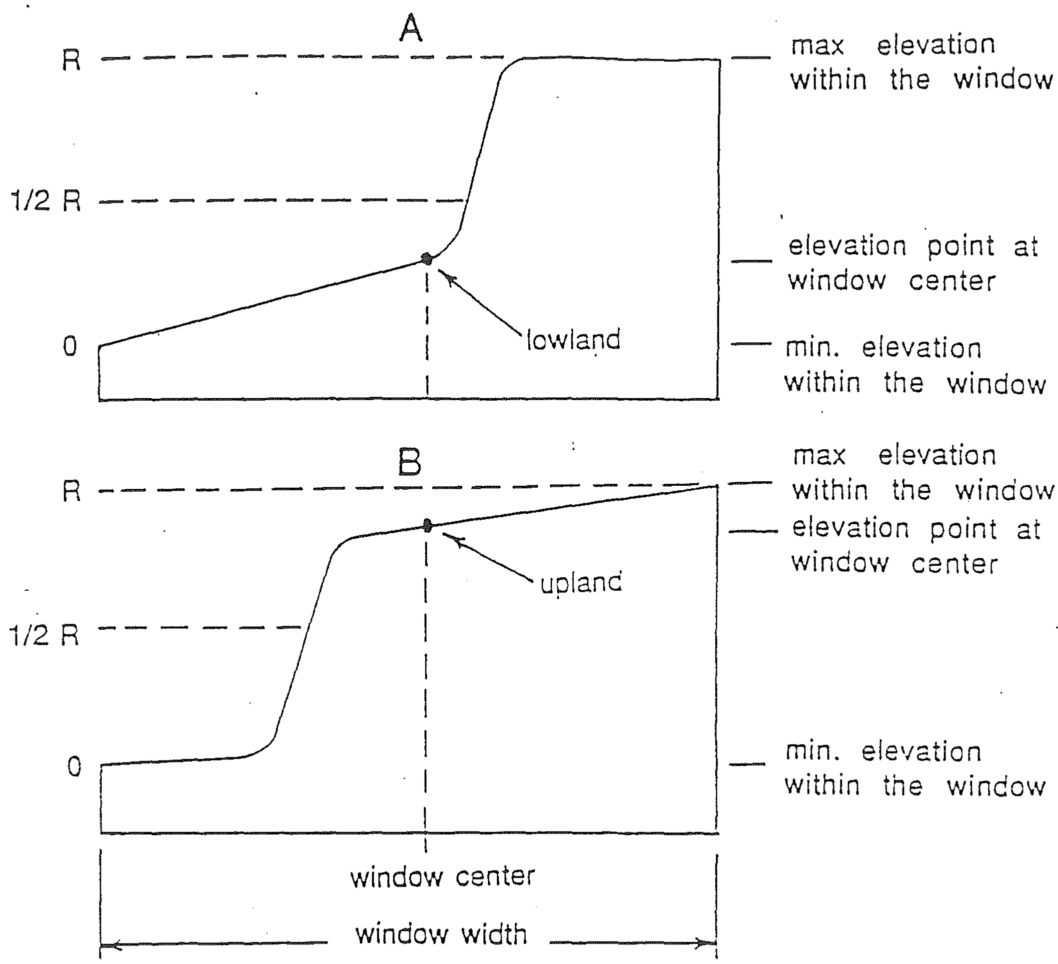


Figure 5.3 The identification of upland and lowland
Source: Dikau et al. (1991)



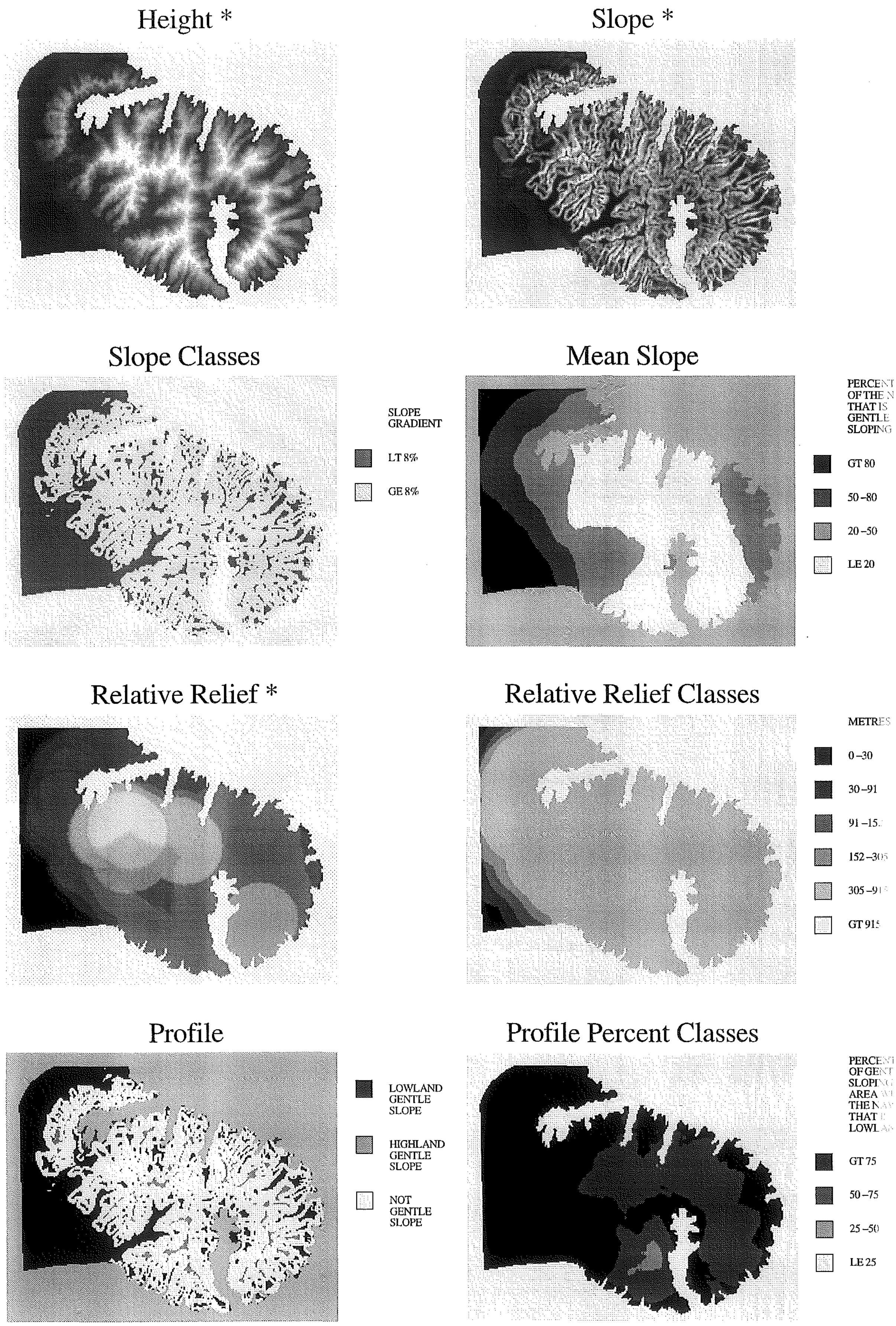
upland: max. elevation - elevation < 1/2 R

lowland: max. elevation - elevation > 1/2 R

when max. elevation = maximum elevation within the moving window
elevation = elevation at moving window center
R = local relief within the moving window

Figure 5.4 Different Stages of the Automated Process

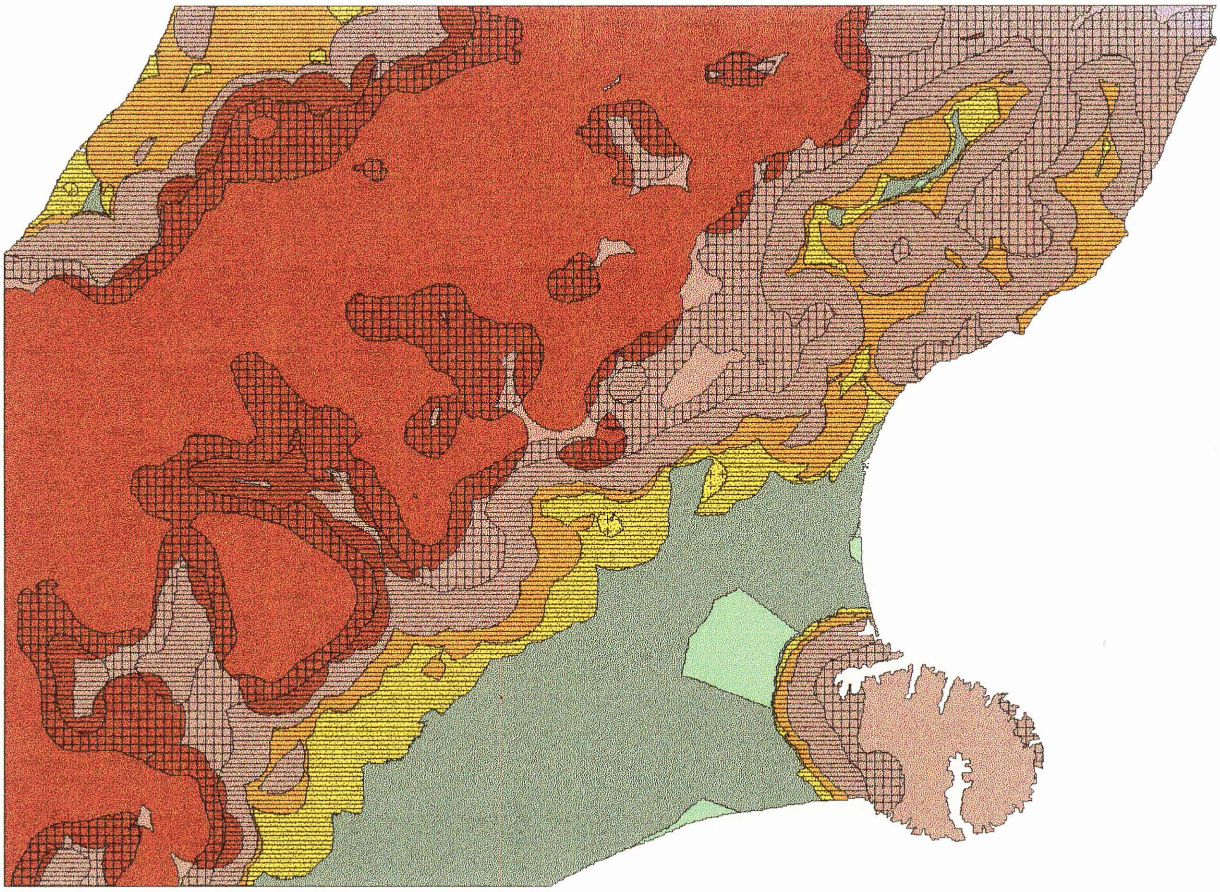
* = Continuous grey scale, with dark as low and bright as high



Area: Banks Peninsula	Cell size: 200m	Slope: 8 percent	Map ID. lfprocessdikau
Data source: DOSLI 1:250,000 (100m contours)			NAW radius: 5600m

Figure 5.5 Landform Classes (Hammond/Dikau)

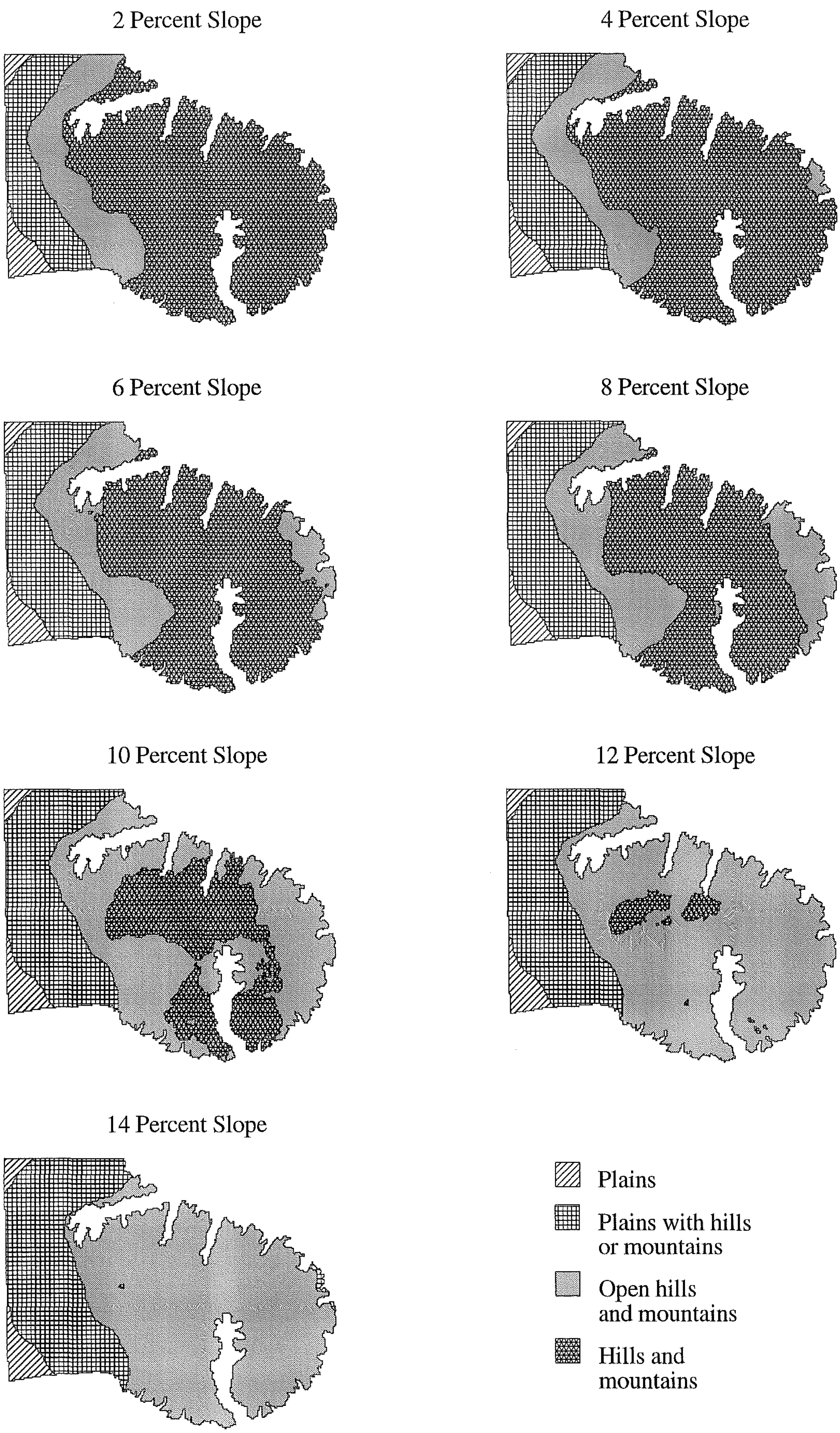
Main road and hydrology layers are added for geographical reference



- Flat or nearly flat plains
- Smooth plains with some local relief
- Tablelands with moderate relief
- Plains with hills
- Plains with high hills
- Plains with low mountains
- Plains with high mountains
- Open high hills
- Open low mountains
- Open high mountains
- Low mountains
- High mountains

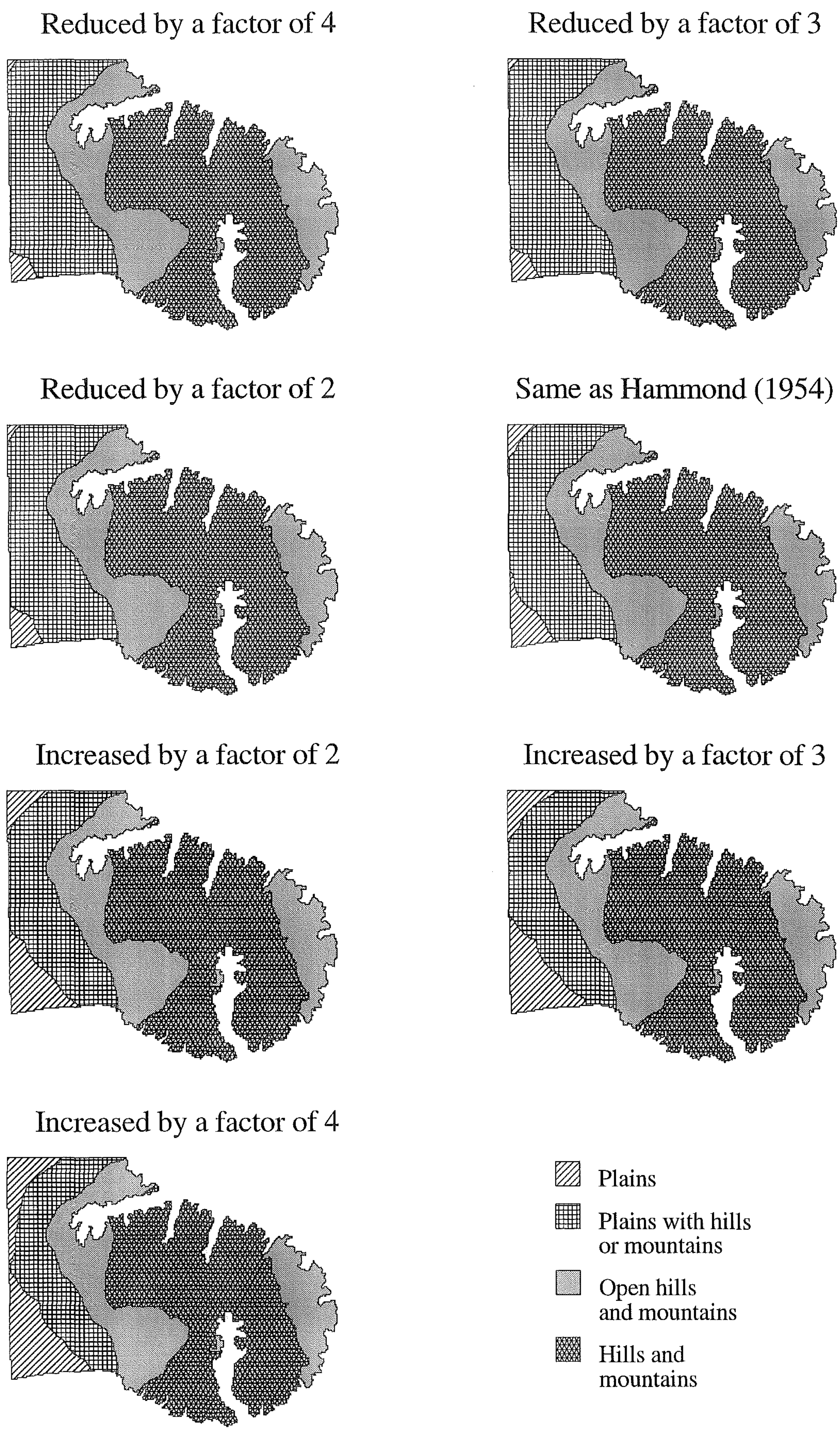
Cell size: 200m	Slope: 8 percent	Map ID. lfdikau2
Data source: DOSLI 1:250,000 (100m contours)		NAW radius: 5600m

Figure 5.6 Effects Of Different Slope Thresholds On The Resulting Landform Type Classification



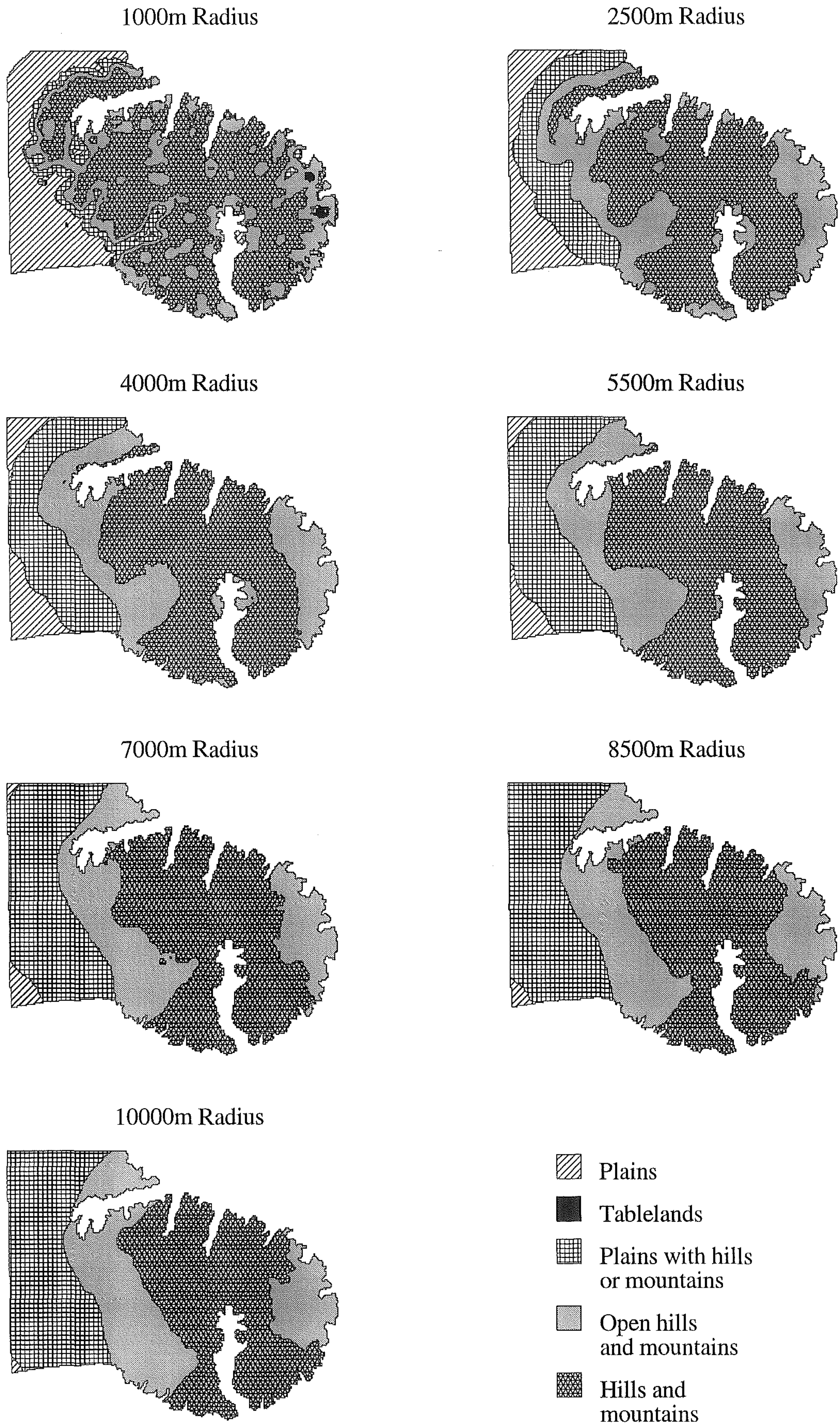
Area: Banks Peninsula	Cell size: 200m	NAW radius: 5600m
Data source: DOSLI 1:250,000 (100m contours)		Map ID. lfslope

Figure 5.7 Effects Of Different Relative Relief Classes On The Resulting Landform Type Classification



Area: Banks Peninsula	Cell size: 200m	Slope: 8 percent	NAW radius: 5600m
Data source: DOSLI 1:250,000 (100m contours)			Map ID. lfrelrel

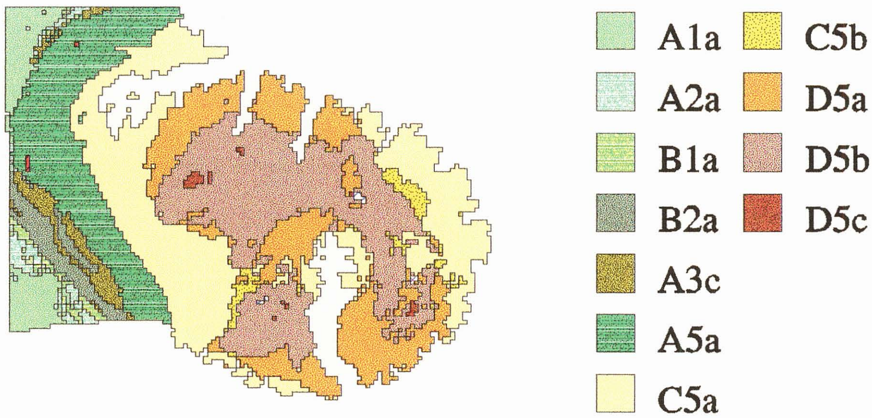
Figure 5.8 Effects Of Different NAW Radii On The Resulting Landform Type Classification



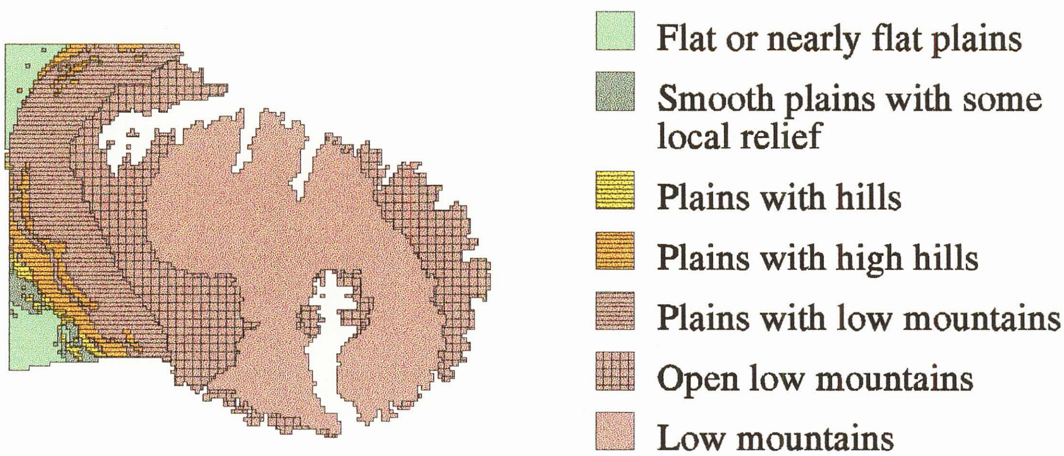
Area: Banks Peninsula	Cell size: 200m	Slope: 8 percent
Data source: DOSLI 1:250,000 (100m contours)		Map ID. lfradii

Figure 5.9 The Majority Resulting From The Combination Of 45 Different Classifications

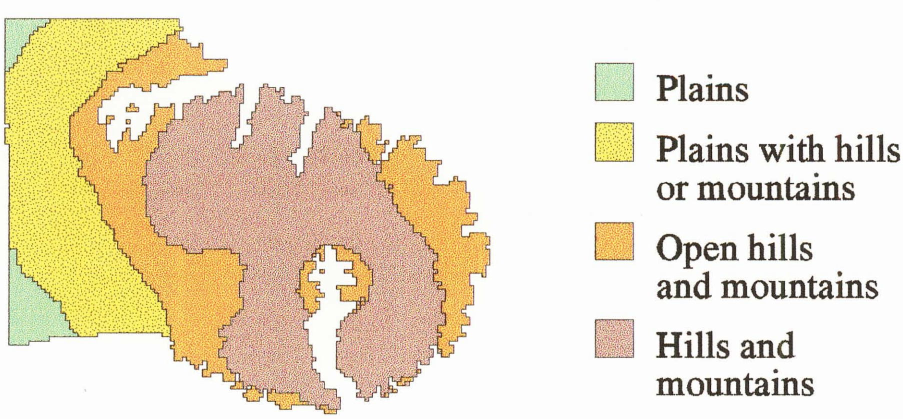
Landform Subclasses



Landform Classes



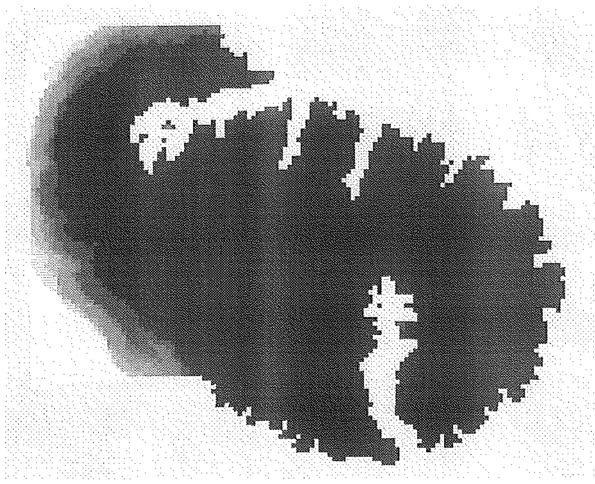
Landform Types



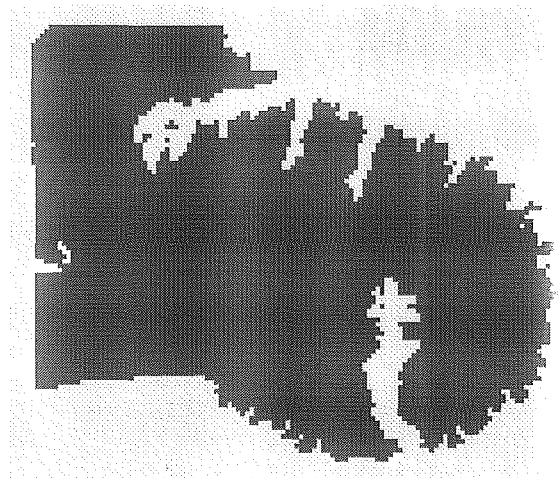
Cell size: 400m	Slope: 4 - 12 percent	NAW radius: 2400 - 8400m
Data source: DOSLI 1:250,000 (100m contours)		Map ID. lfmaj

Figure 5.10 The Membership of Different Landform Types

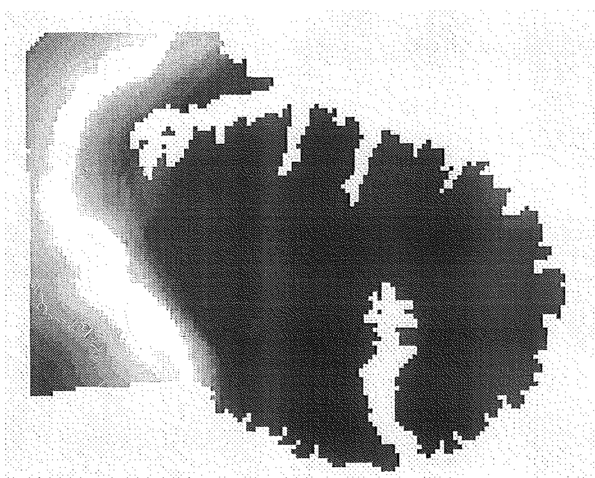
Plains



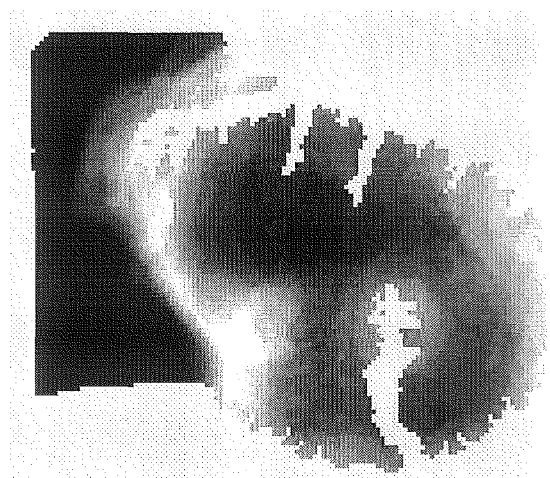
Tablelands



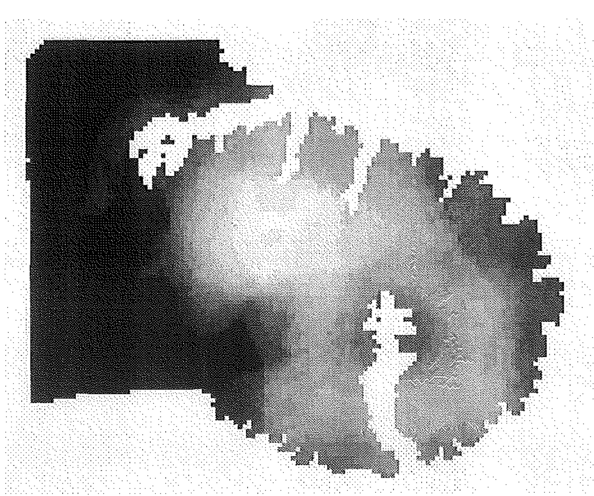
Plains with Hills and Mountains



Open Hills and Mountains



Hills and Mountains



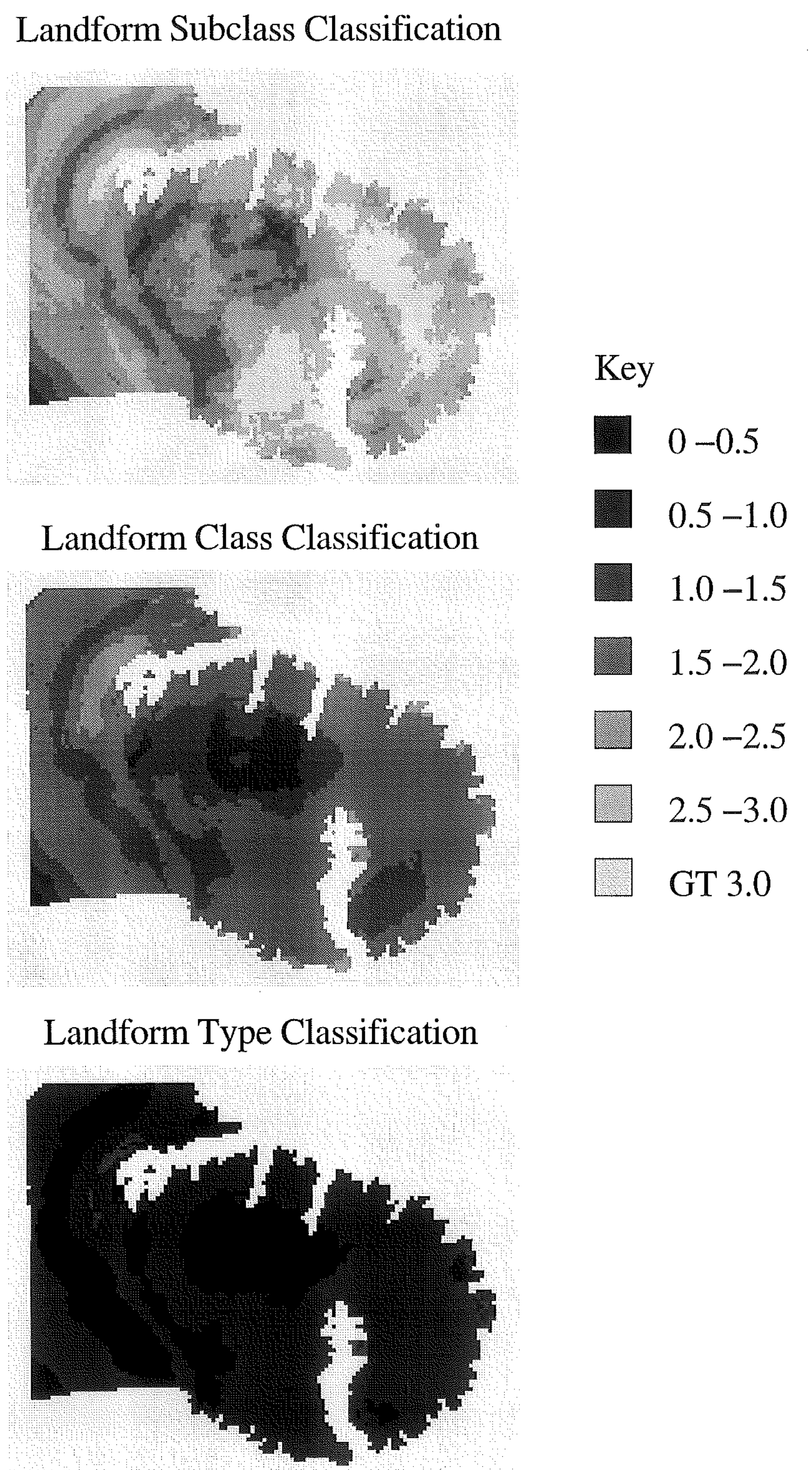
Continuous Grey scale

Darkness represents low membership

Brightness represents high membership

Cell size: 400m	Slope: 4 –12 percent	NAW radius: 2400 –8400m
Data source: DOSLI 1:250,000 (100m contours)		Map ID. lfmemlft

Figure 5.11 Entropy values

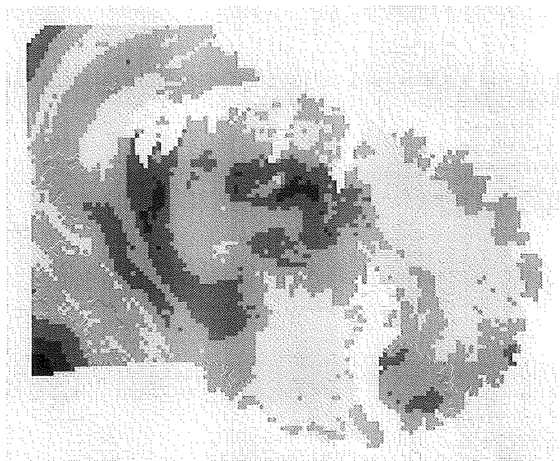


Cell size: 400m	Slope: 4 –12 percent	NAW radius: 2400 –8400m
Data source: DOSLI 1:250,000 (100m contours)		Map ID. lfentropy

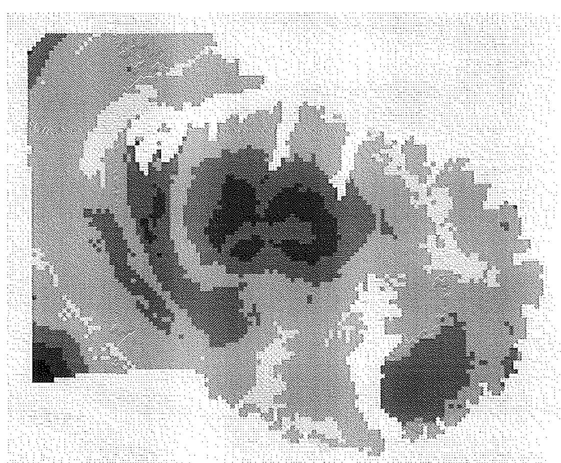
Figure 5.12 Degree of Agreement (Percentage)

0 –20 | 20 –40 | 40 –60 | 60 –80 | 80 –100

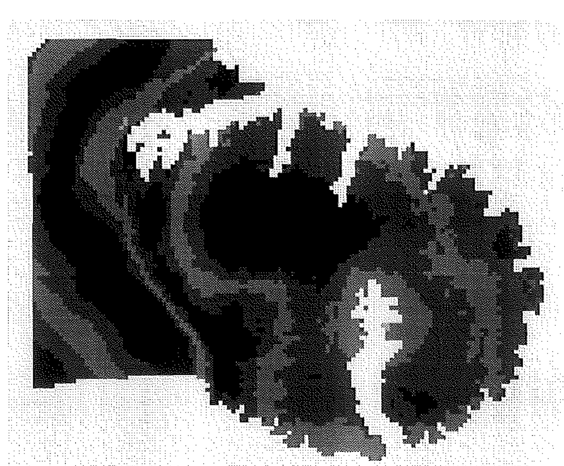
Landform Subclass Classification



Landform Class Classification

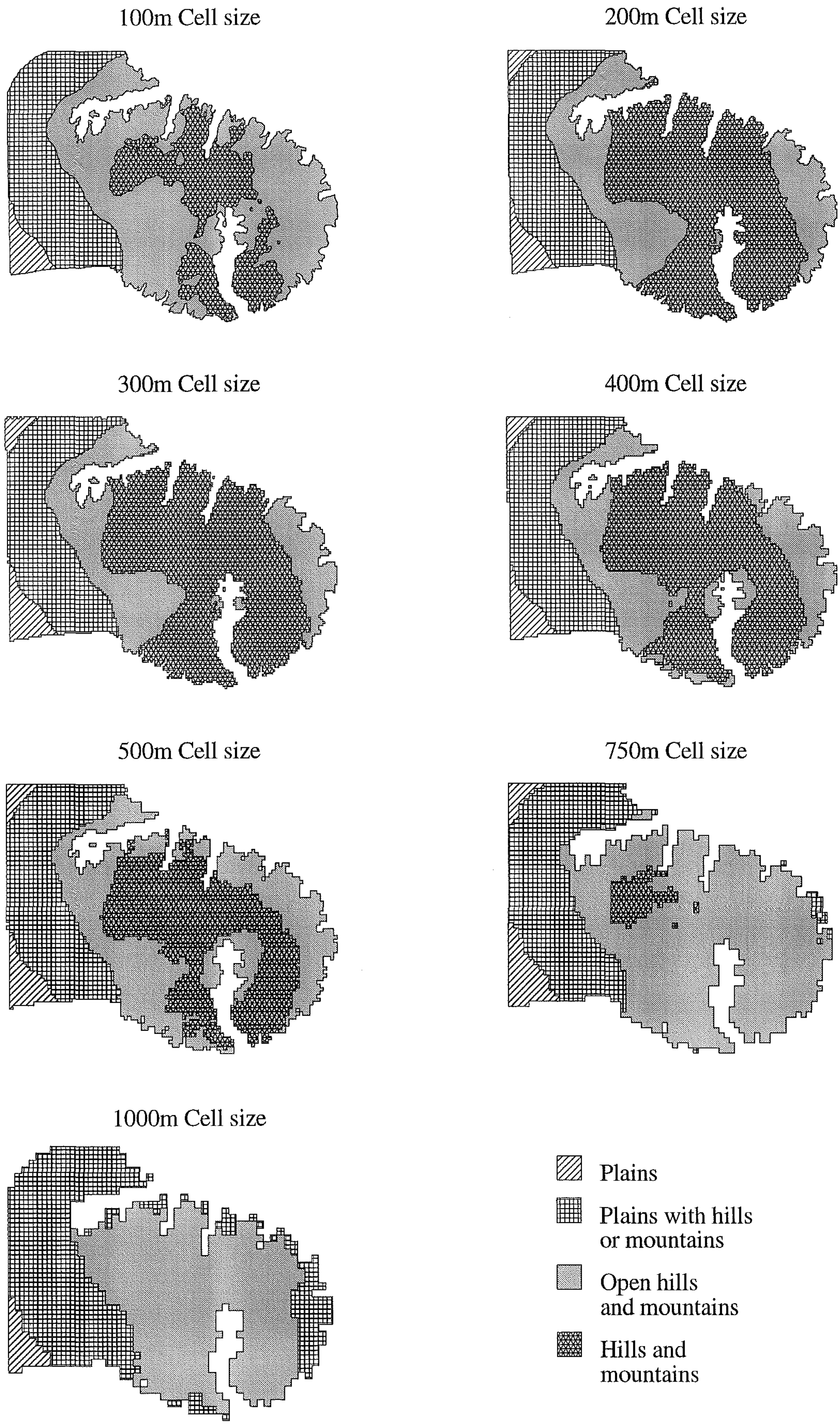


Landform Type Classification



Cell size: 400m	Slope: 4 –12 percent	Map ID. lfpropclass
Data source: DOSLI 1:250,000 (100m contours)		NAW radius: 2400 –8400m

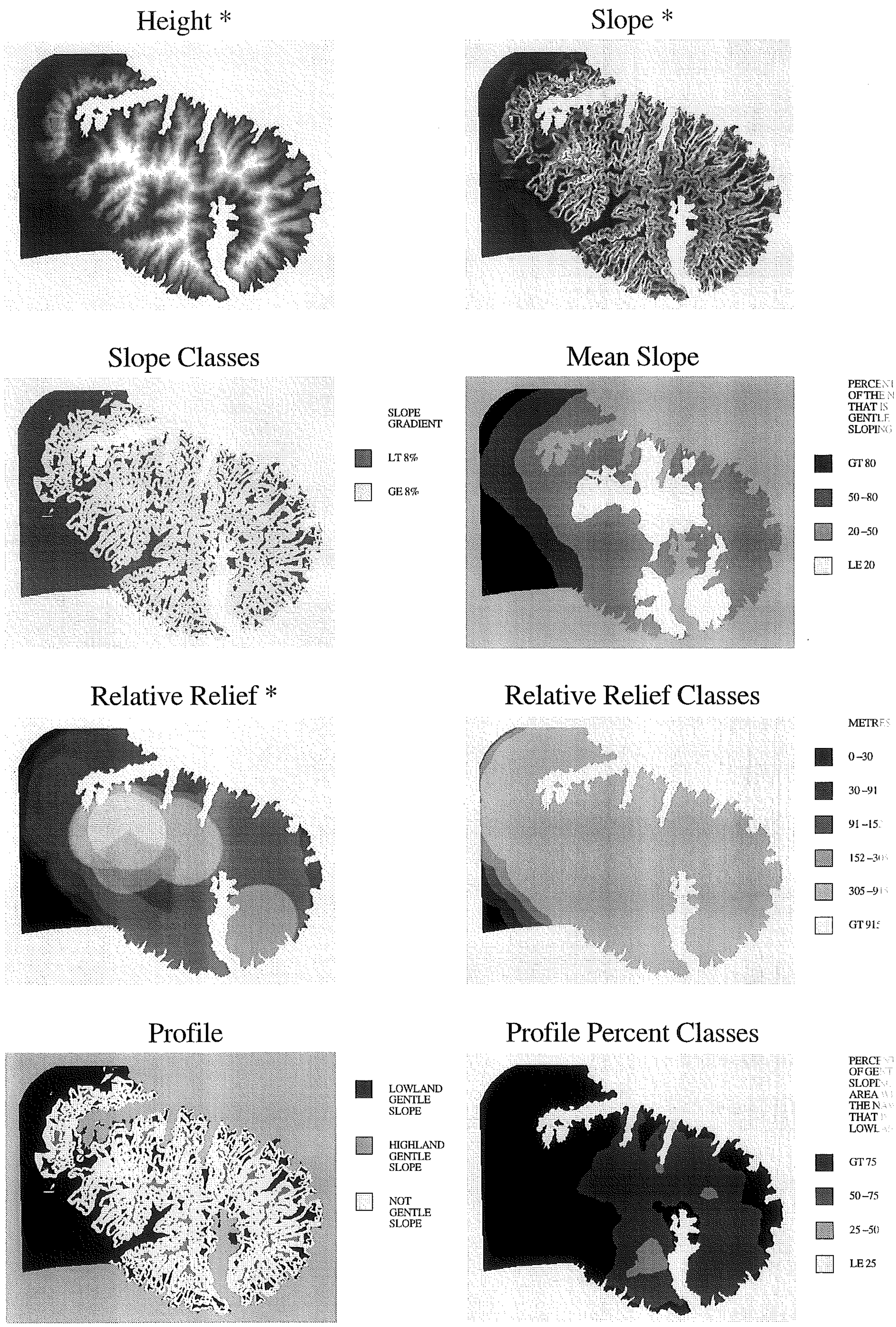
Figure 5.13 Effects Of Different Cell Sizes On The Resulting Landform Type Classification



Area: Banks Peninsula	Slope: 8 percent	Map ID. lfcellsize
Data source: DOSLI 1:250,000 (100m contours)		NAW radius: 5600m

Figure 5.14 Different Stages of the Automated Process (cell size 100m)

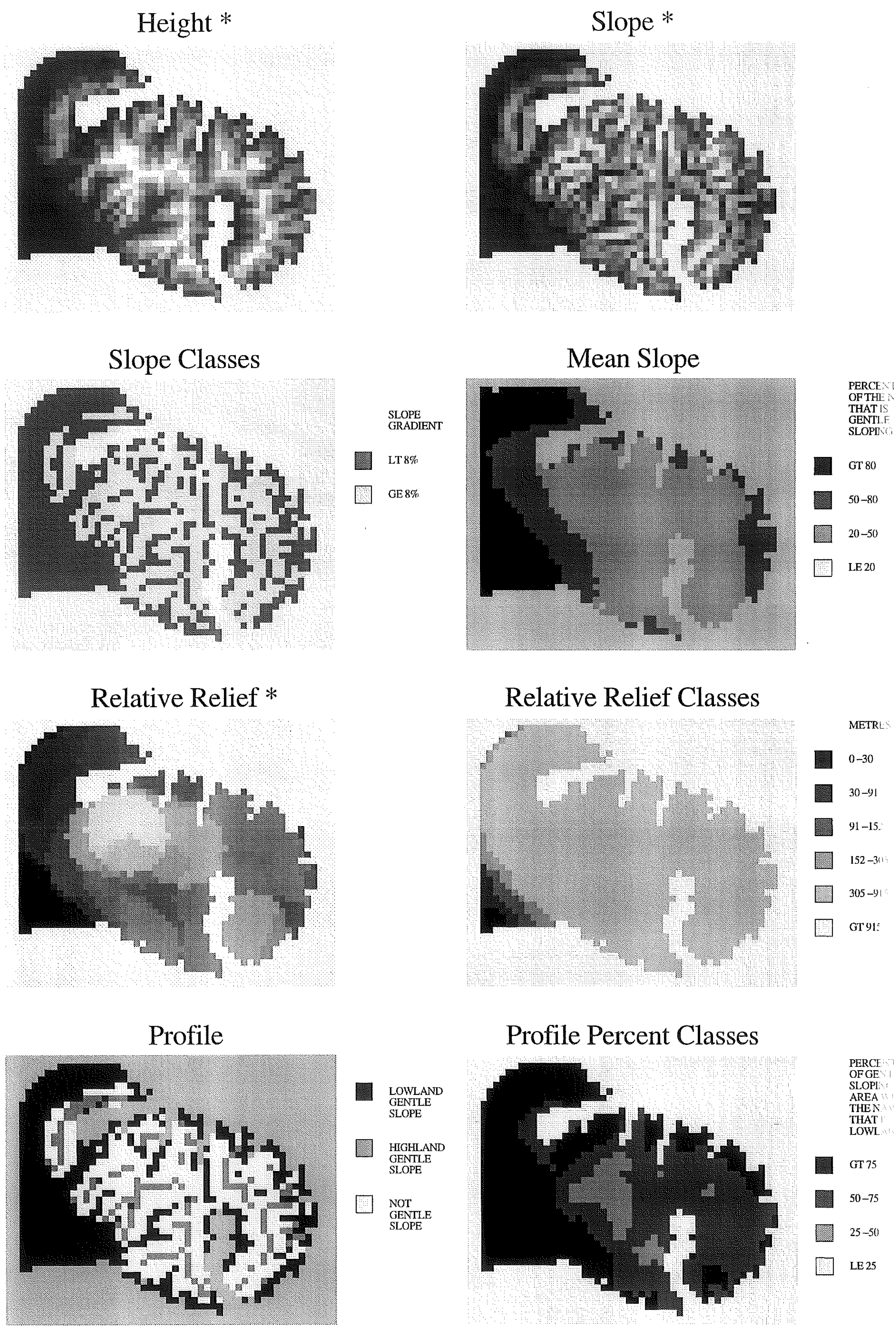
* = Continuous grey scale, with dark as low and bright as high



Area: Banks Peninsula	Cell size: 100m	Slope: 8 percent	NAW radius: 5600m
Data source: DOSLI 1:250,000 (100m contours)		Map ID. lfprocessdikau100	

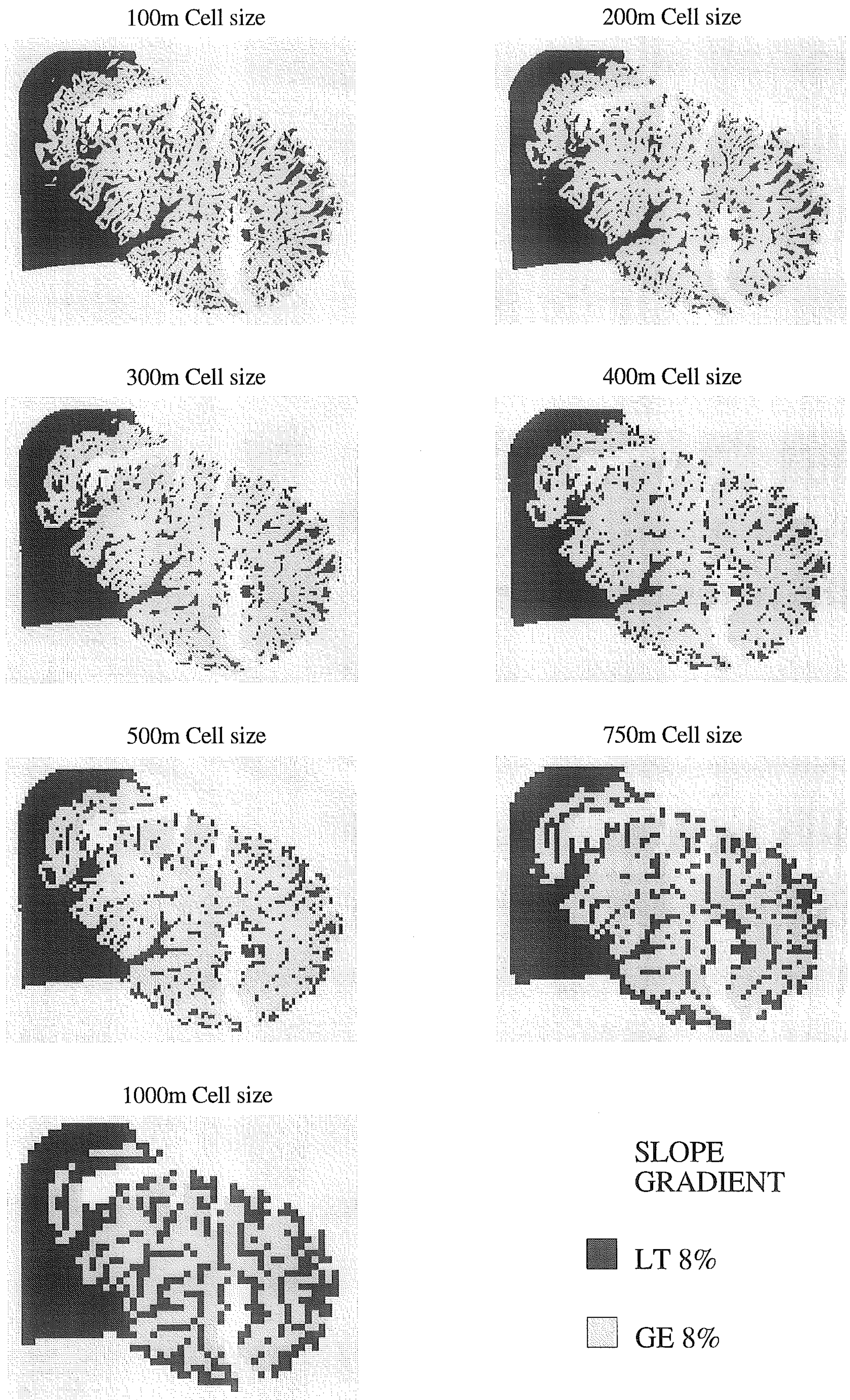
Figure 5.15 Different Stages of the Automated Process (cell size 1000m)

* = Continuous grey scale, with dark as low and bright as high



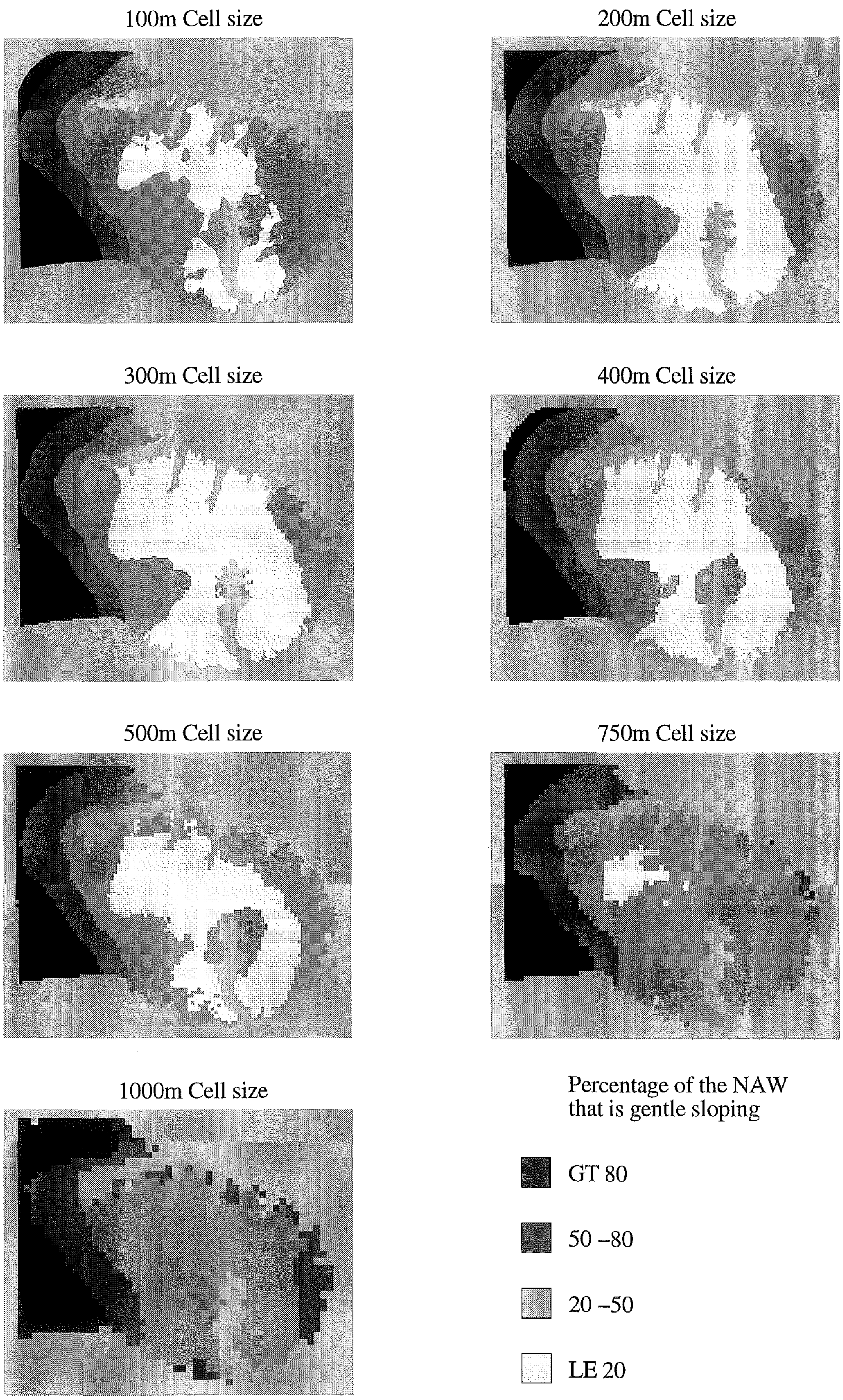
Area: Banks Peninsula	Cell size: 1000m	Slope: 8 percent	NAW radius: 5600m
Data source: DOSLI 1:250,000 (100m contours)		Map ID. lfprocessdikau1000	

Figure 5.16 Effects of Different Cell Sizes on Slope Gradient



Area: Banks Peninsula	NAW radius: 5600m
Data source: DOSLI 1:250,000 (100m contours)	Map ID. lfcellslopeclass

Figure 5.17 Effects of Different Cell Sizes on Mean Slope

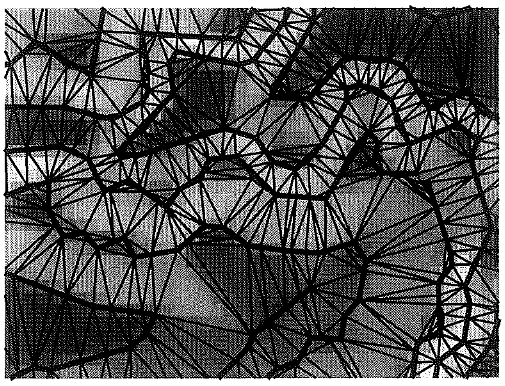


Area: Banks Peninsula	NAW radius: 5600
Data source: DOSLI 1:250,000 (100m contours)	Map ID. lfcellslopemean

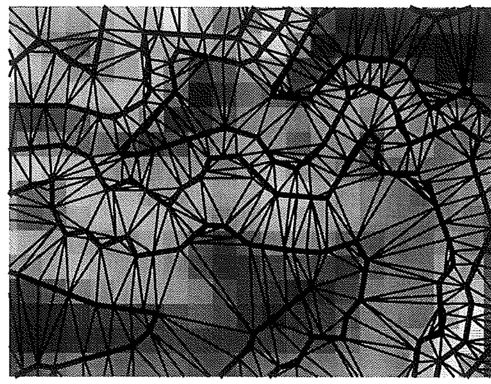
Figure 5.18 Closeup view of the Generalisation Effects Of Different Cell Sizes on Slope

Tin and contour lines are shown to indicate what the slope values should be

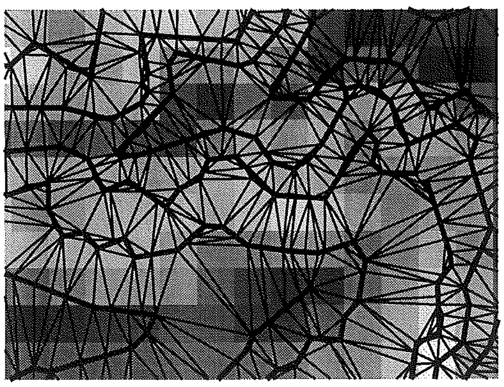
100m Cell size



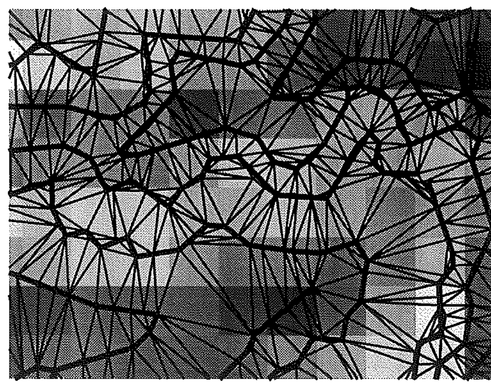
200m Cell size



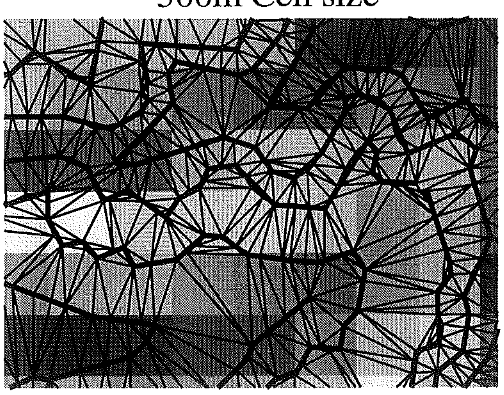
300m Cell size



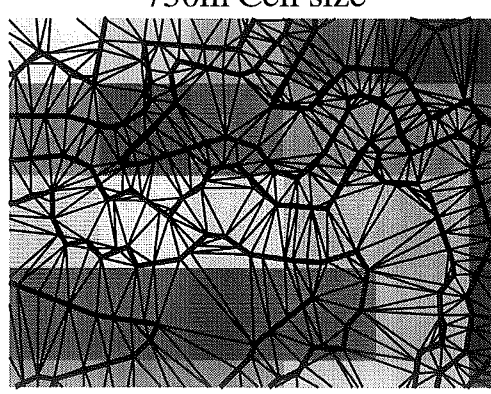
400m Cell size



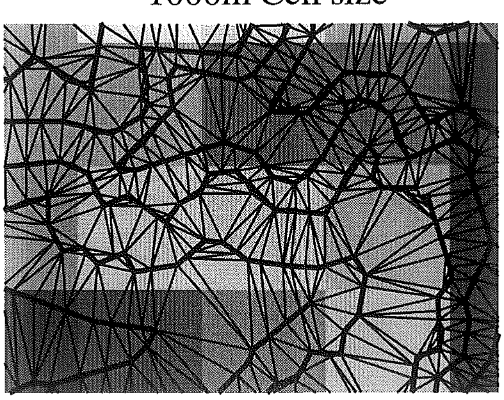
500m Cell size



750m Cell size



1000m Cell size



Continuous Grey Scale

Dark is flat

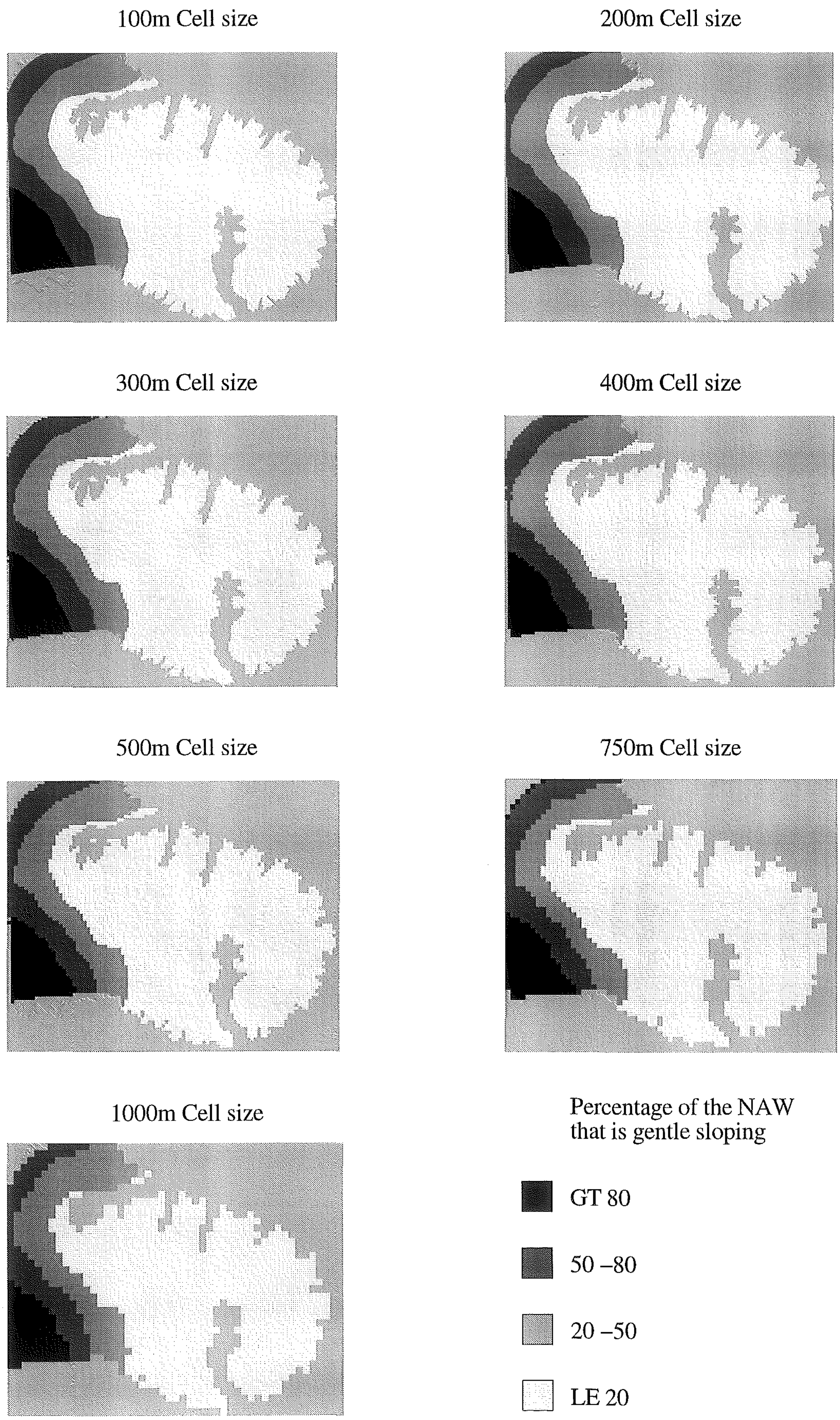
Bright is steep

—— TIN Lines

▬▬ Contour Lines

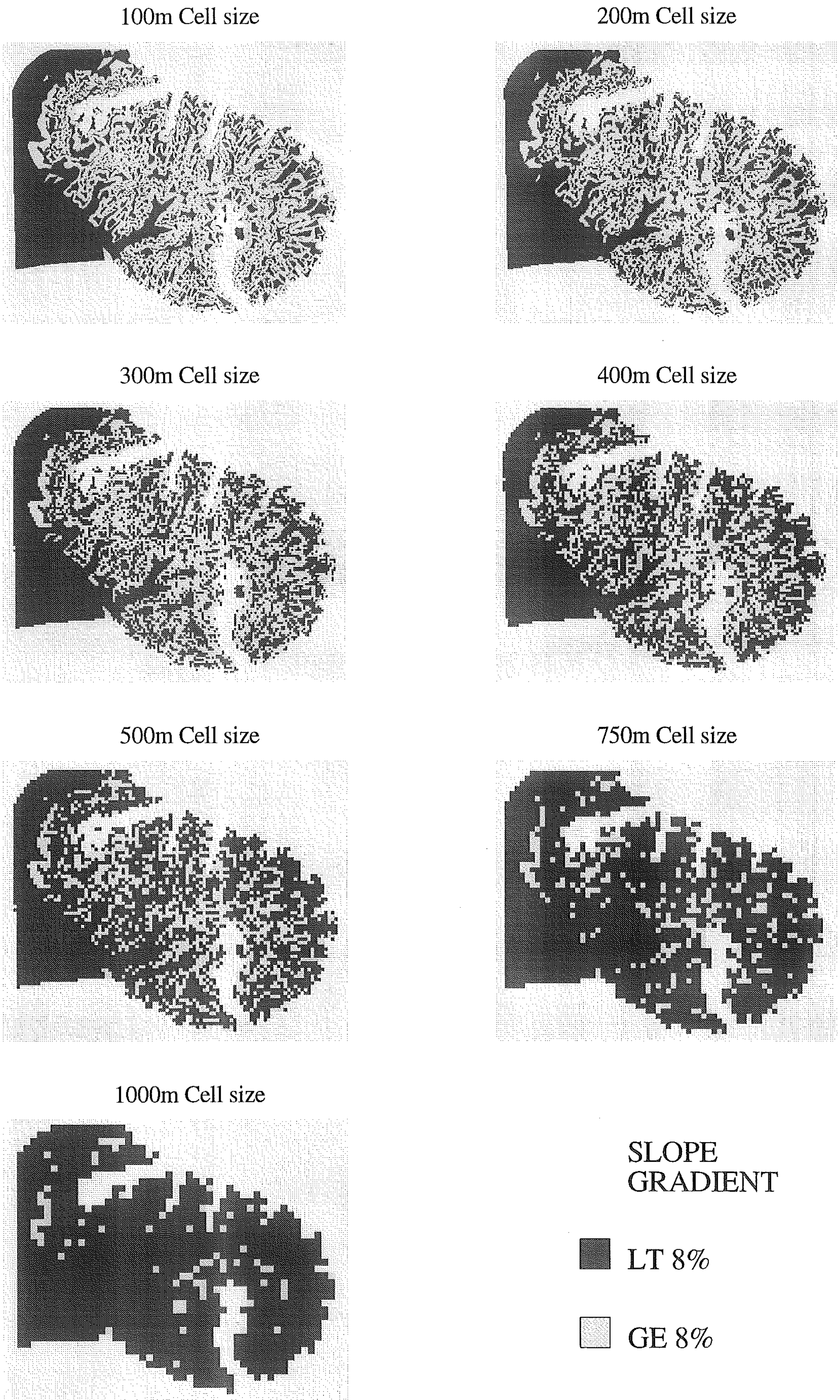
Mapextent: NZMG 2495000, 5713000, 2499000, 5716000
MAP ID. lfgridslope

Figure 5.19 Effects of Different Cell Sizes on Mean Slope (slope information obtained from the LRI)



Area: Banks Peninsula	NAW radius: 5600m
Data source: DOSLI 1:250,000 (100m contours)	Map ID. lfgridslopeclasslri

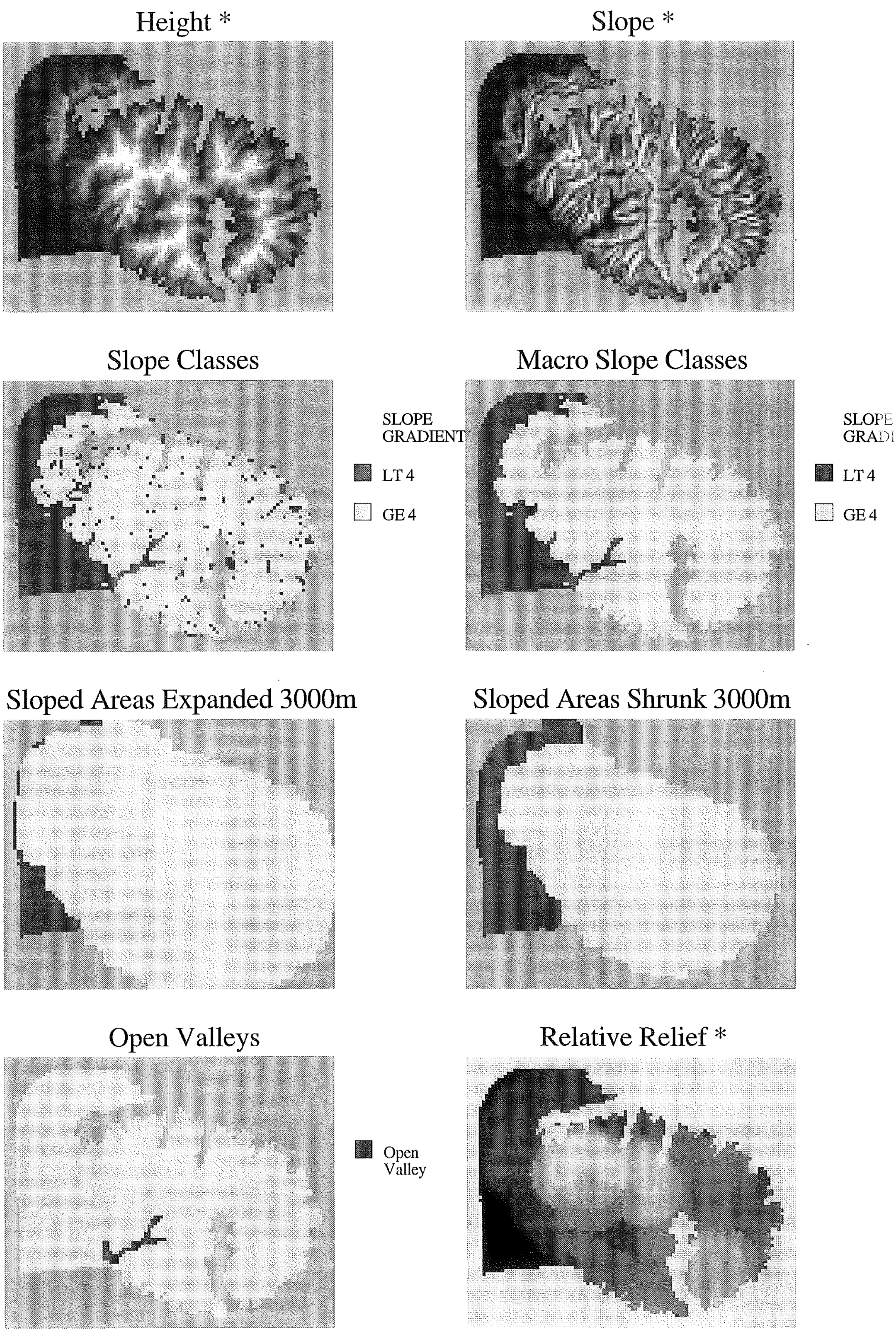
Figure 5.20 Effects of Different Cell Sizes on Slope (slope information obtained from TIN)



Area: Banks Peninsula	NAW radius: 5600m
Data source: DOSLI 1:250,000 (100m contours)	Map ID. lfgridslopeclasstin

Figure 5.21 Different Stages of the Automated Process (Brabyn)

* = Continuous grey scale, with dark as low and bright as high.

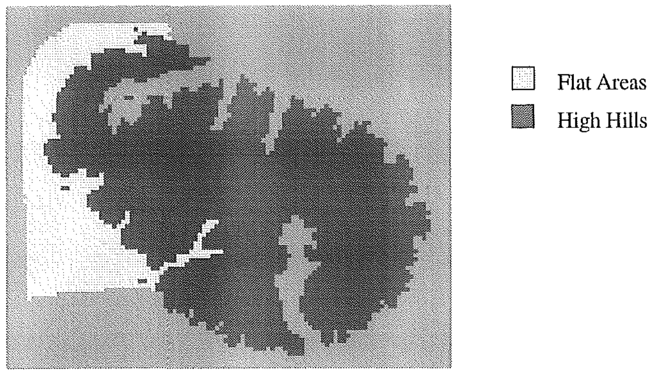


Area: Banks Peninsula Cell size: 500m Slope: 4 percent NAW radius: 5500m
Data source: DOSLI 1:250,000 (100m contours) Map ID. lfprocesspt1

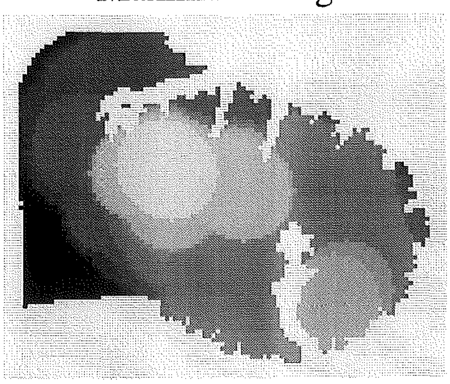
Figure 5.22 Different Stages of the Automated Process (cont.)

* = Continuous grey scale, with dark as low and bright as high.

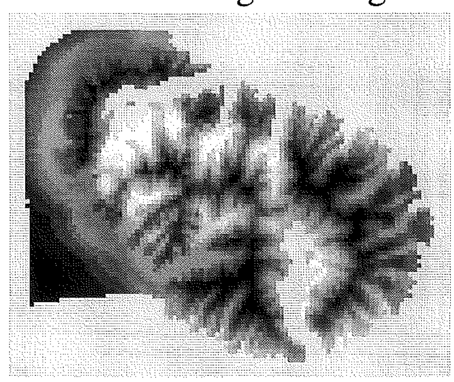
Relief Types



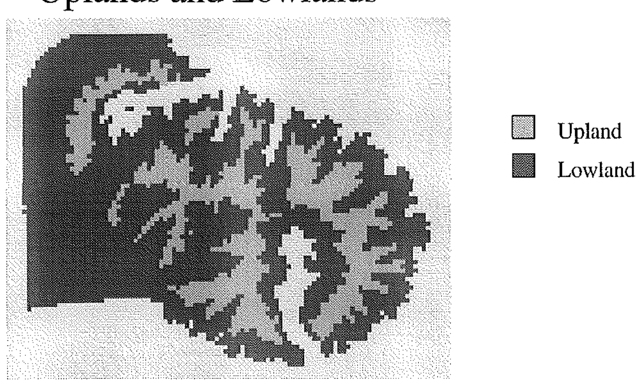
Maximum Height *



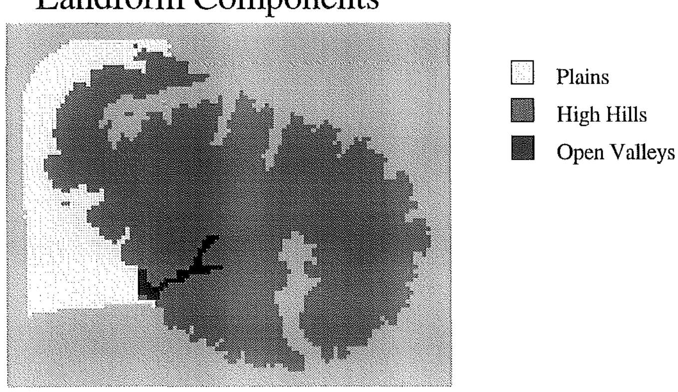
Maximum Height –Height *



Uplands and Lowlands



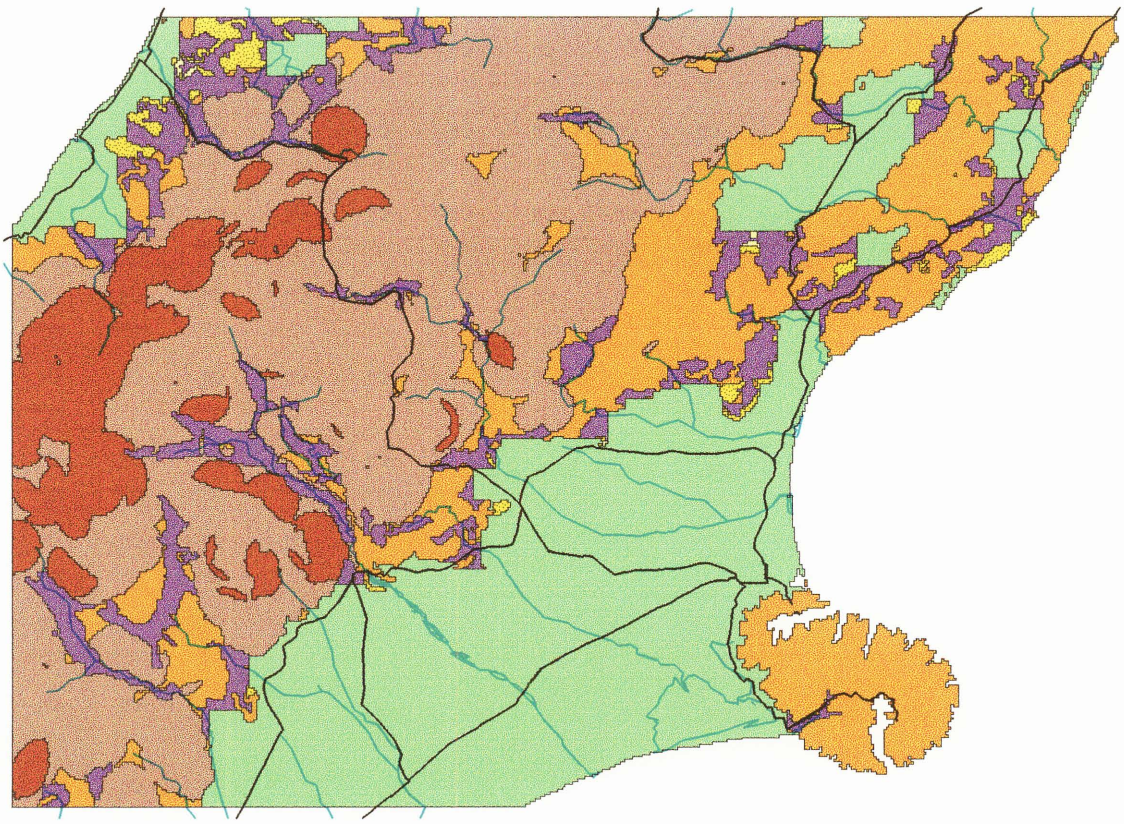
Landform Components



Area: Banks Peninsula	Cell size: 500m	Slope: 4 percent	NAW radius: 5500m
Data source: DOSLI 1:250,000 (100m contours)		Map ID. lfprocesspt2	

Figure 5.23 Landform Components

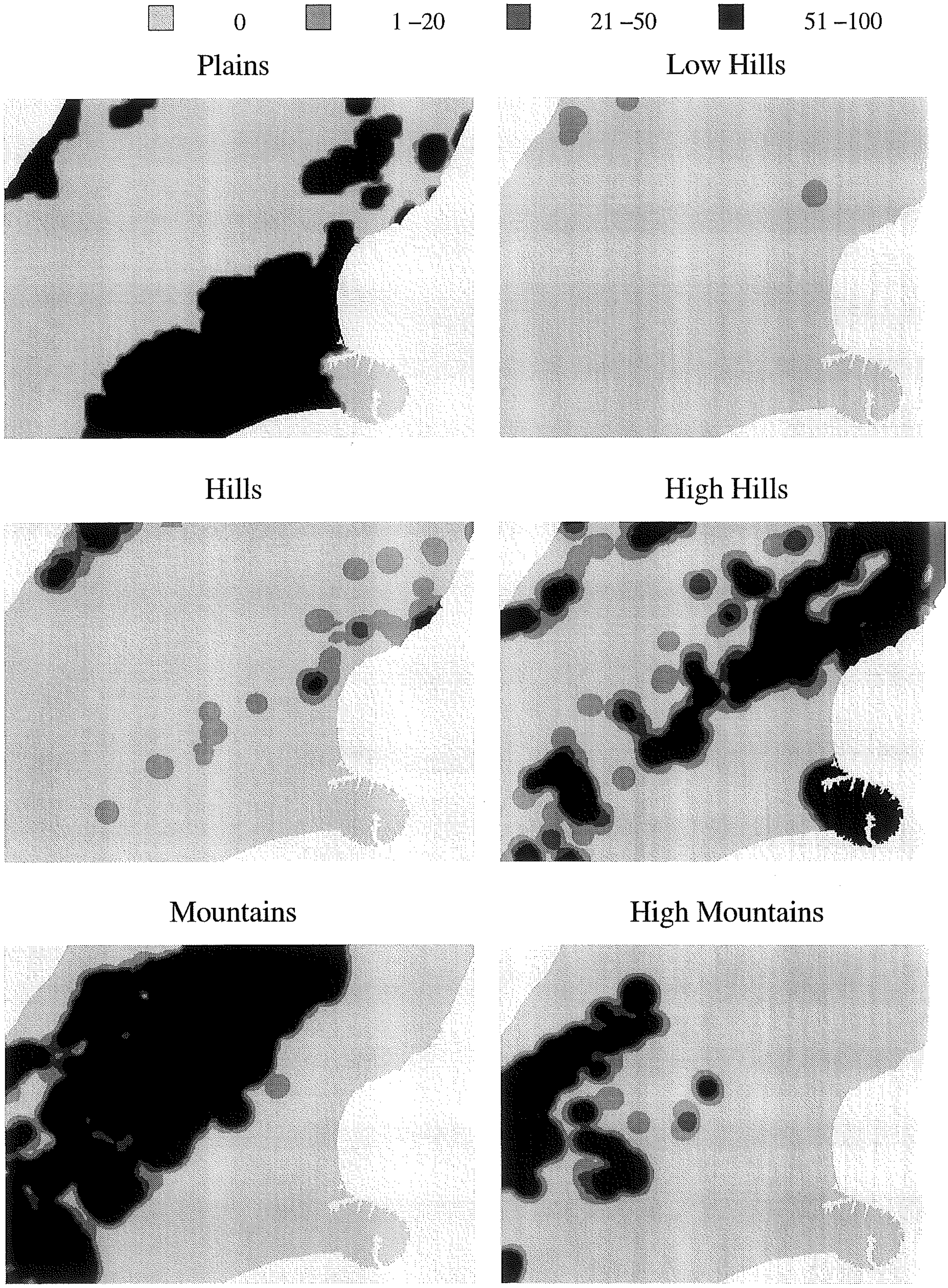
Main road and hydrology layers are added for geographical reference



Cell size: 500m	NAW radius: 5500m	Slope: 4 percent	Map ID. lfcom
Data source: DOSLI 1:250,000 (100m contours)			Valley width criteria: 6000m

Figure 5.24 The Spatial Influence of the Different Landform Componen

(The percentage of cells within the NAW that contain the specified landform class)



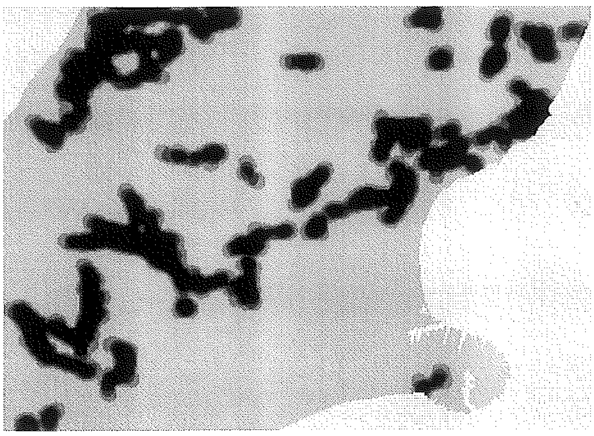
Data source: DOSLI 1:250,000 (100m contours)	Cell size: 500m	NAW radius: 3000m	MAP ID. lfmean

Figure 5.25 The Spatial Influence of the Different Landform Components (con

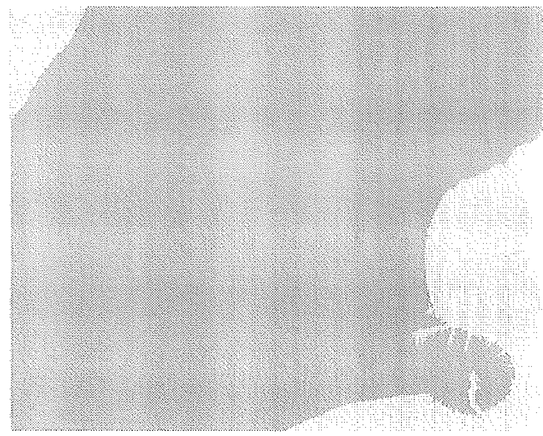
(The percentage of cells within the NAW that contain the specified landform class)

0	1 –20	21 –50	51 –100

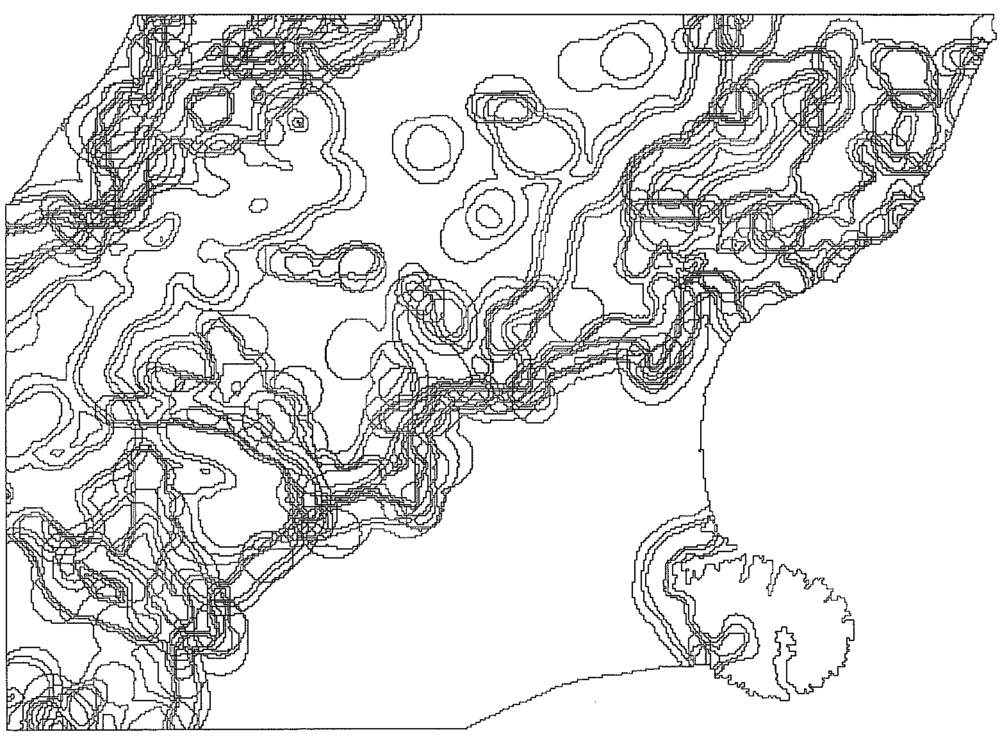
Open Valleys



Tablelands



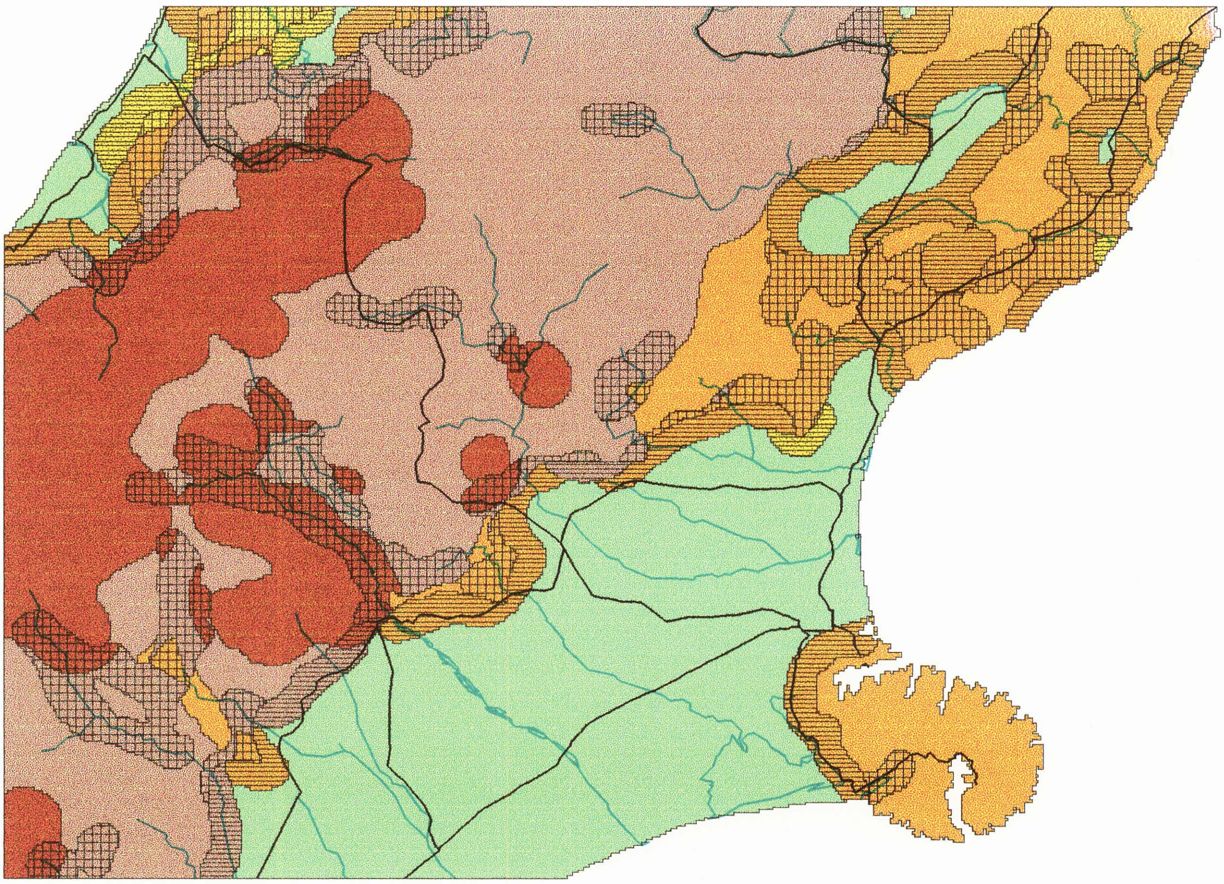
The unique combinations resulting from the overlaying of all the component influence grids (613 unique combinations)



Data source: DOSLI 1:250,000 (100m contours) Cell size: 500m NAW radius: 3000m MAP ID. lfmea

Figure 5.26 Landform Level 1 (Brabyn)

Main road and hydrology layers are added for geographical reference



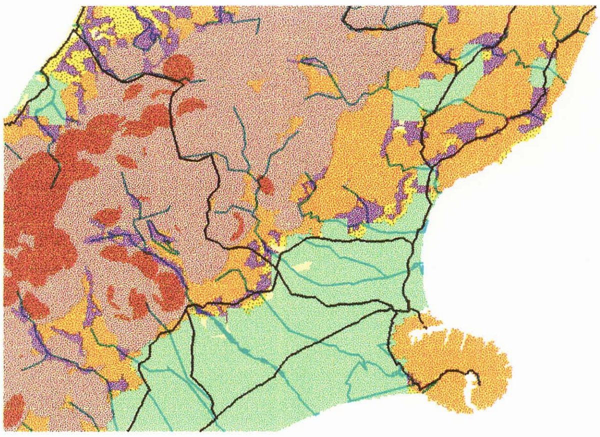
Plains	Open Valley Mountains
Plains Hills	Open Valley High Mountains
Plains High Hills	Hills
Plains Mountains	High Hills
Plains High Mountains	Mountains
Open Valley Hills	High Mountains
Open Valley High Hills	

Cell size: 500m NAW radius: 5500m Slope: 4 percent Map ID. lflev1
Data source: DOSLI 1:250,000 (100m contours) Valley width criteria: 6000m

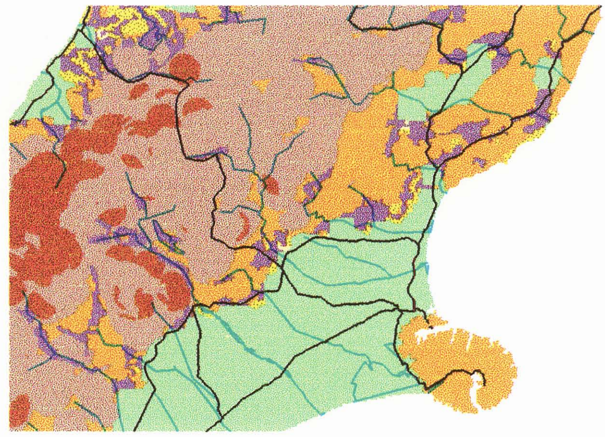
Figure 5.27 The Effects of Slope on Landform Components

Main road and hydrology layers are added for geographical reference

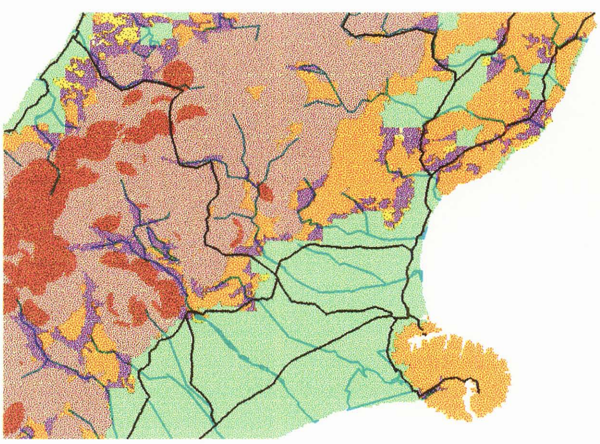
1 Percent slope



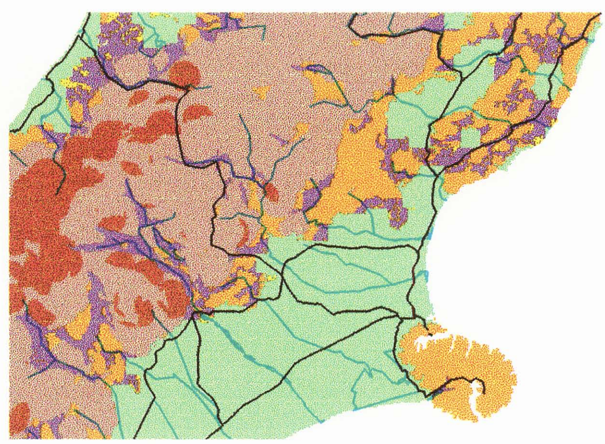
2 Percent slope



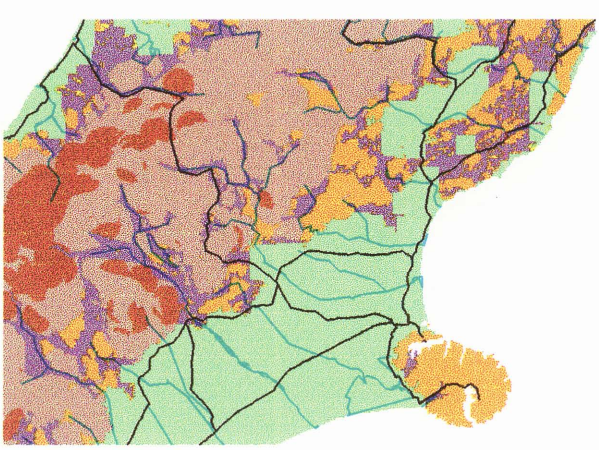
4 Percent slope

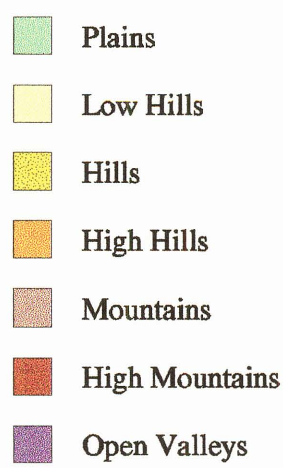


6 Percent slope



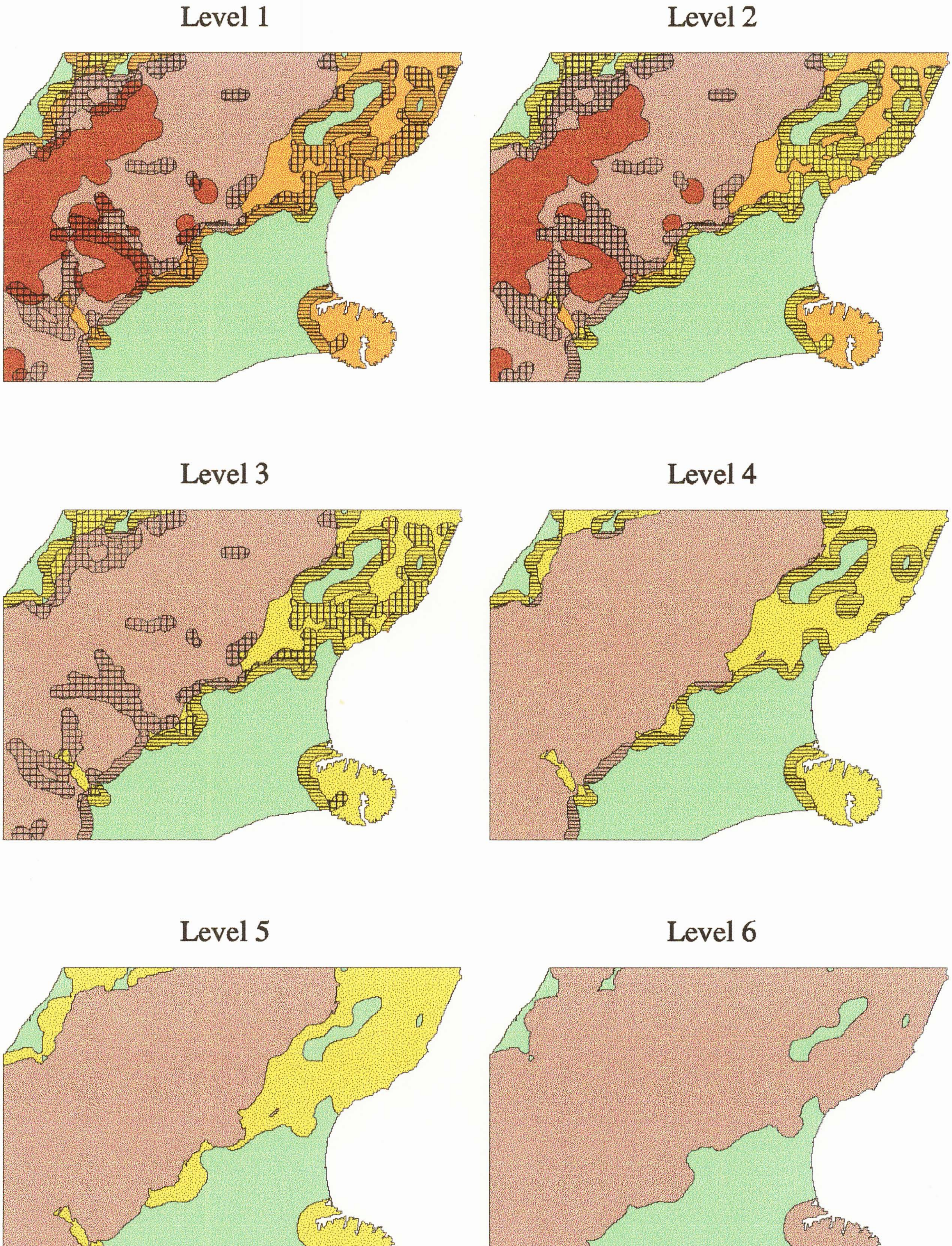
8 Percent slope





Cell size: 500m	NAW radius: 5500m	Map ID. lfslopebra
Data source: DOSLI 1:250,000 (100m contours)		Valley width criteria: 6000m

Figure 5.28 The Effects of Generalisation on Landform



Cell size: 500m	NAW radius: 5500m	Slope: 4 percent	Map ID. lflev1 - 6
Data source: DOSLI 1:250,000 (100m contours)			Valley width criteria: 6000m

Table 5.1 Summary of landform classifications

(- Not stated ? Mentioned but not explicitly)

Wallace (1955)

Parameters	High mountains	Low mountains	Plateaux	Plains with widely spaced high hills	High hills	Dissected tablelands	Plains with widely spaced low hills	Low hills	Plains with moderate relief	Plains with low relief
Slope (percent)	-	-	< 8	< 8	> 8	< 8	< 8	> 8	-	-
Altitude (m)	-	-	?	?	-	?	?	-	-	-
Relative relief (m)	> 900	300-900	200-300	200-300	200-300	100-200	100-200	100-200	30-100	< 30

Linton (1970)

Parameters	Mountains	Bold hills	Hill country	Plateau uplands	Low uplands	Lowlands
Slope	?	?	-	-	-	-
Altitude (m)	-	-	200-600	?	< 300	< 150
Relative relief (m)	> 750	> 400	< 300	< 100	-	-

Crozier and Owen (1983)

Parameters	Very high mountains	High mountains	Mountain land plateaux	Mountain land plains	Low mountain land	High hill country	Hill country plateaux	Hill country plain	Low hills	Moderate relief plains	Moderate relief plain land	Low relief plains	Low relief plain land
Slope (degrees)	> 0	> 0	< 8	< 8	> 0	> 0	< 8	< 8	> 0	> 0	> 0	< 8	> 0
Altitude (m)	-	-	-	-	-	-	-	-	-	-	-	-	-
Relative relief (m)	> 1820	900-1820	300-1820	300-1820	900-1820	300-899	100-299	100-299	180-299	40-99	100-179	40-99	< 40

Table 5.2 Comparison between Hammond's (1964) manual classification process and Dikau et al.'s (1991) automatic process
Source: Dikau et al. (1991)

Item	Hammond	Digital approach
Data source	1 : 250,000 AMS topographic map	100 m DMA digital elevation model (BRABB et al. 1989)
Contour interval or data resolution / data points	Contour interval 15.2 to 61 m	200 m / 8 million pixels
Attributes	Slope, relief, profile type	Slope, relief, profile type
Number of subclasses used	45	96
Unit area (window size)	9.65 km across	9.8 km across
Window movement	9.65 km steps	200 m steps
Map generalization	yes	no
Degree of generalization	Absorbing units < 2072 km^2	none
Final map scale	1 : 5,000,000	variable (in this report 1 : 1,000,000)
Area classified	Entire United States	New Mexico (314,255 km^2)

Table 5.3 Dikau et al.'s (1991) process for automating Hammond's landform classification
Source: Dikau et al. (1991)

Basic attribute and type	Derived attribute	Grid modelling procedure	Data set or layer	Program name
Slope	Slope angle	Moving window with 3 by 3 elevation points	A	GPQUAD
	Percent of < 8 % slope	Moving window with 49 by 49 elevation points, and reclassification into the Hammond slope intervals	B	GPPCTLT GPRCDGRD
Local relief	Range of elevation (local relief within moving window)	Moving window with 49 by 49 elevation points, and reclassification into the Hammond relief intervals	C	GPRELIEF GPRCDGRD
Profile type	(1) Lowland and upland distinction			
	Maximum and minimum elevation within moving window	Moving window with 49 by 49 elevation points	D	GPWINDOW
	Difference between maximum elevation and moving window mid-point elevation from original DEM	Subtraction	E	GPLINCOM
	Difference between maximum and minimum elevation in the moving window (range of elevation)	Subtraction	F	GPLINCOM
	One half of range of elevation within moving window	Scaling	G	GPSCALE
	Lowland/upland within moving window by ratio of E and G	Subtraction	H	GPLINCOM
	(2) Profile type			
	Frequency distribution of A	Moving window with 49 by 49 slope angle points	I	GPPCTLT
	Profile type within moving window by combining H and I	Linear combination	J	GPLINCOM
	Profile type within moving window	Reclassification of J into the Hammond profile type intervals	K	GPRCDGRD
Landform type	Combination of attributes	Linear combination of B, C, K	L	GPLINCOM
		Reclassification of L into the 96 landform subclasses, 24 classes and 5 types used in this report	M	GPRCDGRD

Table 5.4 Dikau et al.'s (1991) landform classes
Source: Dikau et al. (1991)

Landform type (5)	Landform class (24)	Landform subclasses code (96)
Plains (PLA)	Flat or nearly flat plains	A1a, A1b, A1c, A1d
	Smooth plains with some local relief	A2a, A2b, A2c, A2d
	Irregular plains with low relief	B1a, B1b, B1c, B1d
	Irregular plains with moderate relief	B2a, B2b, B2c, B2d
Tablelands (TAB)	Tablelands with moderate relief	A3c, A3d, B3c, B3d
	Tablelands with considerable relief	A4c, A4b, B4c, B4d
	Tablelands with high relief	A5c, A5d, B5c, B5d
	Tablelands with very high relief	A6c, A6d, B6c, B6d
Plains with Hills or Mountains (PHM)	Plains with hills	A3a, A3b, B3a, B3b
	Plains with high hills	A4a, A4b, B4a, B4b
	Plains with low mountains	A5a, A5b, B5a, B5b
	Plains with high mountains	A6a, A6b, B6a, B6b
Open Hills and Mountains (OPM)	Open very low hills	C1a, C1b, C1c, C1d
	Open low hills	C2a, C2b, C2c, C2d
	Open moderate hills	C3a, C3b, C3c, C3d
	Open high hills	C4a, C4b, C4c, C4d
	Open low mountains	C5a, C5b, C5c, C5d
	Open high mountains	C6a, C6b, C6c, C6d
Hills and Mountains (HMO)	Very low hills	D1a, D1b, D1c, D1d
	Low hills	D2a, D2b, D2c, D2d
	Moderate hills	D3a, D3b, D3c, D3d
	High hills	D4a, D4b, D4c, D4d
	Low mountains	D5a, D5b, D5c, D5d
	High mountains	D6a, D6b, D6c, D6d

Level 1 (22 classes)	Level 2 (12 classes)	Level 3 (9 classes)	Level 4 (6 Classes)	Level 5 (4 Classes)	Level 6 (2 Classes)
Plains	Plains	Plains	Plains	Plains	Plains
Plains Low Hills	Plains Hills	Plains Hills	Plains Hills	Hills	Not Flat
Plains Hills					
Plains High Hills					
Low Hills	Low Hills	Hills	Hills		
Hills	Hills				
High Hills	High Hills				
Open Valley Low Hills	Open Valley Hills	Open Valley Hills			
Open Valley Hills					
Open Valley High Hills					
Mountains	Mountains	Mountains	Mountains	Mountains	
High Mountains	High Mountains				
Open Valley Mountains	Open Valley Mountains	Open Valley Mountains			
Open Valley High Mountains					
Plains Mountains	Plains Mountains	Plains Mountains	Plains Mountains		
Plains High Mountains					
Tablelands Plains	Tablelands Hills	Tablelands Hills	Tablelands	Tablelands	
Tablelands Low Hills					
Tablelands Hills					
Tablelands High Hills					
Tablelands Mountains	Tablelands Mountains	Tablelands Mountains			
Tablelands High Mountains					

Table 5.5 Generalisation of Landform Classes

CHAPTER 6

THE RESULTING LANDSCAPE CLASSIFICATION

6.1 Combining the landscape attributes

A landscape classification can be produced by combining the four main attributes of landscape discussed in the previous chapters. The unique combination of these attributes at any chosen level of generalisation forms the basis of individual landscape classes. Figure 6.1 shows graphically this combination process for generalisation level 3. Here, the resulting landscape classification has a total of 536 unique classes, and a total of 3115 discrete areas. It is not feasible to produce a key for this many classes, and therefore Figure 6.2 shows a key for only the top ten classes in total area. A classification code, L3 V3 N3 W3, shows the generalisations used. It means that generalisation level 3 was used for all four attributes: landform (L), vegetation (V), naturalness (N), and water (W).

It should be noted that the results near the boundary of the study area are inaccurate because the classification uses neighbourhood information. Near the boundary, the information beyond the boundary is not available, therefore the classification applied here is inconsistent compared to the centre of the study area. The extent of this inaccuracy is 10km (marked in Figure 6.2 by the inner square), since this was the extent of the largest focal radius used.

Figure 6.1 Combination Process

Main road and hydrology layers are added for geographical reference.

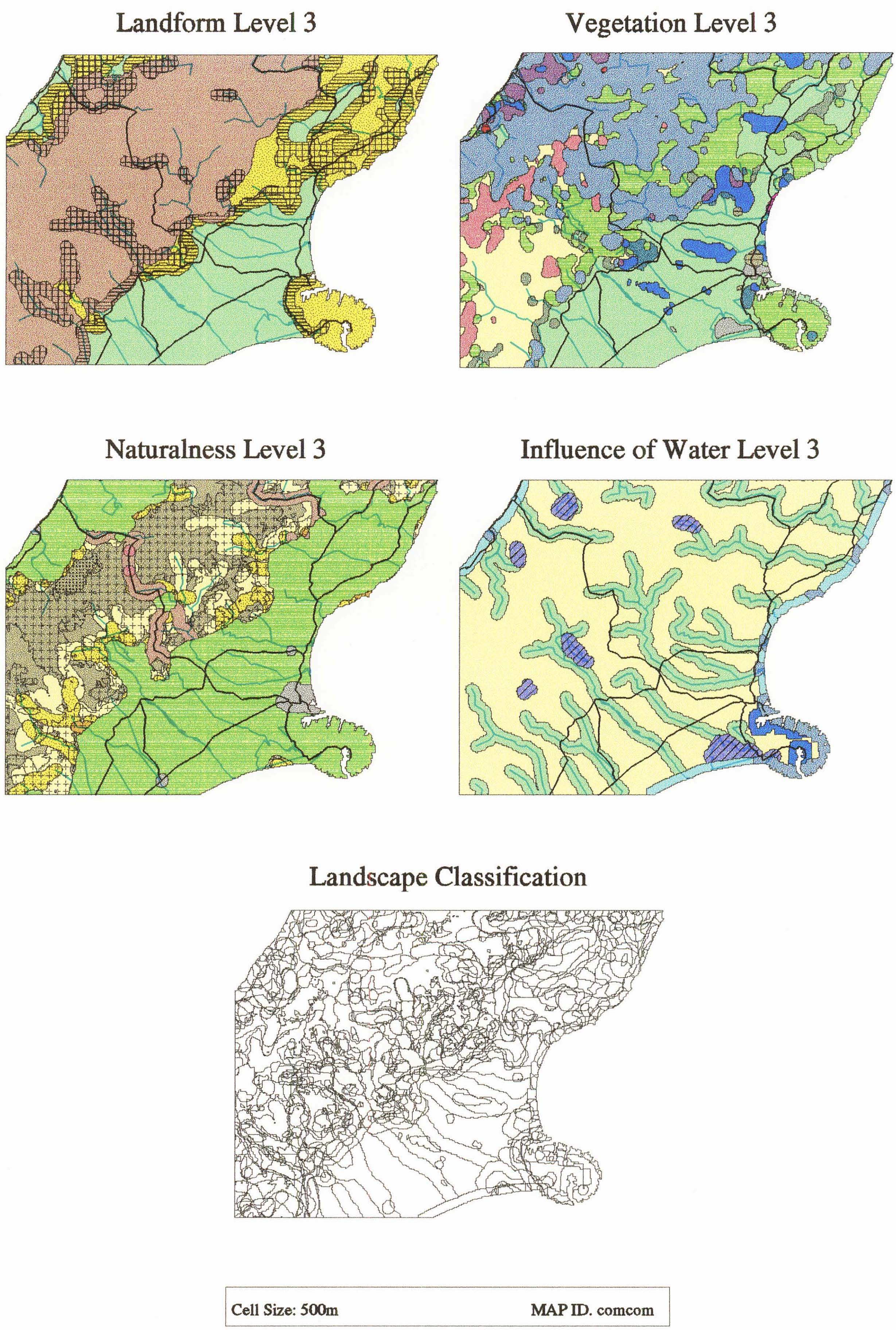
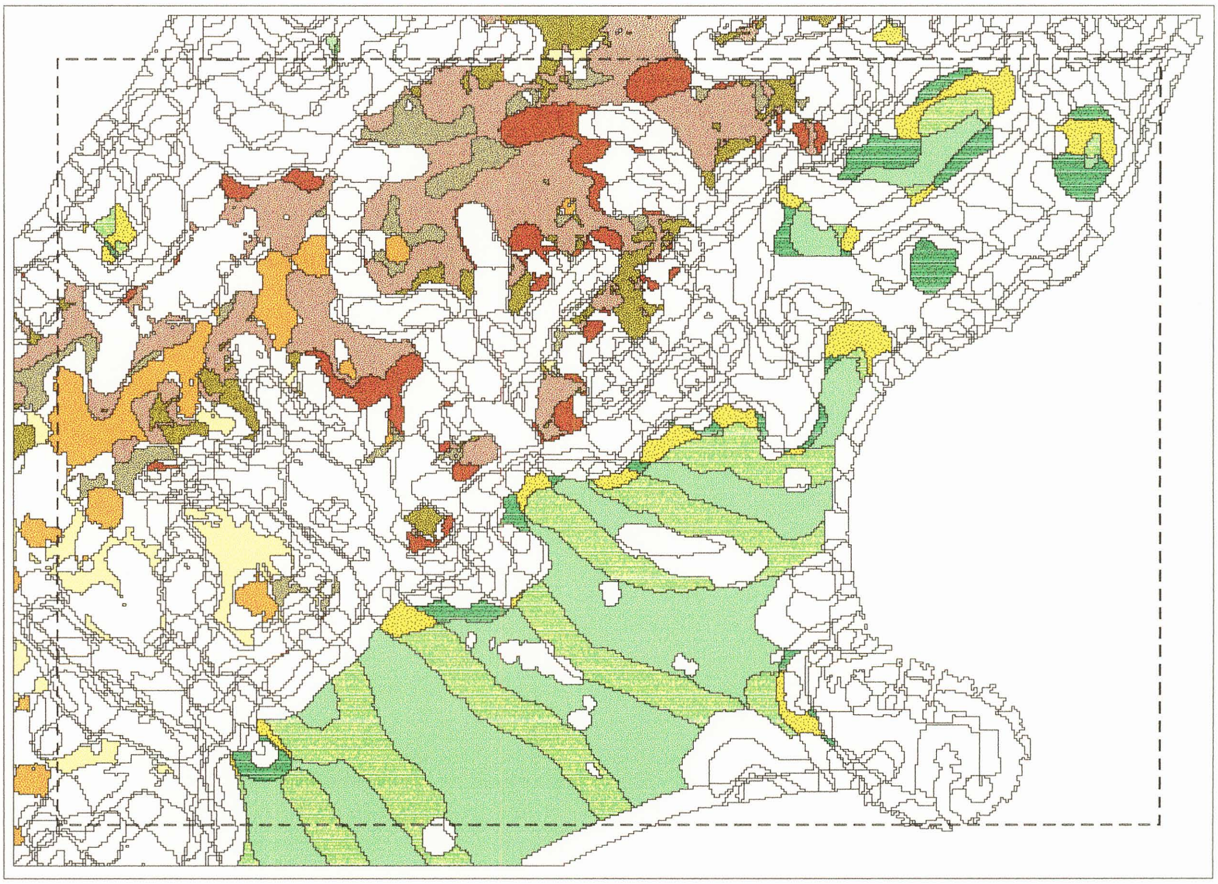


Figure 6.2 Landscape Classification L3-V3-N3-W3

Only the top 10 classes in total area are shaded



	Total Area (Hectares)
Plains/Pasture/Rural	297250
Mountains/Ind.Scrub/Backcountry few Buildings	210100
Plains/Pasture/Rural/River	202925
Mountains/Alpine/Backcountry few Buildings	63450
Mountains/Ind.Scrub/4WD.Track with Buildings	58425
Mountains/Ind.Scrub/Remote	54850
Plains Hills/Pasture/Rural/River	50075
Mountains/Tussock Grassland Scrub/Backcountry few Buildings	49525
Plains Hills/Pasture/Rural	49400
Mountains/Tussock/Remote	42650

Cell Size: 500m   Map ID. com3 - 3 - 3 - 3areac

6.2 Generalisation

Various levels of generalisation can be obtained by combining different generalisation levels of the landscape attributes. Figure 6.3 shows graphically six different outcomes, and the number of discrete areas and unique classes that have resulted from these. The most detailed level has 1302 classes, and the most general level has only 56 classes. It is possible to produce different outcomes by combining various levels of different generalisations. For example, if landform is considered more important than the rest, then level one of landform could be combined with less detailed levels of the other attributes, such as, a L1 V5 N5 W5, or a L1 V6 N2 W6 combination. Thus a range of classifications can be obtained that reflect different generalisations.

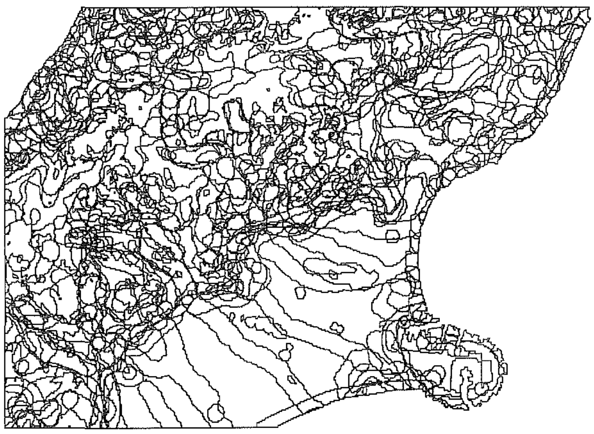
The number of classes identified in Figure 6.3 is only the number identified in the study area. The classification has the potential to identify many more. For level one, there is the potential for 356,224 classes to be identified (the product of 22 landform classes, 46 vegetation classes, 22 naturalness classes, and 16 water classes). However, this is only the tip of the iceberg, because considerable generalisation was required even to produce level one. Without this generalisation the classification would have the potential to produce approximately 6.7×10^{17} different classes.

The question then becomes: What level of generalisation is appropriate? This depends partly on the scale at which the classification will be used (ie. international, national, regional, or local), and partly on the purpose of the classification at the chosen scale. Selection of an appropriate level for a particular investigation might require preliminary cognitive and psychophysical research. The variability in such research between areas of the same class, will demonstrate whether the classifications are actually distinguishing the necessary subtleties. If two areas of the same class are perceived as being different in terms of quality, then the classification has not distinguished the necessary subtleties required. It is difficult to ascertain what subtleties are important at the different levels of investigation.

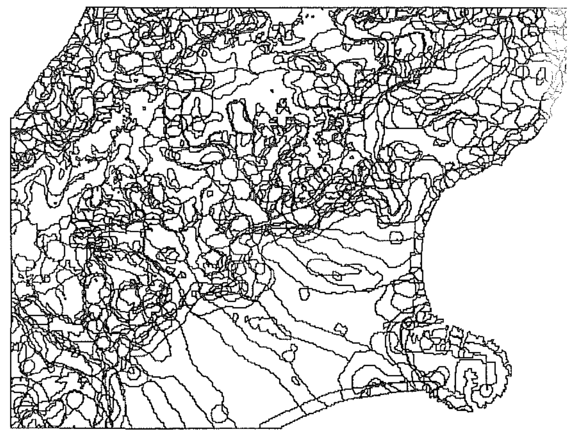
Figure 6.3 The Effects of Generalisation on Landscape Classification

(The figures in brackets are the number of defined areas and the number of unique classes, respectively)

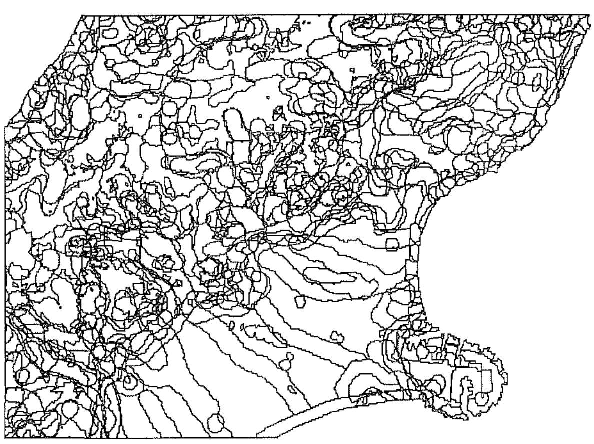
Level L1 V1 N1 W1 (4615, 1302)



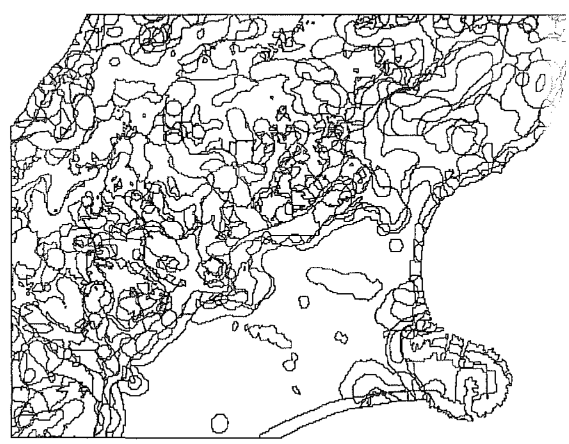
Level L2 V2 N2 W2 (3946, 900)



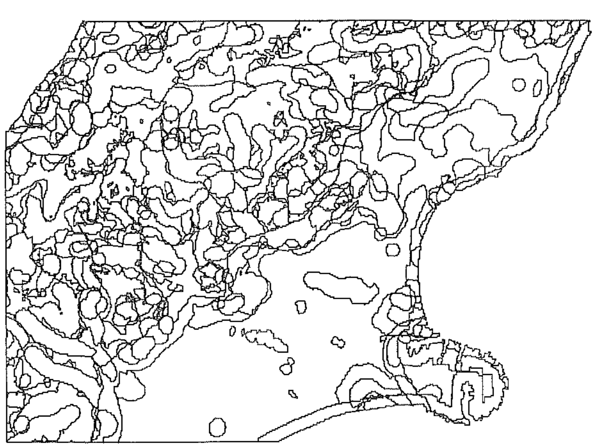
Level L3 V3 N3 W3 (3115, 536)



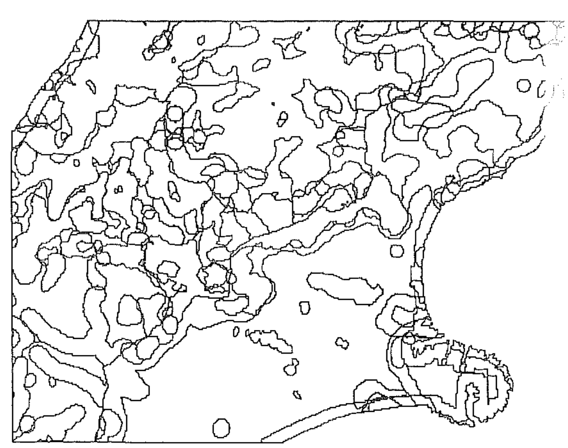
Level L4 V4 N4 W4 (1808, 256)



Level L5 V5 N5 W5 (1257, 144)



Level L6 V6 N6 W6 (593, 56)



Map ID. comcl1 –6

One can speculate on the appropriate generalisation for different levels of scale from two different approaches. It can be said that there is a certain limit to the number of classes that research can cope with, especially for doing psychophysical preference surveys. When surveying public preferences using photos, it is practically not feasible to ask people to rank more than thirty photos (Auckland Regional Authority, 1982). The preferences of different groups of photos can be linked together by having some photos that are common to each group. Therefore, many groups of photos can be used. However, there would be a practical limit to this. The other approach that can be used to determine the appropriate number of classes is to decide what landscape components are really essential in a landscape classification and then to include only these. Large components like mountains, hills and plains, coast, lakes, urban areas, and areas of forests, and grassland have a significant visual impact and therefore should be included. Also, distinctions based on naturalness should be kept since this is a known contentious characteristic. However, because of the lack of landscape content category research in New Zealand, it is difficult to reason with some substantive evidence about this. Trial and error (hypothesis testing) is the only scientific method for determining the appropriate level of generalisation.

Once the landscape attributes have been combined, there is further opportunity for generalisation using definitions based on more than one attribute. Some classes of one attribute may be considered unimportant when a class of another attribute is present. For example, in mountainous regions, it may be considered unnecessary to include rivers since the two are often associated with each other, and perhaps, because mountains are so dominating visually, rivers become insignificant. Zube has suggested the following:

> "As landform increases in dimension from flatlands through hills to mountains, land pattern decreases in importance as an element of visual quality. And, as landform decreases, the diversity of land pattern becomes increasingly important as an element of visual quality." (Zube, 1984b, p. 122)

With the combined coverage, such associations can be identified and generalised. For example, two landscape classifications can be developed with one having detailed vegetation information and the other not. A third landscape classification can be produced by mixing these two classifications so that general vegetation classes exist where there are mountains, and detailed vegetation classes exist where there are not. This has not been done but the option is there. It is possible to do this kind of generalisation before combination by doing a query on two different attribute coverages, for instance "if mountains are present in the landform coverage and river is present in the water coverage then change the water coverage".

Yet another way of generalising is to use a neighbourhood majority function. This replaces the central cell with the class that has the majority of area within a defined search radius. This has the effect of removing the smaller discrete areas. Just because a landscape class occupies only a small area does not mean that it is unimportant. Small size might contribute to its significance, especially if it is unique. It may not therefore be appropriate to generalise using such a filter.

It seems appropriate to generalise the individual landscape attributes before combining them to create a landscape classification. At that stage of the process, the number of classes is more manageable, and the problem is divided into smaller problems. Trying to develop a process that generalises a coverage that has the potential to have 6.7×10^{17} classes would be impractical.

6.3 The application of an agreement model

In section 5.4.2, a method for incorporating fuzzy set theory was demonstrated on a landform classification using entropy and agreement models. This appears useful for researchers for ascertaining the degree of certainty of the classes identified for different areas. Researchers would then be able to locate study areas where there is a high (or low if this is appropriate) consensus over their identification. It was concluded that agreement models are the preferred approach for landform classification. An agreement model could also be applied to the total landscape classification that has been developed. An estimate of the processing time required to do this for the study area would be approximately one week using a Sun Sparc 10 with three parallel processors. This would be if 45 different classifications were produced for each of the four different attributes, which, when combined, produce over four million landscape classification, each with six different levels of generalisation. The overall agreement model could be produced by multiplying together the proportion values of the agreement models of the four individual attributes, and therefore only 180 attribute classifications would need to be produced, rather than four million landscape classifications. This has not been demonstrated in this study because of the amount of processing required.

6.4 Validity

Now that a landscape classification process has been developed and applied to the study area, it is time to discuss the validity of this process and the resulting classifications. In sections 2.8 and 2.9 two sets of criteria were established for this purpose. They are based on general classification principles, and specific landscape criteria that consider the important characteristics of landscapes. In section 3.5.2, it was argued that using these criteria, along with consideration of GIS errors, was the most appropriate means available for this study for assessing the validity of the landscape classification process. A comparison between automated and manual classification, based on these criteria, will indicate whether there has been an improvement.

6.4.1 General classification criteria

Is the classification exhaustive and mutually exclusive?

The classification is exhaustive for the study area. However, without modification it would not be exhaustive for all areas, especially areas outside New Zealand. The classification has been designed for the study area. In some other areas there will be different compositions of landscape components that have not been catered for and modification would therefore be needed. This is particularly the case with vegetation, which has many different components.

The classification is mutually exclusive for the study area. As described in section 4.2.3 checks were made to ensure that an area could not be defined to more than one class.

Is the classification easily understood and applied?

It is questionable whether this classification is easily understood by researchers who have had no training in the concepts of GIS. To them, this classification may appear very complex. The actual fundamentals of this classification are not complex. Basically, it is centred around the use of a focal neighbourhood mean function, which in principle should be easily understood. This is applied to the components of landscapes to identify compositions. The classification is complicated by the lack of existing coverages for two of the landscape components: macro landforms, and indented coasts. Considerable processing has been required to create these. Also, most of the landscape component classes that were available needed considerable generalisation.

The programs written for this classification have been designed for research in order to enable maximum flexibility to explore different options. Once a classification process has been developed, and agreed upon, there are two matters that could be developed to improve the user friendliness of it. Firstly, a user friendly interface can be developed that makes the classification easy to use. Secondly, the type of NDDBs that the process is dependent on need to be standardised. If the topographical databases used standard labels and identified standard entities, then this would make it easier to develop a user friendly classification. Standardisation is now a major consideration of cartography and will most likely be widely applied (Buttenfield and McMaster, 1991). Once a user friendly classification has been developed, then a user would state the names of the input files and the automated classification would do the rest.

As for resources, the automated approach is very quick, therefore requiring minimum human input. The whole of New Zealand could now be classified within a few days. However, the classification does require the uses of expensive computer resources - both hardware and software, and expensive databases. It can be argued that computer resources are becoming cheaper all the time. Also, many resource management institutions already have the necessary computer and data resources, but they are not using them to full capacity. This classification can be run in batch mode in off peak periods. The most human intensive and computer intensive part of automation is usually developing the classification, not applying it. Now that a classification has been developed, it can be easily applied.

Is the classification repeatable and independent of the researcher?

Objectivity is one of the main advantages of automation. The classification process only requires the researcher to start it. The rest is done by the computer. It does not matter who starts the classification as the result will be the same. The design of the classification process and various decisions within the process have been subjective. The classification has as much as possible been based on theory regarding landscape quality. If another researcher designed a landscape classification process based on the same theory there may be some similarity in the resulting classifications. However, there is a lack of such theory so this is unlikely to be the case. The implementation of this classification is, however, totally objective. Automation requires totally explicit instructions, and these have been described comprehensively in this thesis.

Does the classification produce a hierarchical classification?

Yes, and this has been demonstrated (refer to Figure 6.3).

Is the classification flexible so as to cope with new interests and developments?

The classification is flexible because it is modular. The landscape attributes are assessed independently. If one attribute was discovered to be unnecessary, it could be easily dropped from the classification. Similarly, if an attribute was deficient or absent, it could be further developed or added. What also makes this classification flexible is that it is totally explicit. People can see how it works and can improve on it. In this classification, many critical parameter settings have been used in functions, and definitions. These settings can be easily changed. For example, a 3000m search radius was often used. This could be easily changed if cognitive research discovered this value to be deficient. A user friendly program could incorporate a menu interface whereby the user selects appropriate settings, thus enabling the classification to be very flexible. In section 5.4, the flexibility of the classification was demonstrated by producing many different outcomes reflecting different conceptual models of landforms. This can be done for the whole landscape classification process. The classification can also combine different attributes of various generalisation as demonstrated in section 6.2.

Does the classification recognize seasonal or cyclical change?

Such change should be consistent within a class. If a class changes in some areas with seasons, while in other areas it does not, then there are perhaps deficiencies in the classification. The attributes most affected by seasonal change are vegetation and water. Naturalness and landforms do not have cyclical changes. Most of the vegetation and water classes within the classification generally change consistently, however, there are some exceptions. Some exotic tree species are evergreen while others are deciduous. The deciduous species can change quite dramatically with seasons. The same can be said for some indigenous species. With the current vegetation databases that are easily available, it is not possible to distinguish accurately between evergreen and deciduous species. Perhaps individual forest companies could provide coverages of their own forests containing this information. It can be generally said that most forests in New Zealand are evergreen. The need for generalisation may make it impractical to distinguish deciduous and evergreen, and also distinguish between indigenous and exotic. Agricultural landscapes also change seasonally and this will be inconsistent for some classes. The classification does not distinguish between crops and pasture. A field that is being used for growing crops will change in appearance with the seasons differently to a field in pasture. It is not possible to incorporate this distinction with the currently available databases in New Zealand. If this information was available, the generalisation issue may make it impractical to distinguish these classes.

Water bodies also change seasonally - river flows change, lake levels change, and coasts can be rougher at different times of the year. This again can be inconsistent within a class. An obvious reason will be that these components are affected by climate, which in turn is affected by topography. Some rivers are snow fed, while others are not, and therefore flow differently during the spring thaw. The classification, in its present state, does not consider these subtleties. It is probably possible to incorporate them in an automated classification with present technology and knowledge, and even more likely in the future when databases hopefully become more sophisticated. However, one needs to question whether it is useful to include this additional information when it is necessary to generalise.

6.4.2 Specific landscape classification criteria

Does the classification incorporate landform, vegetation, naturalness, and water?

Yes, the classification was designed to do so. The more important question is whether relevant components of these four main attributes have been incorporated? This is difficult to say because landscape content category research has tended to produce results only at a generalised level, and has not provided much insight into how these main attributes should be further classified. Moreover, this research has not been New Zealand based. The classification is limited more by our understanding of landscapes than by GIS capabilities. This thesis has demonstrated the power and flexibility of

GIS for classifying attributes of landscape. When the important landscape components have been decided upon and substantiated by content category research, then GIS will probably be able to incorporate these in a landscape classification. This is the case with landform components as there is now a body of research that has investigated automated landform classification. Vegetation components are usually already identified and are available digitally to a detailed level so can therefore be easily incorporated in an automated classification. The same can be said for components of water. However, with naturalness, automation appears limited by the complexity of the available databases. Some classes of naturalness, such as tourism, mining, and heavy industry, cannot be identified adequately with the current databases available in New Zealand.

Is the classification based on the general public's perception of landscape attributes?

The classification attempts to classify from a general public's perspective by using appropriate levels of generalisation. Since the general public's perspective is not homogeneous, the classification is hierarchically structured so that it can be used for a range of different perceptions. The classification therefore addresses this criterion, even though it is not exactly known how the public perceives landscapes.

Is the classification based on an overall impression of an area perceived from a distance, and does it involve generalisation and composition?

All the components included in the classification have been of large enough size to be seen from a distance. This has been done by generalising or grouping various subtleties so that together these groups are easily visible from a distance. The classification not only incorporates compositions of the four main attributes, but also incorporates compositions of individual components within these main attributes. It also expresses the actual degree of composition (for example, 20% mountain, 50% hill, 20% scrub, 50% remote). The result is that the classification has the potential to identify an astounding number of unique composites - 6.7×10^{17}.

Does the classification recognize landscape as an experience from a multiple of perspectives obtained from movement and exploration?

The classification uses focal mean functions to incorporate movement and exploration. Manual classification techniques have tended to confine analysis of areas to the visual extent of the neighbourhood, while this automated approach has considered areas that are both visible and not directly visible. A neighbourhood focal mean functions can consider all the components in a set neighbourhood that are likely to be encountered through movement and exploration and thereby contribute to the landscape impression. The problem with considering movement and exploration is that the extent and behaviour of these are not known. In this classification a 3000m radius has been commonly used. Is 3000m appropriate? Theoretical understandings provide no answers for this. It could also be argued that exploration is not consistent over an area. People follow roads, and walk on established paths (though their visibility is not confined to these). Also, should areas far from the point of analysis be considered equally as closer areas? To develop an automated classification that considers these aspects, may be possible but would be complicated. Areas can be classified in terms of accessibility to paths and roads. As discussed in section 4.2.4, annuluses and kernels can be used with focal functions, and these can be used to weight different cells by the distance from the central cell. However, before such avenues are researched, the usefulness of the normal focal function should be first ascertained as this may be adequate. Only further research that uses this classification process will answer this.

6.4.3 GIS errors

As mentioned in section 3.5.2, because GIS is particularly powerful with spatial information, errors can be easily propagated. It is therefore necessary to assess these errors and to ensure that the classification is not invalidated by them. These errors can be grouped into three types - database errors, computational errors, and logical errors. These have been called GIS errors but in fact they can also be an issue to manual approaches.

Since perceived landscape is a fuzzy entity it does not permit precise measurement. If the boundary of the classes was changed 200 metres, it would not make too much difference, as it is not exactly known where the boundary should be anyway. This fuzziness has been demonstrated in section 5.4.2, and the notion of entropy and agreement models have been used to address this issue. When considering error, it is important to also consider the error that is acceptable. With landscapes, there is quite a bit of leeway. The amount of leeway appears to have never been discussed in the literature. To determine this figure, requires research that compares results from the use of several different classifications based on different spatial extents. For the time being, this figure is very arbitrary. It is assumed for this thesis that it is about 1000m.

6.4.3.1 Database errors

Common sources of error are associated with data quality. These have been classified as positional error, attribute accuracy, and spatial resolution. They can result from data collection, data input, data storage, data manipulation, and data output (Aronoff, 1991, and Bernhardsen, 1992).

Data quality is a major issue within GIS mainly because GIS uses many different databases (Chrisman, 1991). These databases can be easily shared between different users of GIS and can be easily manipulated. Such is the ease of data sharing and manipulation that it can be very difficult to determine the history of a database and its level of accuracy.

Positional accuracy

With positional accuracy, the databases are adequate. The accuracy of DOSLI's 1:250,000 topographical database is 150m for 90% of un-generalised points (Newsome, 1995). For Landcare's vegetation database, the accuracy is 200m (Newsome, 1995). The accuracy of the Ministry of Forestry databases, which are digitised from a 1:250,000 base map, is unspecified. For the digital chart of the world, the accuracy is specified, in the DCW metadata, as 7100m horizontally and 2000m vertically for the study area. The first two databases were the more important databases for the classification, and 200m is well within the 1000m assumed acceptable error limit.

It should also be considered that mapping agencies specify positional errors for a particular degree of certainty. For example, the DCW is based on the ONC map series, which is used for airplane navigation. The mapping agents, when specifying the positional error of these maps, had to consider the end use of these maps and safeguarded themselves (against lawsuits) by specifying high error intervals. There is not a vertical inaccuracy of 2000m in the New Zealand part of the DCW database, however, this is what has been specified. With landscape classification, a high degree of certainty is not required because human life is not at risk. It is therefore questionable whether the positional accuracy of the databases specified by their publisher is relevant for assessing the error of the resulting landscape classification, which needs considerably less certainty.

Spatial resolution

The spatial resolution of the database, for the purposes of this study, refers to the minimum size for polygons, or the minimum distance between lines and points. Many databases have been obtained from hard copy maps, which can only present a certain amount of information. If a map presents too many roads and polygons, it soon becomes unreadable. If there are too many structures within a certain area, then these are generalised to one structure. A minimum polygon size, and distance between lines and points are used. For DOSLI's 1:250,000 topographic database, this is specified comprehensively (DOSLI, 1984) and varies for different objects. The smallest polygon size is 0.1 ha. All rural roads are recorded because they are not close together in reality. Structures have been generalised if these were too close. Landcare's vegetation database contains polygons 0.05 ha in size, however the spatial resolution is specified by Newsome (1995) as 500 ha. The spatial resolution of these databases is adequate for the classification process because the classification usually uses 20% as a minimum presence within a 282 ha (3000m radius) neighbourhood. Unless accompanied by a sufficient number of small polygons, small polygons will not make a significant difference. Also, detail in the databases may be lost during vector to raster conversion since a 500m cell size was used. This is a computation error and is discussed later. Therefore even if the databases had higher spatial resolution, it is unlikely that this would make a difference to the classification outcome.

Attribute error

Attribute error is concerned with whether a cartographic entity (polygon, line, point, or cell) in a database is labelled correctly. These errors may be present because of cartographic error, or because the database is out of date. All useful databases will have attribute errors because of the need to generalise a complex reality. If an area has mostly forest but also has some grassland, a useful representation of this would be forest, which would not be entirely correct. Very general labels may be used to reduce attribute error (eg. to call the above area vegetation), but might not be useful. Attribute error is related to spatial resolution.

For the purpose of landscape classification, the accuracy of the 1:250,000 topographic database was sufficient for most entities. This database was current in 1990. Some changes may have occurred since then but would not be significant. If any major attribute errors existed, they would have been brought to DOSLI's attention. The only exceptions to this are the tracks, mines and structures layers. Tracks were missing in the more remote areas, and it would have been useful if more specific labels had been used for the mines and structures layer (as discussed in section 4.3.6).

One may question the attribute accuracy of Landcare's vegetation database. It was derived mainly from the Land Resource Inventory and field work done before 1981. Also, considering the nature of vegetation and recent agricultural and afforestation initiatives, many labels will be incorrect. This was discussed in section 4.2.2.

The Ministry of Forestry databases should be reasonably free from attribute error as they were developed in recent years. Supermap2 and the DCW were only used to identify towns and their populations. Supermap2 was derived from the 1991 census, and the DCW was revised in 1991 for the coverages that overlap the study area. Considering their limited use in the classification process, their attribute error would not affect the outcome significantly.

6.4.3.2 Computational errors

Considering the number of computations that can be implemented with a GIS and the degree of complexity of these, there is potential for error to accumulate and become significant. Often with user friendly GIS interfaces, it is easy to instigate a function but not know precisely how that function works and what calculations are involved. With many GIS functions there is a considerable amount of generalization and interpolation involved, and it is possible for the user to be oblivious to this.

Perhaps the most significant computational error in the classification is associated with the conversion of vector data structures to raster data structures. Vector coverages are usually a more precise way of representing landscape components, however, they are more difficult to spatially analyse than raster coverages. It would be very difficult to classify landscapes using only vector coverages. The effects of vector to raster conversion have been mentioned with regard to the spatial resolution of databases. The effect of this operation is dependent not only on the spatial resolution of the databases, but also on the geometry of the polygons and lines, and on chance. With the vector to raster conversion of polygons, the polygon class that contains the greatest area within a cell will become the class assigned to that cell. A long narrow polygon, which might occupy a large area, may be lost as neighbouring polygons might contain more area for each cell. Whether this happens depends also on how the grid overlays the polygon coverage, which is fairly random. It is possible for a polygon to be lost if it is just less than four times the cell size, but the chance of a grid dividing a polygon into exactly four equal size parts is small. Each part would be less than half the size of the cell size, and could be lost if its allocated cell was shared with just one other polygon.

With vector to raster conversion of lines and points in ARC/INFO, vectors are represented in a raster coverage by the overlapping cells. If there is more than one line or point that overlaps a cell, then the majority class (based on length of line or number of points) is allocated to that cell. In the classification process, only single class vector coverages were converted to raster coverages, therefore not too much attribute detail was lost. However, the vector to raster conversion only recorded the presence or absence of a class, so if there were two lines or points of the same class within a cell, the result was the same as if there was only one of these. Whether vector information is lost depends on how the grid overlays which is usually fairly random. The vector to raster conversion also spatially generalises vectors. For example, a twenty metre wide road in a vector coverage can become a 500m road in a raster coverage. However, this was not a problem as all components exerted a spatial influence of at least 3000m.

The other main source of computational error that exists in the classification is terrain interpolation. The representation of a terrain surface using TIN created some obvious errors with the landform classification process developed by Dikau et al. (1991), as discussed in section 5.3.2. There is always error associated with TINs because they interpolate and this can be significant in flat areas that have neighbouring relief. It was necessary to ensure that the new landform classification process was sensitive to this error. This was done by not having too many class that were dependent on subtle changes in slope.

It is difficult to ascertain the error associated with the slope measurements because it is not valid to compare the results with manually calculated results. As discussed in section 5.4.4, the method used for calculating slope affects the resulting slope calculations, and GIS slope functions use a different method to manual slope measurements.

6.4.3.3 Logical errors

A type of error that has been associated with landscape assessment is, for the purposes of this study, called logical error. Hamill (1989) provides an alarming account of these errors that have persisted for a long time within landscape research and have been largely uncontested. He used Leopold's method as an example (Leopold, 1969). Here, numbers were used incorrectly (spurious numbers) as they were assigned arbitrarily to denote different classes. These numbers are therefore nominal numbers, but they were used in mathematical operations as if they were cardinal numbers. The results of these operations were not only meaningless but varied depending on what number was assigned to which class.

Dearden (1980, p.52) also comments on this persistence of error. He states that "these measurement techniques contravene the theories of levels of measurement by using nominal or ordinal scales of measurement and then employing standard arithmetic procedures, such as multiplication and addition. In these circumstances, the methods become invalid."

Lowenthal (1978, p390) sums it up nicely;

> " adding together landform and landuse, panoramic and historical features is like summing apples, oranges, bacon and peppercorns."

Within a GIS environment, it is often necessary to represent words by numbers, especially within raster coverages. The mathematical manipulation of these numbers is very easy within a GIS and care is required to ensure that this is done appropriately. Logical errors do not exist within the classification, although it may appear so. Many operations in the classification are not arithmetical. They combine coverages rather than add coverages. An "and" operation was used instead of a "+" operation. With "apples" and "oranges", the effect of a combine operation is to get a new coverage with a class of "apples and oranges", not some spurious attribute value. When arithmetical operations were used they were done within a class rather than between classes. For instance, the focal mean function was applied to single theme coverages. There is nothing wrong with saying, " an apple plus an apple equals two apples".

6.4.4 Manual versus the automated approach

Manual and automated approaches can be compared by discussing them in relation to the general and specific criteria. The automated approach has been subjected to this and now it is appropriate to do the same with the manual approaches.

Concerning general classification criteria, manual classifications may be mutually exclusive, exhaustive, hierarchical, and able to incorporate seasonal or cyclical changes. However, they are generally not explicit and cannot produce repeatable results that are independent of the observer. This is because they tend to be intuitively based. Some degree of replication may be possible if people have had similar training and some objective criteria are used. However, in relation to an automated method, they are no match. At a national or international scale, there is unlikely to be repeatability with manual methods. It is arguable whether manual methods are easily understood and applied. How can a method be understood and applied if it is intuitive? This must lead to confusion as practitioners seek confirmation on the exact nature of landscape components and composition. Also, manual methods must be very time and resource consuming as many landscape need to be directly observed. Manual methods may be flexible because they are vague, but how can new understandings be gained when it uncertain how the method was implemented.

Regarding specific classification criteria, the manual approach does have some credibility, but it is debateable whether this is more than the automated approach. There is no doubt that the direct perception of landscapes in the field will give a better indication of the nature of a landscape than a computer. However, a landscape in one area has to be classified in relation to the landscapes in other areas. The practitioner therefore has a massive amount of information that needs to be considered. It is questionable whether this can be done manually over a large area, such as the size of New Zealand, or the world.

Manual classification of large areas may not involve field visits to view the entire study area. Representations of reality, such as maps, are often used instead, but the information required is unlikely to be all on one map. Can humans analyse effectively several maps at a time to get an overall impression of landform, vegetation, naturalness, and water? This must be a tedious, and challenging task. It is likely that the manual approaches that use maps have to separate landscape into main attributes, like the automated approach, to make the classification manageable.

One task that manual methods do well is the recognition of pattern. For example, the recognition of a valley floor (relief-flat-relief) can be done easily manually. However, to do this automatically, involves considerable processing. The same can be said for the identification of compositions. As commented in section 5.3.2, some other patterns are probably more effectively recognised using manual methods, for instance the topographic patterns associated with conical volcanoes.

Concerning the notion that landscapes are experienced from movement and exploration, the manual method has been deficient in the past. Most manual methods appear to use direct visibility, and according to the criteria this is inappropriate. Even if manual methods did incorporate exploration, there would not be the resources available to fully explore the whole study area. It would be necessary to rely on maps produced by surveyors that have already done the exploration. The question then returns to whether humans or computers are more effective at analysing maps?

Manual methods require that the practitioner divide the study area into analysis units (usually areas of visual enclosure) before analysis of landscape character begins. This is necessary to make the task manageable. With automation, the GIS divides the study area and the analysis unit is kept very small in comparison. The dramatic subdivision of the study area (into many cells) effectively enables the analysis of landscapes from point perspectives. This is more appropriate and would replicate landscape perception, which is also done from many different points.

In conclusion, it can be said that the automated approach is superior in terms of general classification criteria because it is explicit and repeatable. Regarding the specific criteria, both manual and automated approaches have their pros and cons. Overall, because explicitness and objective repeatability are essential ingredients for a classification, then automation is a significant improvement for landscape classification.

6.5 Applications

6.5.1 Frame of reference

The most significant application for a landscape classification is as a frame of reference for communication within landscape research. This has been discussed in section 2.6. From the results, such as presented in Figure 6.2, it can be seen that if someone was researching landscapes within the study area, then they could use the classification in a variety of ways: for description, mapping, and inventory purposes. Firstly, a particular location can be described by the class within which it is located, or a region can be described by the predominant landscape classes that exist within it. Secondly, all localities with certain landscape characteristics can be located and mapped using the classification. Thirdly, inventories can be created showing areas and numbers of occurrences of landscapes satisfying certain conditions. With GIS, it very easy to generate information on the total area of different classes within a specified region. Within the GRID module of ARC/INFO, a value attribute table is generated that counts the total number of cells that exists for each class. The area of each class can then be calculated by multiplying this number by the area of each cell (25 ha in this study). This has been done in Figure 6.2 for the study area. It can be said, for example, that at generalisation level L3-V3-N3-W3, there is 54,850 hectares of mountain / indigenous scrub / remote landscape within the study area. When the classification is in vector format, the number of occurrences of different classes can be calculated in ARCPLOT simple by selecting them. The area of single polygons can also be easily ascertained.

For this type of communication to be effective on a national or international scale, standardisation is required for the different levels of generalisation, as well as the labels that describe each class. The labels used in Figure 6.2 have been chosen because they describe the actual class. This is a useful coding system but can distort people's interpretation of the class. Each class should be interpreted with the underlying explicit definitions of these classes. By using a descriptive label, people may be inclined to use their own conceptual definitions of these labels to interpret these classes. Non descriptive codes could be used, for example A5R6, but these would make it difficult for people to become familiar with the classification.

6.5.2 Determining uniqueness

As discussed in section 2.6, uniqueness has been used for assessing the value of different landscapes. However, there is not a clear relationship between landscape value and uniqueness. If a landscape is unique and considered ugly, then it is of little value. However, uniqueness can make an average landscape important, or a beautiful landscape extremely important. Information on uniqueness is therefore sought after by landscape practitioners.

Strictly speaking, if something is unique then it is the only one of its kind, therefore, something is either unique or it is not. However, whether something is unique is often expressed on a relative scale, as implied in the phase, "quite unique". The concept of uniqueness has evolved although the term, "rarity" might be more grammatically correct. Uniqueness appears to be on a scale from absolutely unique to very common.

Uniqueness of a class can be expressed using the percentage of the total area a class occupies. This would depend on size of the analysis area (scale), and also on the level of generalisation in which the landscape is perceived. With the classifications that have been generated, it is now possible for people to be explicit about these two considerations, which in turn, could lead to more constructive debates within planning courts. Inventory statistics can be divided by the total area of analysis to give an impression of the uniqueness of that class within the study area. People can now question whether that level of generalisation is important, and whether the extent of the study area is relevant. As discussed previously, the level of generalisation can be easily changed. The same can be said for the extent of the analysis area.

In the above example, the study area was the extent of analysis. With GIS, it is easy to change this extent of analysis by setting an analysis mask. This has the effect of "cutting" the coverage to the required extent. For example, if areal statistics were required just for the Banks Peninsula region, then a coverage that just shows the extent of this region could be used to set the extent of the analysis. A new classification coverage of Banks Peninsula can then be generated by simply entering the command, "coverage (Banks Peninsula) = coverage (study area)". This new classification coverage of Banks Peninsula will automatically have a value attribute table with areal information for each class. Uniqueness information can therefore be generated for all levels of scale - local, regional, national, etc. However, the analysis has to be confined to the extent of the available classification, which at the moment is only for the study area.

With GIS, it is relatively easy to change the extent of the analysis in incremental steps, and for each incremental step generate uniqueness information. To start with, an analysis window of one cell can be used, which can be located over the point of interest. Areal information for a particular class would be either zero or 25 ha (for 500m cell size). The coverage that defines the extent of analysis can then be expanded by one cell in all directions, thereby, generating an analysis window of nine cells. Areal information can then be generated again. This procedure can be repeated hundreds or thousands of times automatically until the size of the analysis window is more than required. Such analysis would be reasonably quick for an area that was the same size as the study area. The resulting information would enable, for a particular class, the uniqueness to be plotted in relation to scale. Instead of asking the question, "is this class at this location unique at this scale?", it is now practically feasible to ask the question, "at what scale is this class located here unique?". However, before these questions can be answered, the question, "what is unique?" would first have to be answered.

The same incremental uniqueness analysis can be calculated for different levels of generalisation, for a given extent of analysis. Uniqueness can be calculated for generalisation level one, and then for level two, and so on, until all generalisation levels have been considered. The uniqueness of a class, for a particular location, can then be plotted against generalisation, and the question that can be asked is: "At what level of generalisation is this class, at this location, for this extent of analysis, unique?". By combining this information with the analysis of scale, as described above, a very interesting model of what is unique would develop. However, before it is worthwhile to develop these models, which are not pushing GIS technology to its limit, it is first necessary to agree on a landscape classification, with its different levels of generalisation.

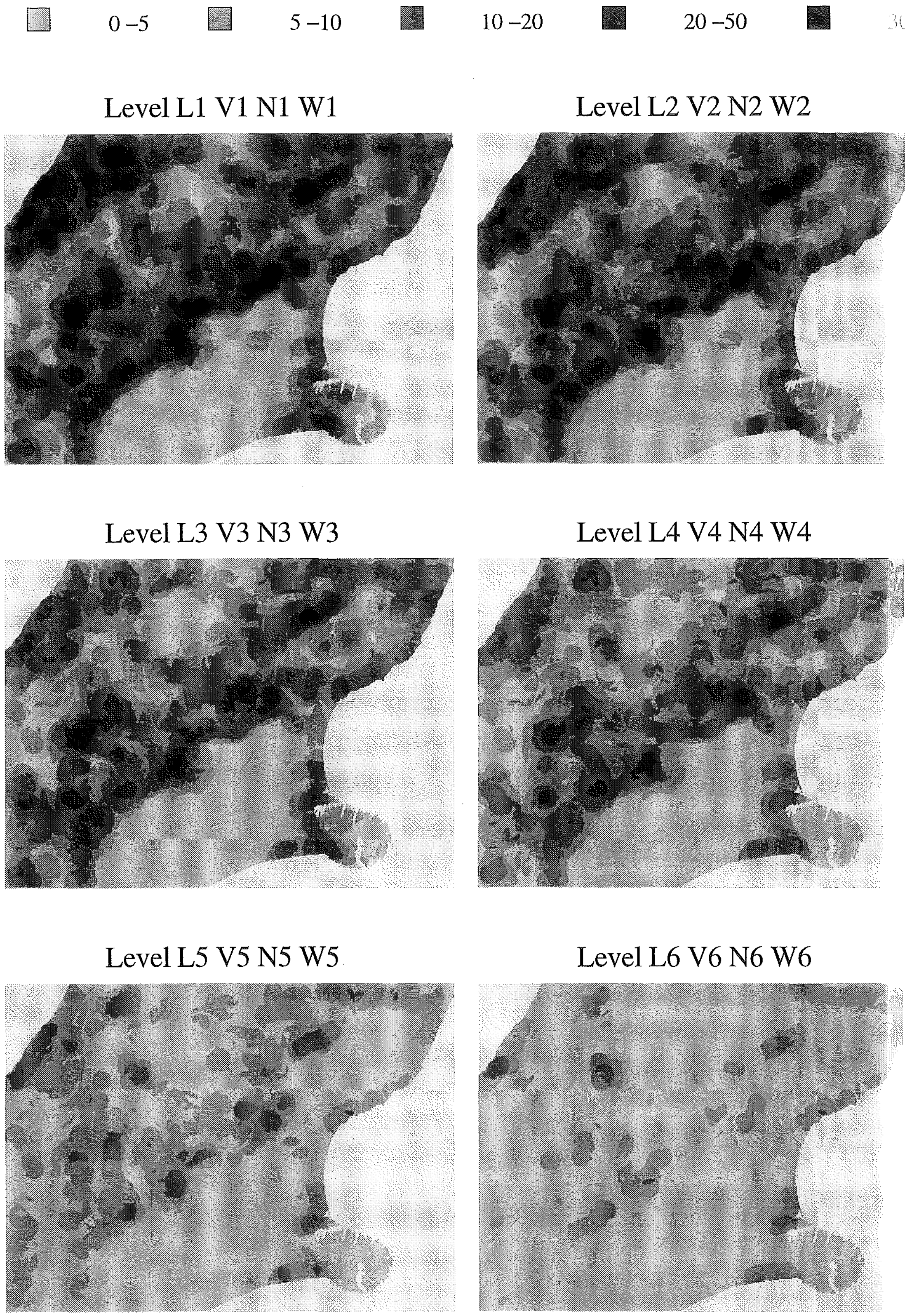
6.5.3 Assessing landscape variety

Landscape variety has been used for assessing landscape quality by the Ministry of Works and Development (1983 and 1987), however, the validity of this has not been proven (Arthur et al., 1977). Whatever the case, there is a demand for information on landscape variety. Variety can be defined as the number of unique classes within a given area. In the past, landscape practitioners have not used quantitative techniques for assessing variety but have used a more intuitive approach. With GIS, variety can be calculated by using a focal variety neighbourhood function (Berry, 1993). This is similar to other focal functions used in this study but assigns the number of unique classes that exist within the analysis window to the central cell that is being processed. The analysis window can be of any size or shape. Figure 6.4 shows the effect of such a function on the landscape classification developed in this study. The analysis window was a 5000m radius circle. If the size of the NAW changes, then so will the variety. Intuitive means for assessing variety are implemented very subjectively, while with GIS, once variety has been defined and a landscape classification agreed upon, then it can be implemented objectively. Since GIS uses an explicit definition of variety, then this definition can be questioned and developed, as our understanding of the nature of landscapes improves. Figure 6.4 also shows the effect of generalisation on variety. As expected, variety is very dependent on this. Obviously, the more detailed a classification and the greater the search radius, then the greater the number of classes that are likely to exist within a given area and therefore the more variety. This figure demonstrates that when practitioners are considering variety, they also need to consider the level of generalisation.

6.5.4 A basis for further manual classification

If automatic landscape classification is considered inadequate, a hybrid of automatic and manual classification could be considered. As discussed in section 6.4.4, both the manual and automatic approaches have their advantages and disadvantages. Perhaps if manual and automatic methods were considered together, then these disadvantages may disappear.

Figure 6.4 The Effects of Generalisation on Landscape Variety



Cell Size: 500m	NAW radius: 5000m	Map ID. comcl1 –6variety

Automation has the advantage that a considerable amount of information on various attributes and over a large area can be treated consistently. With manual classification, it is doubtful whether a practitioner, or a team of practitioners, can match this consistency. However, the exact boundaries of classes may be determined better using manual approaches, especially if areas of visual enclosure are considered a better basis for analysis than focal means. With visual enclosures, the boundaries of classes often correspond with the crest of ridges. While with focal means, this is not so. What is best depends on how one thinks landscapes are perceived. It has been argued in this thesis that landscape perception is derived from movement and exploration, which is affected but not completely restricted by ridges. If visual enclosures are considered more appropriate for landscape classification, an automatically generated classification could form the basis for a classification, which could then be altered manually. The boundaries could be manually edited to ensure they match catchment boundaries. This could be done in digital format using GIS editing capabilities.

As discussed in section 6.4.4, manual approaches appear more appropriate for identifying particular patterns or shapes, such as, conical volcanos. It could be considered that some individual landscape components can be identified better manually, but the spatial extent and composition of these components can be calculated better using GIS. Such a hybrid approach is feasible, however, the landscape components would need to be made available in digital format.

6.5.5 A means for understanding landscapes

An interesting spinoff from trying to classify landscapes automatically is the increased understanding that is obtained about the nature of landscapes. To develop the classification presented in this study, required a considerable amount of "simulation" that considered the effects of using different components, spatial extents, and other parameters. By assessing the effects of these, the importance of different components and parameter settings became apparent. The ability to perform hundreds of classifications is a major advantage of automation and is useful way of exploring the nature of landscapes.

Automation also requires detail about landscapes to be explicitly addressed. With manual approaches, based on intuition, many details have not been considered explicitly. As a result these details have not come out into the intellectual arena and been openly discussed.

It has been argued previously that the best method for validating a classification is to use it. If an unacceptable discrepancy becomes apparent within a class, then this may demonstrate that the nature of landscapes is more complex than the classification portrays. In this way, a classification evolves and an increased understanding of the nature of landscapes is obtained. Automation facilitates this evolution because a classification can be easily redesigned and reapplied.

CHAPTER 7

CONCLUSIONS

7.1 Can GIS classify landscapes?

This thesis demonstrates that GIS can classify landscapes using characteristics that the theoretical literature appears to consider important.

The classification process can be summarized as follows:

1. The selection of the four landscape attributes.

2. For each landscape attribute:

    a) The identification of important components from existing NDDBs using attribute and spatial generalisation. Indented coast and landforms had to be conceptualised using complex routines.

    b) The determination of the spatial influence of each component using a focal neighbourhood mean function.

    c) The identification of classes using complex conditional queries on the spatial influence information.

d) Generalising the classes to six levels by grouping the classes using simple attribute generalisations.

3. Overlaying the landscape attribute classes for each level of generalisation and using the unique combinations as landscape classes.

The classification of landscapes using GIS has considerable advantages over traditional manual methods. The most important advantage is the ability to be explicit and repeatable with complex definitions. Automated classification is also flexible because operational definitions can be easily changed and regions can then be easily reclassified. Once an automated process is developed, it is also considerably quicker to apply than manual methods.

A major issue with developing any classification is validation. A set of criteria using general classification criteria, and specific landscape criteria appear useful for this. Operational errors, and a comparison with previous methods have also been considered. Specific landscape criteria consider whether the classification incorporates the important characteristics of landscapes that affect quality. The classification has generally met these criteria. The classification can now be challenged in terms of whether the criteria are adequate, and/or whether these criteria were adequately applied. Because landscape theory is a "theoretical vacuum" (Appleton, 1975b), it was only possible to include generalised characteristics in the landscape criteria. These criteria need to be met in order for the classification to be valid, however, they are not sufficient to judge whether the classification is actually valid. Validation needs to be completed through independent research.

This research highlighted many deficiencies in the current theoretical understanding of the nature of landscapes, particularly in relation to their composition. These deficiencies have been stated previously in a more general manner by Steinitz (1993), and Mitchell (1993), however, GIS forces one to be very specific about them. It can be considered an advantage of automation that, because GIS needs everything to be explicit, all details need to be addressed.

> "[A] computer-based information system serves to highlight deficiencies in theory, not to hide them, and its use must be seen as an aid to better understanding" (Duffield and Coppock, 1975, p.141).

Specific deficiencies identified in this research relating to the nature of landscapes were:

- What are the exact landscape components that are important to landscape perception and the nature of their contribution?

- What are the appropriate distant decay functions from a given point for each of the landscape components?

- What are the important component compositions?

The lack of landscape theory makes it difficult to substantiate many decisions made in the classification process and therefore the classification is subjective. This does not mean that this classification is inappropriate. For theory to develop, a classification is required to act as a frame of reference. A classification needs to be developed as best as possible with existing theory. Where this is deficient, assumptions should be made. As theory develops, the classification then needs to be revised if assumptions are proven to be incorrect, and it is in this way that a classification evolves.

It appears that GIS can provide an effective research tool for developing landscape theory. A range of possible landscape classifications can be developed and these can then be tested through psychophysical and cognitive research. In this way, an increased understanding of the nature of landscapes can be obtained and the above questions might be answered.

Entropy and agreement models can be used to cope with the gaps in theoretical understanding and the fuzziness of landscapes. This was demonstrated and discussed in section 5.4.2 regarding landforms. It was concluded that the use of an agreement model is more appropriate than the use of entropy for this purpose. This model provides information on the certainty of a classification enabling a researcher to select areas for future investigation. If a researcher wants to research a landscape that has a high consensus (or perhaps a low consensus) over its identity then information on this can be obtained. For example, an evaluation study using psychophysical methods will probably use high consensus areas. However, a researcher may be interested in low consensus areas to develop understanding of landscape character.

Focal neighbourhood functions, in particular the focal mean function, are valuable tools for landscape classification and offer an alternate approach to conventional methods. They can effectively be used to calculate the spatial influence of components from thousands of different points, since relatively small cells (500m) act like points. These spatial influence measurements can then be grouped into classes. Following Zube, Sell, and Taylor's (1982) emphasis that movement and exploration are important parts of landscape perception, it has been argued in this thesis that the focal functions can incorporate these. Focal functions are therefore theoretically and practically preferable to a visibility function.

A significant part of this study has been developing a method for identifying landforms as this attribute had not been conceptualised in existing databases for the study area. This is a reasonably complex operation but, once developed, it is an effective means for classifying macro landforms. A significant problem with automating landform classification is the measurement of slope. It appears that when

GIS is used for measuring slope, the results will be different to manual slope interpretations. This is because methods for manual slope measurements have not been explicitly stated, in terms of scale and location of measurements. An explicit method can be applied with GIS, whereby the scale and location of measurements are stated. Since GIS and manual slope measurements give different results, then slope thresholds used in manual classification are likely to be inappropriate for automated classification. When seeking a GIS measured slope threshold to distinguish flat and non-flat areas, a slope threshold of four percent, rather than Hammond's eight percent, was the most appropriate. A four percent slope threshold measured with GIS had 91% agreement with a seven degree threshold based on the LRI's manually measured slope information. Other problems identified with existing automated landform classification methods were the mixing of macro and meso objects, and the way the spatial influence of relief is determined. These problems can be resolved, and this was demonstrated in a new landform classification developed in this study. It appears that automating a manual classification process, such as Hammond's, may be a good place to start when developing an automated classification but modification is often needed, and improvements can be made to existing manual processes.

This thesis has also revealed options for future research that could improve the existing classification. As previously stated, more cognitive research on the nature of landscapes is needed, but there are also GIS options that could be investigated. These being the use of annuluses and wedges for specifying the extent of the NAW, the use of kernels for incorporating an appropriate distance decay function for landscape components, the identification of particular landforms, such as conical volcanoes, and the use of more complex databases as they come available, eg. the 1:50,000 topographic database. The use of visibility functions could also be considered for providing extra information, but it is doubtful whether they will be more appropriate than focal neighbourhood functions. Perhaps they could be used in combination.

Duffield and Coppock (1975, p.146) made the following comment concerning their computerised landscape assessment package:

> "Perhaps the primary deficiency for landscape assessment lies in the system's inability to cope with the spatial composition of landscape, as opposed to its resource content. It is not unique in this failure; indeed, nearly all existing procedures of assessment of landscape have proved incapable of dealing with this vital aspect of the appeal of landscapes. Clearly the appreciation of landscape is primarily aesthetic and derives as much, if not more, from the spatial relationship of visible resources as from their mere presence in the scene."

GIS can express complex spatial relationships of landscape components, especially if annuluses and wedges are used for defining the NAW. However, the problem is that there is no agreement on what spatial relationships are important. Before more complex spatial relationships are expressed with GIS, it is necessary that our understanding of the nature of landscapes is improved, and this needs to be based on an existing classification.

7.2 Implications for databases

This research used national digital databases, in particular, DOSLI's 1:250,000 topographic database and Landcare's vegetation database. Since NDDBs are relatively new and are still in the process of being developed, it is worthwhile to comment on their worth and possible improvements.

This study has demonstrated that these databases are particularly useful for complex spatial analysis of large areas. These databases in themselves contain a lot of conceptual information, for example towns, roads, etc. However, with spatial analysis, additional concepts can be identified that are useful for resource planning. This has been demonstrated by identifying not only variety and uniqueness, but also certain components of naturalness and landform. In the context of landscape description, it is important that these databases are seen as more than raw material for cartography, and that their true worth is realised.

Despite the fact that digital databases are of such worth, this research has also demonstrated that at the moment there are lost opportunities with digital databases that need to be realised. Perhaps the greatest opportunity that is being lost is that they are not being fully utilized. This is related to access. Based on the experience encountered in this study, access to databases is currently inhibiting their use, and the greatest barrier to access is their cost.

Further opportunities can be gained from digital databases by making them larger and more complex. If this was the case, then they could still be easily analysed. The databases used in this study were fairly large (some were about 20 megabytes). However, this did not pose a major problem with the hardware and software. These databases were significantly reduced in size when converted to raster coverages with a 500m cell size. They could then be easily manipulated and duplicated without lengthy processing times or significant hard disk storage problems. The hard disk space used in this study was not more than 600 megabytes, and this included all the postscript and graphic files generated for presenting maps. The implications of this are that even larger and more complex databases can be used.

Currently, the information in topographical databases in New Zealand is generally limited to the amount of information on hard copy topographical maps. GIS can cope with far more information than this, even over a large area. If there is too much information, this can be easily generalised, but if there is not enough then this can severely restrict its application. The building layer is an example where more information would be useful. Currently, there is very little attribute information associated with this layer. If there had been more, then additional subtleties relating to naturalness could have been included. This study would have benefited if this layer had identified different types of structures rather than just having a single general category called "structures". Attribute information, such as the size, age, and even the number and type of occupants could have enabled a tourism class to be identified. Enhancements to the urban layer, and the mines layer would have also been beneficial. In this study, several different databases had to be used to get information on urban areas. The actual populations were obtained from Supermap2, medium size towns from the DCW, and the large urban areas and very small towns from the topographical database. It would have made analysis easier if all this information had been available within one database. As discussed in section 4.3.6, the mines layer was deficient because it contained only one class of mines. From this it is not possible to distinguish whether the mine is underground, open cast, in use, or abandoned.

General purpose databases (secondary) should contain as little generalisation as possible, and instead leave generalisation in the hands of the users of the databases. GIS is very capable of generalisation. The databases, however, have often been already significantly generalised by cartographers. It is recognised that generalisation is necessary for developing these databases, however, where practically feasible this should be kept to a minimum for users who can use powerful GIS. Inconsistency in the databases is the GIS user's nightmare. Generalisation when applied unevenly will result in inconsistency in the database. Automated generalisation is generally consistent in its application. If a user can apply their own generalisations then they can be sure that this is done consistently, and to the required level. A range of databases with different degrees of generalisation may, however, be appropriate for the benefit of others.

Standardisation of spatial and attribute data within NDDB is absolutely critical and has been highlighted by this study. Rule based automation needs to use consistent databases. Otherwise, it can produce spurious output. If inconsistent databases are used, then automatic processes need to be considerably complex to cope with the diversity of possible data. The simple solution is to ensure that the databases are standardized. It appears that this concern is already being addressed. Land Information New Zealand (LINZ) has released a series of publications specifying the standards that they will use (LINZ, 1985, 1987a, and 1987b). This covers standards for labels, geographical referencing, and measurement and inclusion of area size. Standardisation is also being attempted at an international scale (Murcott, 1995). These initiatives are important for the development of automated geographical abstraction. They, however, need to be applied by all agents that are providing digital databases, and not just the main mapping agents, such as DOSLI. For example, if place names had been standardised between different databases then this would have saved considerable inconvenience. Preferably DOSLI and Supermap2 should be using the same place names.

Databases should also be available in raster format. This study has demonstrated the power of spatial analysis using a raster format. It is doubtful whether such analysis could be done by using only vector coverages. Mapping agents are supplying mainly vector coverages but it is raster coverages that are the most useful for spatial analysis. This did not pose too many problems for this study since a vector to raster conversion function was available. However, this used a powerful GIS that is not available to many GIS users. Many cheap GISs are raster based, therefore, databases should be made available for these systems. A DEM is difficult to obtain accurately from a vector contour coverage as this study demonstrated. It would be preferable if mapping agents supplied a range of DEMs with different cell sizes then the task of creating an accurate DEM from contours would not have to be repeated by different users.

7.3 Implications for GIS

This study has demonstrated that complex geographical abstractions can be implemented within GIS to produce coherent, meaningful results. GIS can use a range of representation techniques that express classification, association, generalisation, and aggregation. This has enabled GIS to express complex structural geographical meaning. The challenge to do this for landscapes was mostly with regard to incorporating generalisation and association. Landscape classification is very much an exercise in generalisation as there is a considerable amount of digital data on different landscape components. Spatial and attribute generalisations using relatively complex conditional queries appear appropriate to bring the complexities of reality down to a manageable level. Simple attribute generalisation based on the grouping of classes is particularly useful for developing a workable hierarchy of levels whereby information relating to classes at specialised levels can be easily linked and applied to classes at a general level. The use of focal neighbourhood functions has been particularly useful for expressing association between components, which in turn has enabled compositions to be identified.

An important method for experimenting with different operational definitions is simulation. GIS with its associated macro languages can simulate complex processes thousands of times within a short period. This has been valuable for accurately "tuning" operational definitions of complex geographical concepts.

With hardware and software, there is always room for improvements. As hardware becomes faster, there is more processing demanded because new computer intensive applications become apparent. Since the process developed in this study was relatively quick, it became feasible to apply the classification several times using different definitions. This consumed significant amounts of CPU time, and a faster hardware platform would have made this task easier. The software could be improved by the removal of limitations on the number of coverages that can be used within a function, for example the majority function in GRID. Also, the removal of limitations on coverage sizes, such as, the number of arcs and size of attribute tables, would also be advantageous. Such limitations might be appropriate when people are experimenting with GIS and are not too sure of outcomes, but can be unwanted when large processing tasks are required. With the development of large national and global databases, extremely large tasks will be expected from GIS.

This study has demonstrated the power and ease with which GIS can manipulate and analyse information traditionally obtainable from hard copy maps. Most measurements that can be obtained manually from a map can now be done automatically. The automation of cartographic analysis significantly enhances the geographer's analytical opportunities. Extensive regions can be analysed, and this analysis can be quite complex. This has been demonstrated in this study with automated landscape classification. The ability of GIS to analyse large areas means that Geographers can realistically, quantitatively examine issues at a national scale, rather than be confined to regional or local scales because of practical constraints. For many environmental issues the national scale is important, and the landscape issue is an example of this. It can be argued that global analysis is also important, and it is only a matter of time before this type of analysis will be available.

This study has revealed how GIS can provide a platform from which a comprehensive landscape classification can evolve. Such a classification can be used for effective communication between landscape researchers, and could contribute to the development of consensus among researchers on landscape issues.

REFERENCES

Amedeo, D. Pitt, D.G. and Zube, E.H. 1989 Landscape feature classification as a determinant of perceived scenic value. Landscape Journal, Vol. 8, No. 1: 36-50.

Appleton, J. 1975a The experience of landscape. J. Wiley and Sons, London.

Appleton, J. 1975b Landscape evaluation: the theoretical vacuum. Transactions of the Institute of British Geographers, 66: 120-123.

Appleton, J. 1980 Landscape in the arts and the sciences. University of Hull, U.K.

Armstrong, M.P. 1991 Knowledge classification and organization. In: Buttenfield, B.P. and McMaster, R.B. (eds.) Map generalization. Longman.

Aronoff, S. 1991 Geographical information systems: A management perspective. WDL Publications, Canada.

Arthur, L.M. Daniel, T.C. and Boster, R.S. 1977 Scenic assessment: An overview. Landscape Planning, 4: 109-129.

Ashby, D.G. 1994 Computer modelling of information that is continuous over a spatial domain. Masters thesis, Dept. of Computer Science. Uni. of Canterbury.

Auckland Regional Authority, 1984: A technical manual on assessment of the Auckland region's landscape.

Band, L. E. 1986. Topographic partitioning of watersheds with digital elevation models. Water resource Res., 22: 15-24.

Barbanente, A. Borri, D. Esposito, F. Leo, P. Maciocco, G. and Selicato, F. 1992 Automatically acquiring knowledge by digital maps in artificial intelligence planning techniques. International Conference in GIS-From space to territory. Theories and methods of spatio-temporal reasoning. Proceedings. Pisa, Italy. Eds. A.U.Frank, I.Campari, and U.Formenti. 1992, New York, Springer-Verlag: 379-401.

Bennett, E. 1985 A practical approach to visual assessment. The Landscape, 9: 5-8.

Bernhardsen, T. 1992 Geographical Information Systems. Viak IT, Norway.

Berry, J.K. 1993 Cartographic Modelling: The analytical capabilities of GIS. In: Maguire, D.J. Goodchild, M.F. and Rwind, D.W. 1993 Geographic Information Systems: Principles and applications. Longman.

Bird, C. Peccol, E. Taylor, J. Brewer, T. and Keech, M. 1994 Monitoring landscape change - the role of GIS. Landscape research 19(3): 120-127.

Bishop, I. D. and Hulse, D.W. 1994 Prediction of scenic beauty using mapped data and geographic information systems. Landscape and urban planning 30: 59-70.

Boffa Miskell Partners Ltd. 1993 Landscape change in the Mackenzie/Waitaki basins. Christchurch.

Bourassa, S. 1991 The aesthetics of landscape. Belhaven Press, London.

Brabyn, L.K. 1991 The costs of a local planning process for rural tourism development. A case study of the new Bealey hotel. An unpublished research project for a postgraduate tourism course, Department of Geography, Canterbury University

Briggs, D.J. and France, J. 1981 Assessing landscape attractiveness: a South Yorkshire study. Landscape Research 6, No. 2: 2-5.

Brooke,D. 1994 A Countryside character programme. Landscape research 19(3): 128-132.

Brown, S.K. 1981 Visual assessment and environmental impact. Thesis for a Dip. L.A. Lincoln University.

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

Burrough, B.A. MacMillan, R.W. 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. and McMaster, R.B. 1991 Map generalization. Longman.

Canterbury Regional Council. 1995 Proposed regional policy statement, incorporating decisions on submissions received and minor amendments. Report 93 (23).

Canterbury Regional Council. 1993 Canterbury regional landscape study. Vol. 1 and 2.

Carlson, A. 1977: On the possibility of quantifying scenic beauty. Landscape Planning, 4: 131-172.

Cary, J. 1995 The construction of landscapes: A review of conceptual and methodological structures for exploring public perception of landscapes. Landcare Research New Zealand Ltd. Unpublished.

Cassettari, S. 1993 Introduction to integrated Geo-information management. Chapman and Hall.

Chrisman, N.R. 1991 The error component in spatial information. In: Maguire, D.J. Goodchild, M.F. and Rwind, D.W. 1993 Geographic Information Systems, Principles and applications. Longman.

Clamp, P. 1976 An evaluation of the impact of roads on the visual amenity of rural areas. Ralp Hopkinson, Newton Watson and Partners/DOE Research Report 7.

Clark, D.M. Hastings, D.A. and Kineman, J.J. 1991 Globial databases and their implications for GIS. In: Maguire, D.J. Goodchild, M.F. and Rwind, D.W. 1993 Geographic Information Systems: Principles and applications. Longman.

Clyd County Council 1978 Structure plan. Report of survey environment. Clyd County Council, England.

Collier, A. 1991 Principles of tourism - a New Zealand perspective, 2nd edition, Pitman, Auckland.

Collins, S. H. 1975. Terrain parameters directly from a digital terrain model. The Canadian Surveyor, Vol. 29, No. 5, December 1975: 507-518

Countryside Commission for Scotland, 1970 A planning classification of Scottish landscape resources. Countryside Commission. Perth, Scotland.

Countryside Commission for Scotland, 1988 A review of recent practice and research in landscape assessment. Countryside Commission. Perth, Scotland.

Cowen, J. 1993 A proposed method for calculating the LS factor for use with the USLE in a grid-based environment. Geography organizing our world. ARC/INFO. Proceedings of the thirteenth annual ESRI user conference. Vol.1, California.

Crozier, M.J. and Owen, R.C 1983 Terrain evaluation for rapid ecological survey. Physical Geography, University of Victoria.

Curran, P. J. 1985 Principles of remote sensing. Longman.

Daniel, T.C. and Vining, J. 1983 Methodological issues in the assessment of landscape quality. In: Altman, I. and Wohlwill, J.F. 1983 Behaviour and natural environments. New York, Plenum Press.

Davis, W.M. 1902 Elementary physical geography. Ginn, New York.

Dearden, P. 1980 A statistical technique for the evaluation of the visual quality of the landscape for landuse planning purposes. Journal of Environmental Management. 10: 51-68.

Dearden, P. 1989 Societal landscape preferences: A pyramid of influences. In: Dearden, P. and Sadler, B. 1989 Landscape evaluation: Approaches and applications. Western Geographical Series. Vol.25, Dept. of Geography, University of Victoria, Canada.

Densem, G. 1980 Landscape evaluation: The Marlborough Sounds landscape study. The Landscape, 9: 8-11.

Department of Statistics and Ministry for the Environment, 1990 State of the Environment reporting in New Zealand. A proposal for a multi-agency project.

Department of Statistics 1992 Documenting the Environment. Department of Statistics, Wellington.

Department of Survey and Land Information 1984 NZMS 270 series 1:250,000 topoplots. Specifications for photogrammetric symbols, procedures and processes. 4th edition, Department of Survey and Land Information, Wellington.

Department of Survey and Land Information 1994 Vegetation information from space's "flying fax". Info News, Issue 22.

Dikau, R. 1989 The application of a digital relief model to landform analysis. In: Raper, J.F. (ed.) 1989 Three dimensional applications in Geographical Information Systems. Taylor and Francis, London: 51-77.

Dikau, R. Brabb, E.E. and Mark, R.M. 1991 Landform classification of New Mexico by computer. U.S. Department of the Interior, U.S. Geological Survey. Open-file report 91-634.

Duffield, B.S. and Coppock, J.T. 1975 The delineation of recreational landscapes: the role of a computer-based information system. Transactions Instit. British Geographers, 66(1): 141-148.

Durham County Council 1974 County structure plan. Landscape survey. England.

Dymond, J.R. and Harmsworth, G.R. 1994 Towards automated land resource mapping using digital terrain models. ITC Journal, 2: 129-138.

Dymond, J.R. DeRose, R.C. and Harmsworth, G.R. 1995 Automated mapping of the land components from digital elevation data. Earth Surface Processes and Landform. Vol. 20: 131-137.

Electricity Corporation of New Zealand Limited 1988 Computer Aided Transmission Line Corridor Study for a Second High Voltage Direct Current Inter-Island Link. Prepared by Boffa Miskell Partners Ltd and GECO NZ.

Environmental Systems Research Institute 1991 ARC/INFO - users guide. Cell - based modelling with GRID. Analysis, display, and management. ESRI, California.

Environmental Systems Research Institute 1993 The Digital Chart of the World. ESRI, California.

Fabos, J. 1979 Planning and landscape evaluation. Landscape Research, 4(2): 4-10.

Fairweather, J.R. and Swaffield, S. 1994 Preferences for land use options in the Mackenzie/Waitaki Basin. Research report No.224, Agribusiness and economics research unit, Lincoln University.

Geiger, R. 1971 The climate near the ground. Harvard University Press, Cambridge.

Glasson, C.R. 1991 A visual assessment of Banks Peninsula. Banks Peninsula Council.

Goodchild, M.F. 1993 Data models and data quality: Problems and prospects. In: Goodchild, M.F. Parks, B.O. and Steyaert, L.T. 1993 Environmental modelling with GIS, Oxford University Press.

Haines-Young, R.H. and Petch, P.R. 1986 Physical geography: Its nature and methods. Paul Chapman Publishing, London.

Hamill, L. 1989 On the persistence of error in landscape aesthetics. In: Dearden, P. and Sadler, B. 1989 Landscape evaluation: Approaches and applications. Western Geographical Series. Vol.25, Dept. of Geography, University of Victoria, Canada.

Hammond, E.H. 1954 Small scale continental landform maps. Annals of Association of American Geographers, 44: 32-42.

Hammond, E.H. 1964 Analysis of properties in landform geography: An application to broadscale landform mapping. Annals of Association of American Geographers, 54: 11-19.

Hay, R. 1990 Sense of place: Cross-cultural perspectives from Banks Peninsula, New Zealand. Ph.D. thesis. Department of Geography, University of Canterbury.

Higuchi, T. 1988 The visual and spatial structure of landscapes. Translated by Terry, T. Massachusetts Institute of Technology Press.

Itami, R. 1989 Scenic perception: Research and application in U.S. Visual Management Systems. In: Dearden, P. and Sadler, B.1989 Landscape evaluation: Approaches and applications. Western Geographical Series. Vol.25, Dept. of Geography, University of Victoria, Canada.

Jackman, T. 1988: Our national landscape. Biological Resources Centre. Department of Scientific and Industrial Research.

Jackson, J.B. 1984 Discovering the vernacular landscape. Yale University Press, New Haven.

Jackson, R.M. 1990 Digital landscapes. In: Kearsley, G. and Fitzharris, B. (editors) 1990 Southern landscapes. Dept. of Geography, University of Otago.

Jeffers, J.N.R. 1982 Modelling, Outline Studies in Ecology. Chapman and Hall, New York.

Jeffers, J.N.R. 1973 Systems modelling and analysis in resource management. J. Env. Mgmt. 1(1): 13-28.

Jones, M. 1991 The elusive reality of landscape. Concepts and approaches in landscape research. Norsk Geografisk Tidsskrift, 45: 229-244.

Kearsley, G.W. and Gray, G. 1993 International visitor flows and infrastructure needs: A New Zealand example. New Zealand Geography Conference Proceedings 1993.

Kilvert, S. and Hartsough, B. 1993 Visual impact of forest operations: measuring concern in New Zealand. Logging Industry Research Organisation. Report Vol. 18, No. 9.

Klemsdal, T. and Sjulsen, O.E. 1988 The Norwegian macro-landforms: definitions, distribution and system of evolution. Norsk geogr. Tidsskr. Vol. 42: 133-147.

Kliskey, A.D. and Kearsley, G.W. 1993 Mapping multiple perceptions of wilderness in southern New Zealand. Applied Geography 13: 203-223.

Land Information New Zealand 1985 LINZ standards for geographical reference. LINZ publication No. 1.

Land Information New Zealand 1987a LINZ standards for area measurement. LINZ publication No. 8.

Land Information New Zealand 1987b LINZ standards for land appellation. LINZ publication No. 6 and 7.

Langridge, D.W. 1992 Classification: Its kinds, systems, elements and application. Bowker Saur, London.

Lay, J. 1991: Terrain feature extraction from digital elevation models: A multi-perspective exploration. GIS/LIS 1991: 191-198.

Leckie, D.G. 1990 Advances in remote sensing technologies for forest surveys and management. Canadian Journal of Forest Research, 20: 463-483.

Lee, Y.C. 1991 Cartographic data capture and storage. In: Taylor, F. (edit.) Geographic information systems. Pergamon.

Leopold, L.B. 1969 Landscape aesthetics. Natural History, 73: 36-45.

Lesslie, R.G. Mackey,B.G. and Shulmeister, J. 1988 Wilderness quality in Tasmania. A report to the Australian Heritage Commission.

Linton, D. 1970: The assessment of scenery as a natural resource. Scottish Geographical Magazine, 84 (3): 219-238.

Lowe, P.D. 1977 Amenity and equity: a review of local environmental pressure groups in Britain. Environmental planning 9: 35-58.

Lowenthal, D. 1978 Finding valued landscapes. Progress in Human Geography 2, No. 3: 373-418.

Maguire, D.J. 1989 Computers in Geography. Longman.

Maguire, D.J. 1991 An overview and definition of GIS. In: Maguire, D.J. Goodchild, M.F. and Rwind, D.W. 1993 Geographic Information Systems: Principles and applications. Longman.

Mather, P.M. 1987 Computer processing of remotely sensed images. Wiley.

Midtbø, T. 1992. Generalization of extensive terrain surfaces represented by triangular irregular networks. Neste generasjon GIS konferanseforedrag, Trondheim, 14-15 December 1992. Institutt for kart og oppmaling, Universitet i Trondheim, Norway.

Miller, D.R. Morrice, J.G. Horne, P.L. and Aspinall, R.J. 1994 The use of geographic information systems for analysis of scenery in the Cairngorm mountains, Scotland. In: Mountain environments and GIS. Edited by Price, M.F. and Heywood,D.I. Taylor and Francis, London.

Ministry of Works and Development, 1983: Natural resources of the Canterbury region: A survey and evaluation for management. Ministry of Works and Development.

Ministry of Works and Development, 1987 VAMPLAN: Visual assessment method for planning; an introductory guide. Ministry of Works and Development.

Mitchell, B. 1993 Geography and resource analysis. Longman Scientific and Technical.

Mitchell, C. 1973 Terrain evaluation. Longman.

Moore, I.D. Turner, A.K. Wilson, J.P. Jenson, S.K. and Band, L.E. 1993 GIS and land-surface-subsurface process modelling. In: Goodchild, M.F. Parks, B.O. and Steyaert, L.T. 1993 Environmental modelling with GIS, Oxford University Press.

Moore, J. and Bennison, T. 1993 Image synthesis. The Landscape, Autumn 1993: 16-18.

Mosley, M.P. 1989 Perception of New Zealand river scenery. New Zealand Geographer. Vol.45, No.1: 2-13.

Murcott, R. 1995 Research meets practice - standards. AURISA/SIRC'95 The 7th colloquium of the spatial information research centre, University of Otago, in association with AURISA New Zealand and Massey University. April 26th-28.

Myers, S.C. Park, G.N. and Overmans, F.B. 1987 New Zealand protected natural areas programme. A guideline for the rapid ecological survey of natural areas. New Zealand Biological Resources Centre.

Naveh, Z. and Lieberman, A. 1994 Landscape ecology. Theory and application. 2nd edition. Springer-Verlag.

Newsome, P. 1987 The vegetation cover of New Zealand. Ministry of Works and Development.

Newsome, P. 1995 Directory of geographic databases within Manaaki Whenua-Landcare. Landcare New Zealand Ltd.

Nyerges, T.L. 1991a Representing geographical knowledge. In: Buttenfield, B.P. and McMaster, R.B. (eds.) Map generalization. Longman.

Nyerges, T.L. 1991b Geographic information abstractions: conceptual clarity for geographic modelling. In: Environment and Planning A 1991 23: 1483-1499.

O'Brien L. 1992 Introducing quantitative geography. Measurements, methods and generalised linear models. Routledge, London.

OECD, 1991 The state of the environment. OECD, Paris

Penning-Rowsell, E.C. and Searle, G.H. 1977 The Manchester Landscape evaluation method: a critical appraisal. Landscape Research, 2 no.3: 6-11.

Petts, G. and Foster, I. 1985 Rivers and landscape. Edward Arnold, London.

Pomeroy, J. Fitzgibbon, J. and Green M. 1989 The use of personal construct theory in evaluating aesthetics. In: Dearden, P. and Sadler, B. 1989 Landscape evaluation: Approaches and applications. Western Geographical Series. Vol.25, Dept. of Geography, University of Victoria, Canada.

Porteous, J. 1990 Landscapes of the mind, worlds of sense and metaphor. University of Toronto Press, Toronto.

Rhind, D.W. 1988 "A GIS research agenda" International Journal of Geographical Systems 2: 23-28.

Rhind, D. and Hudson, R. 1980 Landuse. Methuen, London.

Ribe, R.G. 1982 On the possibility of quantifying scenic beauty - a response. Landscape planning, 9: 61-75.

Robinson, D.G. Laurie, I.C. Wager, J.F. and Trail, A.L. 1976 Landscape evaluation. University of Manchester.

Shea, K.S. 1991 Design considerations for an artificially intelligent system. In: Buttenfield, B.P. and McMaster, R.B. (eds.) Map generalization. Longman.

Shepard F.P. 1937 Revised classification of marine shorelines: a reply. J. Geology 45: 602-624.

Shepard F.P. 1938 A classification of marine shorelines: a reply. J. Geology 46: 996-1006.

Soons, J.M. and Selby, M.J. (editors) 1982 Landforms of New Zealand. Longman.

Southard, R.B. 1987 Automation in Cartography - revolution or evolution. The Cartographic Journal, 24: 59-63.

Stankey, G.H. and Schreyer, R. 1987 Attitudes towards wilderness and factors effecting visitor behaviour: a state of knowledge review. In Proceedings-National Wilderness Conference: Issues, State-of-Knowledge, Future Directions (R.C.Lucas, comp.): 246-293. Ogden, UT:USDA Forest Service, General technical report INT-220.

Statistics New Zealand 1994 New Zealand official year book 94. Statistics New Zealand, Wellington.

Steinitz, C. 1993 A framework for theory and practice in landscape planning. GIS Europe.

Swaffied, S.R. 1991 Roles and meanings of "Landscape". Ph.D. thesis, Lincoln University.

Swaffield, S.R. 1994 Attitudes towards tress: a case study in the New Zealand eastern high country. N.Z. Forestry (February): 25-30.

Tang, L. 1992 Automatic extraction of specific geomorphological elements from contours. Proceedings of the 5th international symposium on spatial data handling. IGU Commission on GIS, 1992, Vol. 2.

Tansley, A.G. 1946 Introduction to plant ecology. George Allen and Unwin, London.

Tomlin, C.D. 1990 Geographic information systems and cartographic modelling. Prentice-Hall, New Jersey.

Tomlin, C.D. 1993 Cartographic Modelling. In: Maguire, D.J. Goodchild, M.F. and Rwind, D.W. 1993 Geographic Information Systems: Principles and applications. Longman.

United Nations Environmental Monitoring Programme. 1990 GRID Global resource information database. Nairobi, Kenya.

Valentine, H. 1952 Die Kusten der Erde. Justus Perthes, Gotha.

Wallace, H.W. 1955 New Zealand landforms. New Zealand Geographer 11(1): 17-27.

Ward, J.C. 1991 Integrated environmental monitoring. Information paper No. 37, Centre for resource management, Lincoln University.

Ward, J.C. and Beanland, B. 1992 Contributions to a national set of environmental indicators to be monitored at a regional level. Information paper No. 36, Centre for resource management, Lincoln University.

Weerakkody, U. 1993 Coastal classification for practical applications. ITC Journal 4: 386-390.

Weibel, R. and DeLotto, J.S. 1988 Automated terrain classification for GIS modelling. GIS/LIS'88: 618-627.

Weibel, R. & Heller, M. 1993. Digital terrain modelling. In: Maguire, D.J. Goodchild, M.F. and Rwind, D.W. 1993 Geographic Information Systems: Principles and applications. Longman.

Whitehouse, I.E. Basher, L.R. and Tonkin, P.J. 1990 A landform classification for PNA surveys in Southern Alps. Department of Conservation, Science and research series, No. 44.

Wilson, A.G. 1970 Entropy in urban and regional modelling. Pion Ltd. London.

Wrigley, N. 1995 Revisiting the modifiable areal unit problem and ecological fallacy. In: Cliff, A.D Gould, P.R. Hoare, A.G. and Thrift, N.J. 1995 Diffusing geography, Blackwell publishers, Oxford, UK.

Zadeh, L.A. 1965 Fuzzy sets and systems. In: Fox, J. Systems theory. Microwave research institute symposia series IV. Polytechnic Press, New York: 29-37.

Zimmermann, E.W. 1951 World resources and industries. Harper and Brothers, New York.

Zimmerman, H.J. 1992 Fuzzy set theory and its application. Kluwer Academic Publishers, London.

Zube, E. Sell, J. and Taylor, J. 1982 Landscape perception: Research, application and theory. Landscape Planning 9: 1-33.

Zube, E.H. 1984a Environmental evaluation: Perception and public policy. Cambridge University Press.

Zube, E.H. 1984b Themes in landscape assessment theory. Landscape Journal, 3: 104-110.

Appendix 1

The definitions of the different vegetation classes

Abbreviations

Hort. = Horticulture
Past. = Pasture
Tuss. = Tussock
I.Sc. = Indigenous Scrub
E.Sc. = Exotic Scrub
A.Sc. = Alpine Scrub
I.Fo. = Indigenous Forest
E.Fo. = Exotic Forest
Alpi. = Alpine herbfields, Rock, and Ice
Wetl. = Wetland
Sand. = Sanddune

(1) Horticulture Pasture

Hort. > 20% and Past. <= 100% and Tuss. <= 20% and I.Sc. <= 20% and E.Sc. <= 20% and A.Sc. <= 20% and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(2) Horticulture Tussock

Hort. > 20% and Past. <= 100% and ((Tuss. > 20% and I.Sc. <= 20% and E.Sc. <= 20% and A.Sc. <= 20%) or (Tuss. > 50% and I.Sc. <= 50% and E.Sc. <= 50% and A.Sc. <= 50%)) and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(3) Horticulture Scrub

Hort. > 20% and Past. <= 100% and ((Tuss. <= 50% and (I.Sc. > 20% or E.Sc. > 20% or A.Sc. > 20%)) or (Tuss. <= 100% and (I.Sc. > 50% or E.Sc. > 50% or A.Sc. > 50%))) and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(4) Horticulture Forest

Hort. > 20% and Past. <= 100% and Tuss. <= 100% and I.Sc. <= 100% and E.Sc. <= 100% and A.Sc. <= 100% and (I.Fo. > 20% or E.Fo. > 20%) and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(5) Horticulture Wetland

Hort. > 20% and Past. <= 100% and Tuss. <= 20% and I.Sc. <= 50% and E.Sc. <= 20% and A.Sc. <= 20% and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. > 20% and Sand. <= 20%

(6) Horticulture Sanddune

Hort. > 20% and Past. <= 50% and Tuss. <= 20% and I.Sc. <= 20% and E.Sc. <= 20% and A.Sc. <= 20% and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. > 20%

(7) Pasture

Hort. <= 20% and Past. > 20% and Tuss. <= 20% and I.Sc. <= 20% and E.Sc. <= 20% and A.Sc. <= 20% and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(8) Tussock

Hort. <= 20% and Past. <= 20% and Tuss. > 20% and I.Sc. <= 20% and E.Sc. <= 20% and A.Sc. <= 20% and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(9) Tussock Pasture

Hort. <= 20% and Past. > 20% and Tuss. > 20% and I.Sc. <= 20% and E.Sc. <= 20% and A.Sc. <= 20% and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(10) Pasture Indigenous Scrub

Hort. <= 20% and Past. > 20% and Tuss. <= 20% and (I.Sc. > 20% or A.Sc. > 20%) and E.Sc. <= 20% and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(11) Pasture Exotic Scrub

Hort. <= 20% and Past. > 20% and Tuss. <= 20% and I.Sc. <= 100% and E.Sc. > 20% and A.Sc. <= 20% and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(12) Tussock Grassland Scrub

Hort. <= 20% and Past. <= 100% and Tuss. > 20% and (I.Sc. > 20% or E.Sc. > 20% or A.Sc. > 20%) and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(13) Lowland Indigenous Scrub

Hort. <= 20% and Past. <= 20% and Tuss. <= 20% and ((I.Sc. > 20% and E.Sc. <= 20%) or (I.Sc. > 50% and E.Sc. <= 50%)) and A.Sc. <= 20% and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(14) Exotic Scrub

Hort. <= 20% and Past. <= 20% and Tuss. <= 20% and I.Sc. <= 50% and E.Sc. > 20% and A.Sc. <= 20% and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(15) Alpine Scrub

Hort. <= 20% and Past. <= 20% and Tuss. <= 20% and I.Sc. <= 100% and E.Sc. <= 20% and A.Sc. > 20% and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(16) Pasture Indigenous Forest

Hort. <= 20% and Past. > 20% and Tuss. <= 20% and I.Sc. <= 20% and E.Sc. <= 20% and A.Sc. <= 20% and I.Fo. > 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(17) Pasture Exotic Forest

Hort. <= 20% and Past. > 20% and Tuss. <= 20% and I.Sc. <= 20% and E.Sc. <= 20% and A.Sc. <= 20% and I.Fo. <= 20% and E.Fo. > 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(18) Tussock Indigenous Forest

Hort. <= 20% and Past. <= 100% and (Tuss. > 20% or Alpi. > 20%) and I.Sc. <= 100% and E.Sc. <= 100% and A.Sc. <= 100% and I.Fo. > 20% and E.Fo. <= 20% and Wetl. <= 20% and Sand. <= 20%

(19) Tussock Exotic Forest

Hort. <= 20% and Past. <= 100% and Tuss. > 20% and I.Sc. <= 100% and E.Sc. <= 100% and A.Sc. <= 100% and I.Fo. <= 20% and E.Fo. > 20% and Alpi. <= 100% and Wetl. <= 20% and Sand. <= 20%

(20) Indigenous Forest Indigenous Scrub

Hort. <= 20% and Past. <= 20% and Tuss. <= 20% and I.Sc. > 20% and E.Sc. <= 20% and A.Sc. <= 20% and I.Fo. > 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(21) Indigenous Forest Exotic Scrub

Hort. <= 20% and Past. <= 20% and Tuss. <= 20% and I.Sc. <= 100% and E.Sc. > 20% and A.Sc. <= 100% and I.Fo. > 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(22) Indigenous Forest Alpine Scrub

Hort. <= 20% and Past. <= 20% and Tuss. <= 20% and I.Sc. <= 100% and E.Sc. <= 20% and A.Sc. > 20% and I.Fo. > 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(23) Exotic Forest Scrub

Hort. <= 20% and Past. <= 20% and Tuss. <= 20% and (I.Sc. > 20% or E.Sc. > 20% or A.Sc. > 20%) and I.Fo. <= 20% and E.Fo. > 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

()
24 Indigenous Forest Scrub Pasture

Hort. <= 20% and Past. > 20% and Tuss. <= 20% and (I.Sc. > 20% or E.Sc. > 20% or A.Sc. > 20%) and I.Fo. > 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(25) Exotic Forest Scrub Pasture

Hort. <= 20% and Past. > 20% and Tuss. <= 20% and (I.Sc. > 20% or E.Sc. > 20% or A.Sc. > 20%) and I.Fo. <= 20% and E.Fo. > 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(26) Indigenous Forest

Hort. <= 20% and Past. <= 20% and Tuss. <= 20% and I.Sc. <= 20% and E.Sc. <= 20% and A.Sc. <= 20% and I.Fo. > 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(27) Exotic Forest

Hort. <= 20% and Past. <= 20% and Tuss. <= 20% and I.Sc. <= 20% and E.Sc. <= 20% and A.Sc. <= 20% and I.Fo. <= 20% and E.Fo. > 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(28) Tussock Scrub Alpine

Hort. <= 20% and Past. <= 100% and (Tuss. > 20% or A.Sc. > 20% or I.Sc. > 20%) and E.Sc. <= 20% and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. > 20% and Wetl. <= 20% and Sand. <= 20%

(29) Alpine Herbfields, Ice, and Rock

Hort. <= 20% and Past. <= 20% and Tuss. <= 20% and I.Sc. <= 20% and E.Sc. <= 20% and A.Sc. <= 20% and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. > 20% and Wetl. <= 20% and Sand. <= 20%

(30) Mixed Forest

Hort. <= 20% and Past. <= 20% and Tuss. <= 20% and I.Sc. <= 20% and E.Sc. <= 20% and A.Sc. <= 20% and I.Fo. > 20% and E.Fo. > 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(31) Mixed Forest Grassland

Hort. <= 20% and (Past. > 20% or Tuss. > 20%) and I.Sc. <= 20% and E.Sc. <= 20% and A.Sc. <= 20% and I.Fo. > 20% and E.Fo. > 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(32) Mixed Forest Scrub

Hort. <= 20% and Past. <= 20% and Tuss. <= 20% and (I.Sc. > 20% or E.Sc. > 20% or A.Sc. > 20%) and I.Fo. > 20% and E.Fo. > 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

(33) Mixed Forest Grassland Scrub

Hort. <= 20% and (Past. > 20% or Tuss. > 20%) and (I.Sc. > 20% or E.Sc. > 20% or A.Sc. > 20%) and I.Fo. > 20% and E.Fo. > 20% and Alpi. <= 100% and Wetl. <= 20% and Sand. <= 20%

(34) Wetland

Hort. <= 20% and Past. <= 20% and Tuss. <= 20% and I.Sc. <= 20% and E.Sc. <= 20% and A.Sc. <= 20% and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. > 20% and Sand. <= 20%

(35) Wetland Grassland

Hort. <= 20% and (Past. > 20% or Tuss. > 20%) and I.Sc. <= 20% and E.Sc. <= 20% and A.Sc. <= 20% and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. > 20% and Sand. <= 20%

(36) Wetland Indigenous Scrub

Hort. <= 20% and Past. <= 100% and Tuss. <= 100% and (I.Sc. > 20% or A.Sc. > 20%) and E.Sc. <= 20% and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. > 20% and Sand. <= 20%

(37) Wetland Exotic Scrub

Hort. <= 20% and Past. <= 100% and Tuss. <= 100% and I.Sc. <= 100% and E.Sc. > 20% and A.Sc. <= 100% and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. > 20% and Sand. <= 20%

(38) Wetland Exotic Forest

Hort. <= 20% and Past. <= 100% and Tuss. <= 100% and I.Sc. <= 100% and E.Sc. <= 100% and A.Sc. <= 100% and I.Fo. <= 100% and E.Fo. > 20% and Alpi. <= 100% and Wetl. > 20% and Sand. <= 20%

(39) Wetland Indigenous Forest

Hort. <= 20% and Past. <= 100% and Tuss. <= 100% and I.Sc. <= 100% and E.Sc. <= 100% and A.Sc. <= 100% and I.Fo. > 20% and E.Fo. <= 20% and Alpi. <= 100% and Wetl. > 20% and Sand. <= 20%

(40) Sanddune

Hort. <= 50% and Past. <= 20% and Tuss. <= 20% and I.Sc. <= 20% and E.Sc. <= 20% and A.Sc. <= 20% and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. <= 50% and Sand. > 20%

(41) Sanddune Grassland

Hort. <= 20% and (Past. > 20% or Tuss. > 20%) and I.Sc. <= 20% and E.Sc. <= 20% and A.Sc. <= 20% and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. > 20%

(42) Sanddune Indigenous Scrub

Hort. <= 20% and Past. <= 100% and Tuss. <= 100% and (I.Sc. > 20% or A.Sc. > 20%) and E.Sc. <= 20% and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. > 20%

(43) Sanddune Exotic Scrub

Hort. <= 20% and Past. <= 100% and Tuss. <= 100% and I.Sc. <= 100% and E.Sc. > 20% and A.Sc. <= 100% and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. <= 100% and Wetl. <= 20% and Sand. > 20%

(44) Sanddune Exotic Forest

Hort. <= 20% and Past. <= 100% and Tuss. <= 100% and I.Sc. <= 100% and E.Sc. <= 100% and A.Sc. <= 100% and I.Fo. <= 100% and E.Fo. > 20% and Alpi. <= 100% and Wetl. <= 20% and Sand. > 20%

(45) Sanddune Indigenous Forest

Hort. <= 20% and Past. <= 100% and Tuss. <= 100% and I.Sc. <= 100% and E.Sc. <= 100% and A.Sc. <= 100% and I.Fo. > 20% and E.Fo. <= 20% and Alpi. <= 100% and Wetl. <= 20% and Sand. > 20%

(46) Urban Area, Water, or not Classified

Hort. <= 20% and Past. <= 20% and Tuss. <= 20% and I.Sc. <= 20% and E.Sc. <= 20% and A.Sc. <= 20% and I.Fo. <= 20% and E.Fo. <= 20% and Alpi. <= 20% and Wetl. <= 20% and Sand. <= 20%

Appendix 2

The definitions of the different landuse classes

Abbreviations

urban = Urban roads.
popl = Towns.
popm = Large settlements.
pops = Small settlement.
roadnh = National highways.
roadph = Provincial highways.
roadsecs = Secondary sealed roads.
roadsecm = Secondary metal roads.
roadcom = All roads (except 4WD. tracks) combined .
track = Four wheel drive tracks.
pylon = Pylons.
mast = TV. and radio masts.
rail = Railways.
build = Buildings (except huts).
huts = Mountain huts and holiday baches.

(1) Urban

urban > 10%

(2) Developed Rural

urban <= 10% and roadcom > 20%

(3) Rural

urban <= 10% and roadcom > 10% and <= 20%

(4) Town

urban <= 10% and roadcom <= 10% and popl > 0%

(5) Large Settlement

urban <= 10% and roadcom <= 10% and popl = 0% and popm > 0%

(6) Small Settlement

urban <= 10% and roadcom <= 10% and popl = 0% and popm = 0% and pops > 0%

(7) Highway with Utilities

urban <= 10% and roadcom <= 10% and popl = 0% and popm = 0% and pops = 0% and (roadnh > 0% or roadph > 0%) and (pylon > 0% or mast > 0% or rail > 0%)

(8) Highway with Isolated Buildings

urban <= 10% and roadcom <= 10% and popl = 0% and popm = 0% and pops = 0% and (roadnh > 0% or roadph > 0%) and pylon = 0% and mast = 0% and rail = 0% and (build > 0% or huts > 0%)

(9) Highway

urban <= 10% and roadcom <= 10% and popl = 0% and popm = 0% and pops = 0% and (roadnh > 0% or roadph > 0%) and pylon = 0% and mast = 0% and rail = 0% and build = 0% and huts = 0%

(10) Utility with Sealed Secondary Roads

urban <= 10% and roadcom <= 10% and popl = 0% and popm = 0% and pops = 0% and roadnh = 0% and roadph = 0% and roadsecs > 0% and (pylon > 0% or mast > 0% or rail > 0%)

(11) Utility with Unsealed Secondary Road

urban <= 10% and roadcom <= 10% and popl = 0% and popm = 0% and pops = 0% and roadnh = 0% and roadph = 0% and roadsecs = 0% and roadsecm > 0% and (pylon > 0% or mast > 0% or rail > 0%)

(12) Utility with 4 wheel drive track

urban <= 10% and roadcom <= 10% and popl = 0% and popm = 0% and pops = 0% and roadnh = 0% and roadph = 0% and roadsecs = 0% and roadsecm = 0% and track > 0% and (pylon > 0% or mast > 0% or rail > 0%)

(13) Utility

urban <= 10% and roadcom <= 10% and popl = 0% and popm = 0% and pops = 0% and roadnh = 0% and ROADpH = 0% and roadsecs = 0% and roadsecm = 0% and track = 0% and (pylon > 0% or mast > 0% or rail > 0%)

(14) Sealed Secondary Road with Buildings

urban <= 10% and roadcom <= 10% and popl = 0% and popm = 0% and pops = 0% and roadnh = 0% and roadph = 0% and roadsecs > 0% and roadsecm >= 0% and track >= 0% and pylon = 0% and mast = 0% and rail = 0% and (build > 0% or huts >0%)

(15) Unsealed Secondary Road with Buildings

urban <= 10% and roadcom <= 10% and popl = 0% and popm = 0% and pops = 0% and roadnh = 0% and roadph = 0% and roadsecs = 0% and roadsecm > 0% and track > 0% and pylon = 0% and mast = 0% and rail = 0% and (build > 0% or huts > 0%)

(16) Four Wheel Drive Track with Buildings

urban <= 10% and roadcom <= 10% and popl = 0% and popm = 0% and pops = 0% and roadnh = 0% and roadph = 0% and roadsecs = 0% and roadsecm = 0% and track > 0% and pylon = 0% and mast = 0% and rail = 0% and (build > 0% or huts > 0%)

(17) Sealed Secondary Road

urban <= 10% and roadcom <= 10% and popl = 0% and popm = 0% and pops = 0% and roadnh = 0% and roadph = 0% and roadsecs > 0% and roadsecm > 0% and track > 0% and pylon = 0% and mast = 0% and rail = 0% and build = 0% and huts = 0%

(18) Unsealed Secondary Road

urban <= 10% and roadcom <= 10% and popl = 0% and popm = 0% and pops = 0% and roadnh = 0% and roadph = 0% and roadsecs = 0% and roadsecm > 0% and track > 0% and pylon = 0% and mast = 0% and rail = 0% and build = 0% and huts = 0%

(19) Four Wheel Drive Track

urban <= 10% and roadcom <= 10% and popl = 0% and popm = 0% and pops = 0% and roadnh = 0% and roadph = 0% and roadsecs = 0% and roadsecm = 0% and track > 0% and pylon = 0% and mast = 0% and rail = 0% and build = 0% and huts = 0%

(20) Backcountry Many Buildings

urban = 0% and roadcom <= 10% and roadsecs = 0% and roadsecm = 0% and popl = 0% and popm = 0% and pops = 0% and roadnh = 0% and roadph = 0% and track = 0% and pylon = 0% and mast = 0% and rail = 0% and (build > 8% or huts > 8%)

(21) Backcountry Few Buildings

urban = 0% and roadcom <= 10% and roadsecs = 0% and roadsecm = 0% and popl = 0% and popm = 0% and pops = 0% and roadnh = 0% and roadph = 0% and track = 0% and pylon = 0% and mast = 0% and rail = 0% and (build > 0% and <= 7% or huts > 0% and <= 2%)

(22) Remote

urban = 0% and roadcom <= 10% and roadsecs = 0% and roadsecm = 0% and popl = 0% and popm = 0% and pops = 0% and roadnh = 0% and roadph = 0% and track = 0% and pylon = 0% and mast = 0% and rail = 0% and build = 0% and huts = 0%

Appendix 3

The definitions of the different influence of water classes

(1) Lake

river is absent and lake is present and coast is absent

(2) Lake/River

river is present and lake is present and coast is absent

(3) Lake/River/Non-indented Coast

river is present and lake is present and coast is non-indented

(4) Lake/River/Indented Coast

river is present and lake is present and coast is indented

(5) Lake/River/Very Indented Coast

river is present and lake is present and coast is very indented

(6) Lake/Non-indented Coast

river is absent and lake is present and coast is non-indented

(7) Lake/Indented Coast

river is absent and lake is present and coast is indented

(8) Lake/Very Indented Coast

river is absent and lake is present and coast is very indented

(9) River

river is present and lake is absent and coast is absent

(10) River/Non-indented Coast

river is present and lake is absent and coast is non indented

(11) River/Indented Coast

river is present and lake is absent and coast is indented

(12) River/Very Indented Coast

river is present and lake is absent and coast is very indented

(13) Non-indented Coast

river is absent and lake is absent and coast is non-indented

(14) Indented Coast

river is absent and lake is absent and coast is indented

(15) Very Indented Coast

river is absent and lake is absent and coast is very indented

(16) Absent

river is absent and lake is absent and coast is absent

Appendix 4

The definitions of the different landform classes

Abbreviations

pla = plains
val = open valleys
lhill = low hills
hill = hills
hhill = high hills
mount = mountains
hmount = high mountains
tab = tablelands

(1) Plains

pla > 20% and val <= 100% and lhill <= 20% and hill <= 20% and hhill <= 20% and mount <= 20% and hmount <= 20% and tab <= 20% or (pla <= 100% and val > 0% and lhill = 0% and hill = 0% and hhill = 0% and mount = 0% and hmount = 0% and tab = 0%)

(2) Plains Low Hills

pla > 20% and val <= 100% and lhill > 20% and hill <= 20% and hhill <= 20% and mount <= 20% and hmount <= 20% and tab <= 20%

(3) Plains Hills

pla > 20% and val <= 100% and lhill <= 100% and hill > 20% and hhill <= 20% and mount <= 20% and hmount <= 20% and tab <= 20%

(4) Plains High Hills

pla > 20% and val <= 100% and lhill <= 100% and hill <= 100% and hhill > 20% and mount <= 20% and hmount <= 20% and tab <= 20%

(5) Plains Mountains

pla > 20% and val <= 100% and lhill <= 100% and hill <= 100% and hhill <= 100% and mount > 20% and hmount <= 20% and tab <= 20%

(6) Plains High Mountains

pla > 20% and val <= 100% and lhill <= 100% and hill <= 100% and hhill <= 100% and mount <= 100% and hmount > 20% and tab <= 20%

(7) Open Valley Low Hills

pla <= 20% and val > 20% and lhill > 0% and hill = 0% and hhill = 0% and mount = 0% and hmount = 0% and tab <= 20%

(8) Open Valley Hills

pla <= 20% and val > 20% and lhill <= 100% and hill > 0% and hhill = 0% and mount = 0% and hmount = 0% and tab <= 20%

(9) Open Valley High Hills

pla <= 20% and val > 20% and lhill <= 100% and hill <= 100% and hhill > 0% and mount = 0% and hmount = 0% and tab <= 20%

(10) Open Valley Mountains

pla <= 20% and val > 20% and lhill <= 100% and hill <= 100% and hhill <= 100% and mount > 0% and hmount = 0% and tab <= 20%

(11) Open Valley High Mountains

pla <= 20% and val > 20% and lhill <= 100% and hill <= 100% and hhill <= 100% and mount <= 100% and hmount > 0% and tab <= 20%

(12) Low Hills

pla <= 20% and val <= 20% and lhill > 20% and hill <= 20% and hhill <= 20% and mount <= 20% and hmount <= 20% and tab <= 20%

(13) Hills

pla <= 20% and val <= 20% and lhill <= 100% and hill > 20% and hhill <= 20% and mount <= 20% and hmount <= 20% and tab <= 20%

(14) High Hills

pla <= 20% and val <= 20% and lhill <= 100% and hill <= 100% and hhill > 20% and mount <= 20% and hmount <= 20% and tab <= 20%

(15) Mountains

pla <= 20% and val <= 20% and lhill <= 100% and hill <= 100% and hhill <= 100% and mount > 20% and hmount <= 20% and tab <= 20%

(16) High Mountains

pla <= 20% and val <= 20% and lhill <= 100% and hill <= 100% and hhill <= 100% and mount <= 100% and hmount > 20% and tab <= 20%

(17) Plains Tableland

pla > 20% and val <= 20% and lhill <= 20% and hill <= 20% and hhill <= 20% and mount <= 20% and hmount <= 20% and tab > 20%

(18) Tablelands Low Hills

pla <= 20% and val <= 100% and lhill > 20% and hill <= 20% and hhill <= 20% and mount <= 20% and hmount <= 20% and tab > 20%

(19) Tablelands Hills

pla <= 20% and val <= 100% and lhill <= 100% and hill > 20% and hhill <= 20% and mount <= 20% and hmount <= 20% and tab > 20%

(20) Tablelands High Hills

pla <= 20% and val <= 100% and lhill <= 100% and hill <= 100% and hhill > 20% and mount <= 20% and hmount <= 20% and tab > 20%

(21) Tablelands Mountains

pla <= 20% and val <= 100% and lhill <= 100% and hill <= 100% and hhill <= 100% and mount > 20% and hmount <= 20% and tab > 20%

(22) Tablelands High Mountains

pla <= 20% and val <= 20% and lhill <= 100% and hill <= 100% and hhill <= 100% and mount <= 100% and hmount > 20% and tab > 20%