Spatial Ecology of Adélie Penguin Breeding Colonies:

The Effects of Landscape, Environmental Variability and Human Activities

by

Phillippa Kate Bricher B.A., Grad. Cert. GIS.

(University of Tasmania)

A thesis submitted in partial fulfilment of the requirements for a Masters of Environmental Management Degree at the School of Geography and Environmental Studies, University of Tasmania (December, 2006).

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma in any tertiary institution, and to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Signed

Phillippa Kate Bricher, BA, Grad. Cert. GIS.

Adélie penguin, Whitney Pt, Windmill Is, Dec. 2005.

Abstract

Adélie penguins have been widely studied as an "indicator" species for the health of the Southern Ocean ecosystem. However, the effects of climatic variability and human activities on Adélie penguin populations are poorly understood. As many of the Adélie penguin colonies used for long-term demographic studies are located near research stations, there is a need to be able to disentangle the effects of human activities and environmental variability on Adélie penguin populations. This study investigates the landscape properties that drive the locations of Adélie penguin colonies in the Windmill Is, East Antarctica. It also examines whether potential changes in snow cover and/or proximity to human activities best explain the varying population trends of colonies in two breeding localities. While some colonies have been abandoned, or have undergone strong population decreases, the populations of others have grown by more than 1000% in the past 38 years.

This study uses Geographic Information Systems to generate spatial data of landscape, snow accumulation patterns and proximity to human activity parameters. Landscape parameters are derived from fine-scale digital elevation models (DEMs) and snow accumulation patterns are modelled using a complex physically-based GIS model. The parameters are then combined into multivariate statistical models to generate predictions of habitat suitability.

Individually, the landscape attributes, such as elevation, slope, solar radiation, and wetness index, have little power to predict the distribution of colonies within a breeding locality. On the other hand, multivariate models (discriminant analysis and decision tree) derived from these landscape attributes predict the presence or absence of colonies in test grid cells with up to 78.9% accuracy. General rules to describe the distribution of Adélie penguin colonies are not easily derived, as habitat suitability appears to be driven by complex interactions between landscape attributes.

At Whitney Pt, the study site farthest from Casey, modelled snow accumulation parameters explain most of the variation in population trends among colonies (up to 83.7% accuracy, for five classes). At Shirley I, 500 m from Casey, models derived from proximity to human activity parameters correctly predict the trend classes for up to 83.8% of test cells, while models derived from snow accumulation parameters correctly classify up to 57.8% of test cells. This suggests that while snow accumulation patterns are a primary driver of variation in population trends among colonies, the

effect of snow accumulation is outweighed by the effects of human activities near Casey.

Acknowledgments

This study would not have happened without my two dedicated, passionate and patient supervisors – Dr Arko Lucieer and Dr Eric Woehler. This thesis has greatly benefited from Arko's gentle encouragement, answers to an endless string of questions and excellent problem solving skills. Without Eric, the project would never have started. I'd like to thank him for taking me to Antarctica in the first place, and for a thousand explanations and corrections.

I'd like to thank all members of the 59th ANARE to Casey, especially the volunteer penguin counters, the dive team's water taxi service and the station leader, Marilyn Boydell, for facilitating my fieldwork.

The Australian Antarctic Data Centre played an invaluable role by providing me with data, equipment and advice and by answering all my questions. Dr Ben Raymond's statistical expertise and teaching helped enormously with designing the project and with interpreting the results.

At the University of Tasmania's School of Geography and Environmental Studies, the support of several colleagues is gratefully acknowledged. Mr Rob Anders' efforts to generate the photogrammetric DEMs, despite less than ideal photographs and ground control data made this project possible and proved to be character-building for us both. Many thanks also to Dr Jon Osborn for his unstinting support and advice on the photogrammetric aspects of the study. Dr Neil Adams (IASOS) was outstandingly kind in speedily providing me with data and advice on how to use the NCEP/NCAR weather reanalysis data. Peter Morgan gathered vital ground control data at short notice. In addition, there were numerous hallway conversations throughout the year with staff and fellow students that produced new insights. Thanks to Gary Sjoberg (Flinders University) for his lessons on formatting large and unwieldy documents.

Finally, thanks go to my parents for their belief and support.

This project was financially and logistically supported by the Australian Antarctic Division, with authorisation under ASAC project 1219 and animal ethics approval by the University of Tasmania Animal Ethics Committee (Animal Use Permit No A0008581). This project has also received financial and logistic support from School of Geography and Environmental Studies and the Centre

for Spatial Information Science at the University of Tasmania.

Some or all of the data used within this paper was obtained from the Australian Antarctic Data Centre (IDN Node AMD/AU), a part of the Australian Government Antarctic Division (Commonwealth of Australia). The data is described in the following metadata records:

"Windmill Islands 1:50000 GIS Dataset" Ryan, U. (1999, updated 2006).

"Windmill Islands penguin colonies digitised from 1990 Linhof aerial photography" Ryan, U. and Woehler, E. (2000, updated 2006).

Contents

1 Introduction

The effect of climate change on the environment is an issue of current public and scientific concern. Some of the most significant climate alterations have been observed in polar regions (e.g. Fraser and Patterson, 1997; Croxall. et al., 2002; Ainley, 2002; Forcada et al., 2006). In addition, Antarctica is designated as a wilderness zone to be protected under the Antarctic Treaty. Therefore, an understanding of the effects of climate change is particularly critical for management of Antarctic environments.

Seabirds have been widely used as indicators of changes in the Southern Ocean ecosystem (e.g. Micol and Jouventin, 2001; Croxall et al., 2002; Kato et al., 2002; Kato et al., 2004). There are a number of reasons for this, including their perceived primary role in the Southern Ocean ecosystem and the ease with which they can be monitored (Micol and Jouventin, 2001; Kato et al., 2002). However, it has also been acknowledged that using birds as bioindicators of climate change is problematic because of the complex nature of the numerous interactions in the Southern Ocean ecosystem (Croxall et al., 2002) and the potential confounding effects of human impacts at local scales. Adélie penguin (*Pygoscelis adeliae*) colonies are known to be abandoned and recolonised as the climate changes (Ainley, 2002; Emslie and Woehler, 2005) and have hence been termed "bellwethers of climate change" (Ainley, 2002).

1.1 Climate variability and Adélie penguin populations

The localities used by breeding Adélie penguins are affected by interactions between the terrain and local climatic conditions. One of the key facets of this study is the investigation of the impact of snow accumulation on the distribution of colonies and on their population trends over 46 years. One previous study attempted similar analyses (Fraser and Patterson, 1997). That study used a hillshade model as a surrogate for wind exposure to penguin colonies. Wind exposure was, in turn, used as an indicator for areas where snow would be abraded. The present study applies a more complex snow accumulation model, based on the physics of drifting snow and available meteorological data (Wallace, 2005).

Most recent investigations of Adélie penguin population trends have focused on climate change, and especially changes in sea-ice extent (e.g. Trivelpiece and Fraser, 1996; Croxall et al., 2002;

Kato et al., 2002; Kato et al., 2004; Forcada et al., 2006). Others have attempted to separate the effects of climate changes and human activities (e.g. Fraser and Patterson, 1997; Micol and Jouventin, 2001). The present study followed the latter approach, and attempts to differentiate between climatic and human-induced effects.

Many studies of the effects of climatic variation on Adélie penguin population trends have focused on broader-scale variables, such as sea-ice extent. These studies have generated somewhat contradictory results. There is debate about the extent to which these studies have been able to show clear patterns or causal mechanisms for population fluctuations (Croxall et al., 2002; Ainley et al., 2003). It is possible that part of the reason for these contradictory results is that other environmental factors are confounding or exacerbating the effects of sea-ice changes at different sites and in different years (Fraser and Trivelpiece, 1996; Fraser and Patterson, 1997; Ainley et al., 2003). This study attempts to increase understanding of local environmental effects that alter the suitability of individual colony sites. This will, in turn, improve the interpretation of local, regional and ecosystem-scale population trends.

1.2 Human impacts and Adélie penguins

Human activities have had substantial impacts on the physical environment of Antarctic coastal areas (e.g. Young, 1990; Wilson et al., 1990; Micol and Jouventin, 2001). For Adélie penguins, this effect has been most severe where penguin colonies have been destroyed for the construction of research stations and associated infrastructure (Wilson et al., 1990; Micol and Jouventin, 2001). There is argument about the potential impact of human activities outside the immediate footprint of research stations (Wilson et al; 1989; Culik et al., 1990; Wilson et al., 1991; Woehler et al., 1994; Giese, 1996; Fraser and Patterson, 1997; Micol and Jouventin, 2001; Pfeiffer and Peter, 2004). Woehler et al. (1994) proposed that decreasing populations in some penguin colonies were the result of pedestrian visits by station personnel.

As the number of people visiting Antarctica increases, so does concern about the potential impacts of human disturbance on Antarctic wildlife (Pfeiffer and Peter, 2004). In 2005/06, 26 245 tourists visited Antarctica on tourism vessels (IAATO, 2006). In addition to this, almost 4000 live in research stations located throughout Antarctica during summer (COMNAP, 2006). The debate on the potential impact of human activities has been particularly intense on the

Antarctic Peninsula, where tourism is concentrated (Fraser and Patterson, 1997; Pfeiffer and Peter, 2004). However, it is also potentially an issue around all Antarctic research stations.

Long-term studies of Adélie penguin populations have typically been generally conducted near research stations, where human activities are focused (e.g. Woehler et al., 1994; Fraser and Patterson, 1997; Micol and Jouventin, 2001; Woehler et al., 2001). Studies of Adélie penguins have also generally involved nesting birds. This means that it may be difficult to disentangle any effects of climatic variability and the role of human activities on numbers of breeding Adélie penguins. Clarke and Kerry (1994) raised concerns about the effects of invasive monitoring procedures on the validity of scientific observations of Adélie penguins at Béchervaise Island, near Mawson.

1.3 The role of GIS in studying these phenomena

In recent years, Geographic Information Systems (GIS) have been used extensively for habitat analysis of plant and animal species across the globe (e.g. Glenz et al., 1991; Manel et al., 1999; Lenton et al., 2000), as the development of GIS software has made it possible to include spatial variability data into ecological studies (Maurer, 1994). However, GIS has rarely been used to examine the land-based habitat requirements of Adélie penguins, with the exception of Fraser and Patterson (1997). Historically, attempts to study the nest-site requirements of Adélie penguins have been forced to ignore the spatial variability of terrain in and among colonies. This was largely because of the inability of available analytical techniques and computing power to adequately examine spatial data (Yeates, 1975; Moczydlowski, 1986 and 1989; Maurer, 1994; Evans, 1991). Some studies of human impacts have incorporated some limited assessment of the spatial variability of human activities (e.g. Wilson et al., 1990; Young, 1990; Woehler et al., 1994; Fraser and Patterson, 1997; Patterson et al., 2003).

A high-resolution digital elevation model allows fine-scale features to be captured and quantified, and their role in the distribution and population trends of penguin colonies to be investigated. GIS modelling of landscape parameters such as drainage and snow accumulation is more efficient than manually measuring these phenomena, and in the case of snow accumulation, allows historical trends and relationships to be examined from available long-term datasets (Orndorff and Van Hoesen, 2001). This enables an examination of the physical landscape characteristics of Adélie penguin colonies in much greater detail than in previous studies, and an assessment and quantification of the spatial variability of nesting sites (Yeates, 1968; Moczydlowski, 1986, 1989).

Historically, most studies have looked at population trends for what are here termed breeding localities – areas that contain several colonies (using the definition of Woehler et al. (1991, 1994). There has been little examination of the spatial variability of demographic data within breeding localities. Exceptions to this include Woehler et al. (1994) who reported correlations between the population trends of colonies and their distance from Casey, Fraser and Patterson (1997) who compared the role of variability in wind exposure with the population trends of colonies, Wilson et al. (1990) who investigated the effects of human disturbance associated with the Cape Hallett research station on penguin breeding success and Patterson et al. (2003) who examined the relationships between snow accumulation, tourist visits and colony population trends.

In the Windmill Is, individual Adélie penguin colonies have exhibited different population trends. Several of the colonies closest to Casey have undergone population decreases during the 50 years of human occupation in the Windmill Is region (Woehler et al., 1994; E.J. Woehler, unpub. data). However, the overall Adélie penguin population of the Windmill Is trebled between 1961/62 and 1989/90, with the populations of many colonies increasing, and new colonies established at many breeding localities (Woehler et al., 1991). This trend has continued to the present (E.J. Woehler, unpub. data). An analysis of fine-scale processes is needed to contribute to our understanding of the observed variability.

1.4 Adélie penguins

Adélie penguins have been intensively studied because of their perceived primary role in the Southern Ocean ecosystem, and the ease of access to them for study (Giese, 1996; Micol and Jouventin, 2001; Kato et al., 2002; Ainley, 2002). Adélie penguins are considered an 'indicator' species for the health of the whole ecosystem (e.g. Ainley, 2002). With the exception of Emperor penguins (*Aptenodytes forsteri*) penguin colonies are typically located on coastal ice-free sites (Trivelpiece and Fraser, 1996; Ainley, 2002). These are similar to the requirements for research stations, and many stations have penguin colonies nearby. Adélie penguins are strongly philopatric, and their colonies are easily observed, unlike cryptic-nesting species such as Wilson's storm-petrels (*Oceanites oceanicus*).

One of the advantages in studying the distribution of Adélie penguins is that their current and former spatial distributions can be easily mapped. The birds form colonies of up to thousands of pairs. Adults build nests from small pebbles, collected from surrounding areas, to raise their eggs/chicks above the ground and so protect them from snow and meltwater (Ainley, 2002). In the Windmill Islands, these accumulations of nest pebbles have been shown to be up to 9000 years old (Emslie and Woehler, 2005). The perimeters of existing and former colonies can be clearly seen in aerial photographs and on the ground.

Regional trends have been identified in Adélie penguin populations across Antarctica. East Antarctic populations have shown sustained increases; populations on the Antarctic Peninsula have increased and decreased, and those in the Ross Sea showed no clear pattern (Woehler and Croxall, 1998; Woehler et al., 2001). Most population studies have examined regional trends, and there is a need for better understanding of finer-scale variation within regions.

Adélie penguins can feed a considerable distance out to sea. A tracking study at Shirley I, near Casey, found that breeding penguins travelled between 31 and 144 kilometres from the colony (Wienecke et al., 2000). Another study of the foraging range of penguins at Shirley I found that they had a maximum foraging range of 135 kilometres (Kerry et al., 1997). These findings are broadly consistent with studies in other parts of the continent which have found that Adélie penguins feed between 2 and 100 kilometres from the colonies, with the short distances associated with extensive fast-ice (e.g. Kerry et al., 1995; Watanuki et al., 1997; Ainley, 2002). Given the large distances Adélie penguins travel to feed, and the short distances between colonies within a breeding locality, it is considered that land-based influences on Adélie penguin colonies are more likely to explain population trend differences among colonies than differences in the marine environment.

1.5 Definitions

In the penguin literature, the terms "colony", "rookery", "sub-colony" and "breeding locality" have been used in contradictory and ambiguous ways. The definition of what constitutes a colony is unclear for many seabird species. Wittenberger and Hunt (1985) proposed that a continuum exists from solitary to colonial nesting, and that the decision on whether neighbouring groups of birds should be described as discrete colonies depends on the degree of interaction among the groups. This study uses the definitions in Woehler et al. (1991) and Woehler et al. (1994): A breeding colony is here defined as an area of contiguous nest territories. In turn, a nest territory is defined as an area containing a nest, and which is defended by a breeding pair, and is typically approximately 1m^2. A breeding locality is a geographical feature, either an island or a discrete area of mainland, on which breeding colonies are found. Thus, Whitney Pt contains 48 colonies (sensu Woehler) and is considered to be one breeding locality. This contrasts with the definition of Ainley (2002) who used the term "colony" for what is here termed a breeding locality (Fig. 1.1), and with the term "rookery" which was historically used to describe breeding localities (e.g. Penney, 1968).

Sites that contain conspicuous and clearly outlined agglomerations of nest pebbles, and are not known to have been used by breeding pairs of penguins during the period of human occupation in Antarctica have often been described as "relict" colonies (e.g. Penney, 1968; Woehler et al., 1994; Emslie and Woehler, 2005). This study follows that definition, but uses the term "relic" rather than "relict" to refer to these unused colonies. This change follows the Oxford Dictionary's (Pearsall, 1999) definitions as follows:

> *"**Relic n.** 1 an object of interest surviving from an earlier time"*
>
> *"**Relict n.** 1 an organism or other thing which has survived from an earlier period. > Ecology: a population, formerly more widespread, that survives in only a few localities."*

By these definitions, the surviving remnant of a penguin colony whose population is decreasing might be defined as relict. "Relic" is a more appropriate term to describe colonies that are not used presently, and is thus used here.

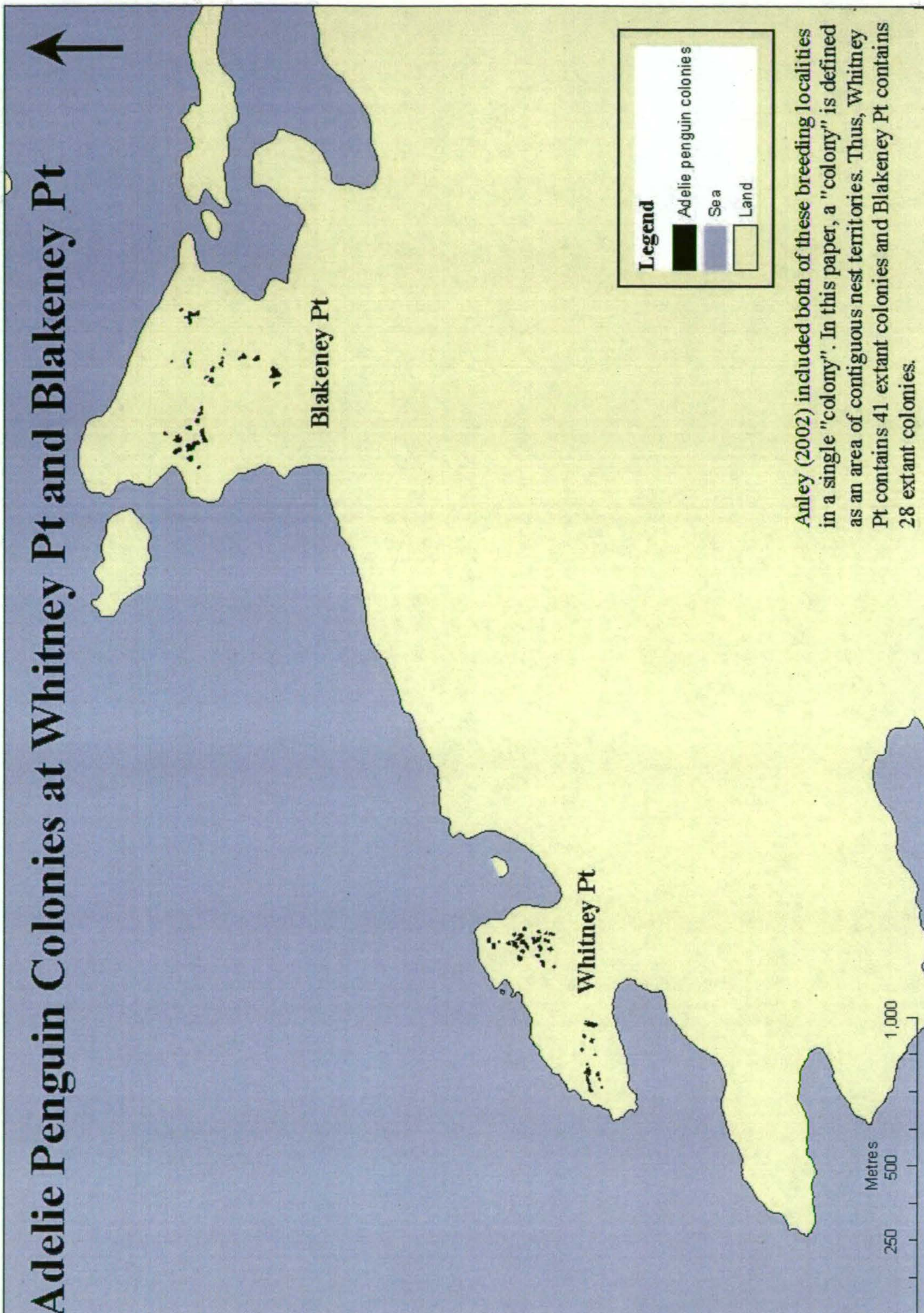


Fig. 1.1: This area of the Windmill Is, East Antarctica, was identified by Ainley (2002) as a single colony. However, using the definitions of Woehler et al. (1994), this map shows two breeding localities, each containing several colonies.

1.6 Aims and objectives

The study aims to:

- Quantify spatial landscape parameters (slope, drainage, aspect, solar radiation, planar and profile curvature, surface roughness) and climatic parameters (wind exposure and snow accumulation) from fine-scale digital elevation models (DEMs) of the two study sites

- Apply multivariate statistical analyses to investigate the importance of static landscape parameters in influencing the distributions of Adélie penguin colonies at the two study sites

- Determine the contribution of selected climatic variables (snow accumulation patterns and wind exposure) to the observed long-term population trends of Adélie penguin colonies at the two study sites, using multivariate statistical analyses

- Investigate the ability of proximity to Casey and the main Shirley I access point and exposure to potential air-borne emissions from Casey to explain the observed population trends of colonies on the island using multivariate statistical analyses

1.7 Hypotheses

This study investigates selected aspects of the spatial ecology of Adélie penguin breeding localities. It examines whether selected parameters of the landscape can predict the locations of Adélie penguin colonies within the breeding localities at Whitney Pt (66° 15'S, 110° 32'E) and Shirley I (66°17'S, 110°29'E) near Casey, Wilkes Land, East Antarctica. The study also investigates whether the interaction of these parameters and snow accumulation patterns, or proximity to human activities can predict the population trends of penguin colonies at two sites.

This can be expressed as the following null hypotheses:

H_{NULL} 1 Static landscape variables (slope, drainage, aspect, planar and profile curvature, surface roughness, wind exposure, snow cover and solar radiation) cannot predict the locations of current and relic Adélie penguin colonies at Shirley I and Whitney Pt.

H_{NULL} 2 Interactions between the shape of the land and the weather conditions that drive snow accumulation patterns cannot predict the population trends of Adélie penguin colonies at Shirley I and Whitney Pt.

H_{NULL} 3 Proximity and exposure to human activities associated with Casey cannot predict the population trends of Adélie penguin colonies at Shirley I and Whitney Pt.

2 Literature Review

2.1 Spatial ecology of Adélie penguins and the effects of landscape processes

2.1.1 Adélie penguin distribution

Adélie penguins have a circumpolar breeding distribution between 60° and 77°S. The global population has been estimated at approximately 2.4 million breeding pairs, at some 170 breeding localities. The birds nest on ice-free rocky shores with landing beaches, where there is access to open water for feeding. Colony sites are believed to be chosen because they have ready access to the sea, are exposed to prevailing winds, have gentle slopes that allow good drainage and discourage snow accumulation, and have a supply of suitable pebbles for nest construction (Yeates, 1975; Trivelpiece and Fraser, 1996; Ainley, 2002).

The spatial distribution of any species can be viewed at a nested hierarchy of scales, with the spatial pattern varying according to the scale (Maurer, 1994). At the broadest scale – that of the entire Antarctic continent – Adélie penguins breed where there are exposed rocky areas with landing beaches (Falla, 1937, in Ainley, 2002). Viewed from a regional scale, in an area such as the Windmill Islands, breeding localities are patchily distributed. Within each breeding locality, penguins are clustered into colonies and within an individual colony; the nest territories of penguins are mostly contiguous. Most studies that address the spatial distribution of Adélie penguin colonies have been conducted at the broader scales.

Many of those studies focused on the marine environment and variability in parameters such as prey availability and sea-ice extent (e.g. Ainley and Le Resche, 1973; Fraser et al., 1992; Kerry et al., 1995; Fraser and Trivelpiece, 1996; Kato et al., 2002; Forcada et al., 2006). Earlier studies suggested that the distributions of seabirds were primarily controlled by prey availability (e.g. Voous, 1965, in Fraser and Trivelpiece, 1996). From the 1970s onwards, studies suggested that distributions were also constrained by variability in the marine environment, such as sea-ice extent, and variations in sea temperatures, salinity and mixing depths (e.g. Ainley and LeResche, 1973; Fraser and Trivelpiece, 1996)

Ainley (2002) argued that breeding localities were geographically structured by a combination of available resources and by intra- and inter-species competition (Ainley, 2002). The resources included physical factors, such as suitable nesting sites, and biological factors, such as prey availability. Prey availability appeared to be the primary driver of the total number of birds in a region, and competition for food exacted a negative effect on population clumping (Ainley et al., 1995; and reported in Ainley, 2002). Where a locality had a large breeding population, the localities within a 150-200 kilometre radius were typically found to have small populations (Ainley et al., 1995). In addition to the breeding birds, each colony had a population of non-breeding birds that visited the colony and fed farther out to sea (Birt et al., 1987; Ainley, 2002). In contrast to the negative effect of competition for resources, the natal philopatry and social tendencies of Adélie penguins were found to have a positive effect on the clumping of colonies, in that while nesting sites and prey resources were available, Adélie penguins remained close to their birth colony (Ainley, 2002).

During the summer chick-rearing period, breeding Adélie penguins are central place foragers (Ropert-Coudert et al., 2004). The distances they travel to feeding areas vary throughout the breeding season and among breeding localities. Satellite-tracking studies have found that the birds travel up to 200 kilometres from the colony to feed during the chick provisioning period, with the shortest distances associated with areas of fast-ice, where the penguins walk to the foraging grounds (e.g. Kerry et al., 1995; Watanuki et al., 1997; Ainley, 2002). Tracking studies at Shirley I (Wienecke et al., 2000) found that breeding penguins feed up to 31-110km from the colony during the guard stage and 94-144km during the crèche stage (Kent et al., 1998; Wienecke et al., 2000). If Adélie penguins in the Windmill Islands feed between 30 and 140 kilometres from the breeding localities, it appears likely that differences in population trends among colonies located less than 100m apart are driven by factors related to the terrestrial and social environment, rather than marine environment.

2.1.2 Coloniality

Adélie penguins are strongly colonial birds. Theories of why they nest colonially include nest-site availability; anti-predator strategies; access to mates; and social factors related to breeding, such as information transfer (Ainley et al., 1995). Wittenberger and Hunt (1985) noted that 98% of marine birds nest in colonies. They proposed that many seabird nest sites are more clumped than they would be if simply constrained by the available suitable habitat. This is demonstrated in situations where neighbouring potential nest habitat remains unoccupied while one colony becomes crowded. This pattern occurs in the Windmill Islands, where relic colonies occur within 50m of extant colonies that contain hundreds of pairs. Wittenberger and Hunt also noted that colonies provide protection against predation in the form of increased vigilance, but at the same time they attract predators by providing a concentration of available food and they may also be more prone to disease.

It may be that for Adélie penguins, a shortage of available rocky coast forces some degree of nest clumping, that makes them unable to take advantage of one of the benefits of solitary nesting – that of concealment from predators. Studies have found that when Adélie penguin breeding localities are under stress, the effects are most strongly exhibited in smaller colonies (Giese, 1996; Fraser and Patterson, 1997). Fraser and Patterson argued that a population below 25-30 pairs was unable to maintain the colony's defences against predation by skuas (*Catharacta spp.*) on the Antarctic Peninsula.

2.1.3 Topographic influences on colony locations

Fraser and Patterson (1997) used the term "landscape effect" to describe the influence that the shape of the land exerts on Adélie penguin colonies. This described a phenomenon recognised by the earliest Antarctic explorers – that Adélie penguins not only require ice-free rocky areas for nesting, but that they also select those sites where snow does not accumulate (Levick, 1915).

It has since been argued that snow accumulation, meltwater runoff and solar radiation influence the selection of Adélie penguin nesting sites, and that the abandonment of a colony can occur rapidly after two or more years of failed breeding (Yeates, 1975; Moczydlowski, 1986, 1989; Trivelpiece and Fraser, 1996; Fraser and Patterson, 1997).

Snow cover has repeatedly been found to be one of the most important drivers of nest site selection by Adélie penguins (Levick, 1915; Yeates, 1975; Moczydlowski, 1986; Moczydlowski, 1989; Trivelpiece and Fraser, 1996; Fraser and Patterson, 1997; Ainley, 2002). Therefore it is not surprising that changes in snowfall and wind regimes should also have been found to alter their breeding success and influence population trends, as recently shown by Fraser and Patterson (1997).

Ainley (2002) wrote that Adélie penguin colonies typically occur on ridges and higher ground, and that where they share breeding grounds with congeneric birds, Adélies are found farther from landing beaches. He argued that this is related to the conditions at the sites during the Last Glacial Maximum (19 000 y bp) when the species was under greater ecological pressure, and when land-ice lowered the height of the land by several metres, submerging gently sloping beaches. In more southern areas, where Adélie penguins nest in single-species colonies and land is in greater demand, he found that they nest closer to sea level .

Adélie penguins are confined to areas where glaciers have formed moraines near the coasts, to provide nest-pebbles. Ainley (2002) argued that the availability of nest pebbles is a crucial driver of colony locations. In the Windmill Islands, most relic and extant colonies occur on raised-beach formations covered in rock debris measuring 2-6cm (Fig. 2.1). Keage (1982) argued that the preference for these formations demonstrates the importance of the availability of nest pebbles in determining colony locations. He wrote that the colony size and nest density are directly related to the availability of nest pebbles, with colony populations increasing with distance from the ice cap. However, he did not address the potential role played by penguins in building up these raised-beach formations by collecting nest-pebbles from surrounding areas.

The present study did not address the availability of nest pebbles for two reasons. First, it is difficult to measure pebble availability and requirements. The number of pebbles used in nest construction is highly variable and ranges from nests built with a few pebbles to nests built from several hundred stones. It is also difficult to measure the spatial distribution of suitably sized pebbles over the area of a breeding locality. Second, if nest pebble availability were a limiting factor on nesting sites in the Windmill Islands, it is likely that the relic colonies would either be occupied or denuded of stones. At both Shirley I and Whitney Pt, there are numerous relic colonies in close proximity to colonies that contain hundreds of breeding pairs, and these relic colonies contain huge numbers of pebbles.

Fig 2.1: Adélie penguin colony at Whitney Pt, January, 2006. This colony is undergoing a strong population increase (3900% more breeding pairs than when the colony was first counted in 1959/60).

Early studies of the relationship(s) between the shape of the landscape and nest site selection by Adélie penguins occurred between the 1960s and 1980s (Yeates, 1975; Moczydlowski, 1986; Moczydlowski, 1989). Both Yeates and Moczydlowski noted the importance of snow distribution in determining the locations of colonies. Yeates found that microclimatic effects were vital for successful breeding, with nest sites exposed to high levels of wind exposure and solar radiation, compared with unoccupied areas. He argued that while microclimate played a vital role in site selection, interannual variability was likely to be caused by macroclimatic variation. Yeates did not address the role microclimate plays in determining the ultimate effect of macroclimate events. It may be that some colonies are in locations that are more prone than others to being covered by snow and hence become more or less suitable as snow cover conditions change (Fraser and

Patterson, 1997).

Moczydlowski (1986; 1989) investigated the terrain properties of colony locations in the South Shetland Is. He found that the common features of all colonies were good drainage and high levels of solar radiation, with Adélie penguins nesting in the sites with the thinnest snow cover at the end of winter. In addition, their colonial nature helped shed snow because their faeces carried high levels of sodium chloride, which lowered the freezing point of water. This in turn aided dispersion of snow from colonies. From this, he concluded that Adélie penguins selected nest sites that are naturally likely to have the least amount of snow, and then, through the deposition of faeces, further increase the site's suitability. In 1986, Moczydlowski found that when penguins were not present, there was no difference in air temperature between colony sites and other parts of the landscape. He also argued that Adélie penguins did not nest in the most exposed sites. Instead, he proposed that they prefer sites with the least snow cover, but also with lower winds. Both Yeates and Moczydlowski's studies were conducted before the widespread availability of GIS as a tool for analysing spatial data. Their studies therefore did not take account of spatial variability within and among colonies.

At Cape Hallett, in the Ross Sea, Adélie penguins are found on well-drained mounds, and where these were flattened by human activities associated with the now-abandoned research station, Adélie penguins did not recolonise after the station closed. However, where those mounds remained or were rebuilt as part of habitat rehabilitation, the penguins reoccupied after the humans left (Wilson et al., 1990).

2.1.4 Climate variability and snow accumulation

As with studies of the geographic distribution of Adélie penguin colonies, most investigations of the relationships between climate variability and Adélie penguin breeding success or population trends have been conducted at broad scales. Many studies have examined the responses of breeding localities to changes in sea-ice extent and other climatic variables (e.g. Ainley and LeResche, 1973; Fraser et al., 1992; Croxall et al., 2002; Kato et al., 2002; Kato et al., 2004; Olmastroni et al., 2004; Forcada et al., 2006). Such studies are useful in studying population trends for entire regions. They are of limited value in attempting to interpret differences in population performance among colonies within a breeding locality, where all the colonies are likely to be subject to similar sea ice conditions and other, finer-scale processes are likely to be involved.

A few studies have attempted to separate the effects of changes in sea-ice extent and snow accumulation patterns at local scales (Trivelpiece and Fraser, 1996; Fraser and Patterson, 1997; Patterson et al., 2003). On the Antarctic Peninsula, rising temperatures over the past 50 years have been accompanied by increasing snowfall and decreases in sea-ice extent – both factors which have been implicated in decreasing Adélie penguin breeding populations (Fraser and Patterson, 1997). Trivelpiece and Fraser (1996) studied population trends at Litchfield I, near Palmer Station on the Antarctic Peninsula. They noted that 18 of 21 colonies that have recently been abandoned were in the lee of prominent topographic features.

Patterson et al. (2003) used a GIS hillshade model as a proxy for snow accumulation on Litchfield I and nearby Torgersen I, in a bid to determine whether snow accumulation or the effects of human visitation could best explain the observed changes in colony population trends. The hillshade model could be seen more accurately as a surrogate for exposure to the prevailing winds. Their study found a strong correlation between wind exposure and population trends, and no statistically significant relationship between the rates of human visitation and population trends. One of the limitations of a hillshade model as a surrogate for snow accumulation is that it models light – which has a laminar flow, whereas wind flows turbulently, and the transport of snow is physically complex (e.g. Kind, 1986; Liston and Sturm, 1998; Green et al., 1999). Thus, a hillshade can only give a first approximation of the patterns of snow accumulation. Such an approximation is appropriate in sites with simple topography, but potentially less useful for fine scale studies in sites with more complex or finer-scale topographies, such as those around Casey, where the maximum altitude is about 35m above mean sea level, and the landscape is dominated by a mix of low cliffs and gently undulating plateaux.

2.1.5 Human impacts on Adélie penguins

The Adélie penguin colonies, for which long-term population records are available, such as those used in examinations of responses to climate variability, are typically located near Antarctic research stations (Ainley and LeResche, 1973; Fraser et al., 1992; Trivelpiece and Fraser, 1996; Fraser and Patterson, 1997; Fraser, 1998; Micol and Jouventin, 2001; Olmastroni et al., 2004; Emslie and Woehler, 2005). This makes it potentially difficult to separate the effects of climate variability from the effects of human activity at or near the colonies.

A number of studies have investigated the effect of human activities on Adélie penguins. Reasons given for this research include public concern that man's presence in Antarctica may damage the ecosystem (Wilson et al. 1991), the obligations of national research programs under the Antarctic Treaty and concerns that the effects of human activity may affect the results of scientific research (Clarke and Kerry, 1994; Wilson et al., 1989). National Antarctic science programs are obliged to minimise their effect on wildlife and the environment, under the Antarctic Treaty and the Madrid Protocol. The International Association of Antarctica Tour Operators (2006) stated that tourism operators are also obliged to meet the requirements of the Antarctic Treaty. Young (1990) noted that Adélie penguins and humans have very similar requirements in Antarctica, namely access to ice-free terrain near water, and that as human activities increase, so too do the chances of significant effects on Adélie penguins.

Adélie penguins have often been considered to be relatively immune to human disturbances because they do not always display overt distress behaviours (Giese, 1996). However, numerous studies have attempted to investigate the effects of human activities on pygoscelid penguins. These activities include the destruction of penguin colonies (Micol and Jouventin, 2001); other alterations to the terrain from station construction (Wilson et al., 1990); aircraft flying over colonies (Culik et al., 1990; Wilson et al., 1991); manipulation of the birds during scientific studies (Wilson et al., 1990; Clark and Kerry, 1994, Giese, 1996) and pedestrian visits to colonies (Culik et al., 1990; Wilson et al., 1991; Woehler et al., Giese, 1996; 1994; Fraser and Patterson, 1997; Holmes et al., 2006). Measures used to determine the effects on penguins include behavioural changes, physiological changes such as heart rate (Wilson et al., 1991), changes in feeding behaviour (Wilson et al., 1989); and changes in breeding success or colony population trends (Woehler et al., 1994; Giese, 1996; Fraser and Patterson, 1997; Patterson et al., 2003).

The results of these studies have been somewhat contradictory and ambiguous. Many of the studies that have found a negative impact have used short-term measures, such as heart rates or behavioural responses. For example, Wilson et al., (1991) found that while birds may appear unconcerned by the approach of humans, heart rates increased from 80 to 127 beats per minute when approached by a researcher on foot. Another experiment showed that comparatively brief and non-invasive human handling resulted in an approximate 50% increase in the duration of foraging trips (Wilson et al., 1989). The researchers in that study concluded that this represented a "psychological" effect on the birds.

There is also some direct evidence of a long-term effect on Adélie penguin numbers as a result of human visits. Giese (1996) studied the effects of scientific nest checks conducted every second day and tourist-style visits two to four times every day in Adélie penguin colonies that had previously been exposed to little human activity. She found that colonies subjected to both treatments had lower breeding success than control colonies. This difference was significant in small colonies (~40 pairs) and not significant for larger colonies (~70 pairs). Giese argued that the effect of disturbance was exacerbated in smaller colonies and that it was most closely linked to the frequency of disturbance rather than the intensity of the disturbance.

Woehler et al., (1994) proposed that visits to colonies by station personnel were responsible for observed decreases in breeding success and populations of colonies at the end of Shirley I nearest Casey. An examination of Adélie penguin population trends on two islands near Palmer Station on the Antarctic Peninsula was unable to find a link between population trends and human activities (Fraser and Patterson, 1997). In that study, the most heavily visited island, Torgersen I, was also the one with the smallest decrease in Adélie penguin numbers. Young (1990) found that Adélie penguin numbers in colonies close to the research station at Cape Bird declined significantly, while the overall number of penguins in the breeding locality increased. Those colonies closest to the station were the ones that had been most intensively studied and were also within 200m of a helicopter landing pad.

At Cape Hallett, in the Ross Sea, Adélie penguins are known to nest on well-drained mounds (Wilson et al., 1990). The total population decreased from 62 900 pairs to 37 000 between 1959 and 1968, when the Cape Hallett research station was in use. Station construction work led to the destruction of some mounds and the construction of buildings near others changing the snow accumulation regime. The station was abandoned in 1973 and demolished in the 1980s. Wilson et al. found that penguins reoccupied few areas where humans had altered the terrain. A combination of land acquisition by man, disturbance and scientific study was blamed for the observed population decline. Scientific practices such as banding and handling of the birds were most closely correlated with population declines in individual colonies. It took 12-14 years after the research station was abandoned for the penguin population to reach the size it was before human occupation.

Micol and Jouventin (2001) investigated the effects of station activities and the construction of a runway on the populations of 7 seabird species nesting near Dumont d'Urville, Antarctica. They found that despite the destruction of 10% of the region's Adélie penguin nests, the total number of Adélies had increased by 50% during the study period. On Ile des Pétrels, where the station buildings are located and helicopters operate each summer, the number of Adélie penguins increased by 250%. Micol and Jouventin noted that they could not quantify the effect of environmental factors such as sea-ice extent, food availability and nest-site availability. However, they suggested that in the long term, these factors outweighed the apparently significant short-term effects of human construction activities.

It is unclear why the penguin populations of Cape Hallett and Dumont d'Urville should display such different responses to station activities. It may be that other environmental factors have confounded the results of one or both of these studies. Little of the literature has involved examination of variation among colonies within one breeding locality, which is the scale at which such impacts are most likely to be seen.

2.2 GIS habitat modelling

Geographic Information System analysis has rarely been used to examine the spatial ecology of Adélie penguins (Patterson et al., 2003). However, GIS has been widely used to investigate relationships between the landscape and numerous other species across the world (e.g. Aspinall and Veitch, 1993; Baker et al., 1995; Bian and West, 1997; Store and Kangas, 2001). GIS has enabled ecological studies to quantitatively analyse spatial variability (Burrough and McDonnell, 2000; Vogiatzakis, 2003). Store and Kangas (2001) also noted that GIS applications can generate new data by spatial analysis of existing data. Typically, GIS-based habitat analyses have involved the creation of spatial data layers, each of which represents one habitat parameter. These layers have then been combined, using some function – derived from either expert knowledge or statistical testing - to produce a map showing the relative quality of habitat for the species being examined (Store and Kangas, 2001).

GIS habitat modelling has been widely used in autecology (Guisan and Zimmermann, 2000; Kidd and Ritchie, 2000; Lenton et al., 2000; Osborne et al., 2001; Lauver et al., 2002; Gibson et al., 2004). It has also been used for many reasons related to the interactions between human land-uses and the environment (Guisan and Zimmermann, 2000). These have included the identification of conservation priorities (Bian and West, 1997; Walker and Craighead, 1997), field work and management planning (Curnutt et al., 2000), distribution and abundance modelling (Aspinall and Veitch, 1993), environmental impact assessments, as a step in refining habitat maps (Breininger et al., 1991), and to model changes in habitat suitability and fragmentation through time (Hansen et al., 2001).

Habitat prediction models have been found to produce stronger results for species that are common, range-restricted or more specialised than for species that are rare, wide-ranging and/or more generalised in their habitat requirements (Pereira and Itami, 1991; Debinski et al., 1999). Adélie penguins can be considered to be common within breeding localities, though their specific habitat requirements are poorly understood (Yeates, 1975; Moczydlowski, 1986, 1989

Numerous approaches have been used to drive GIS-based habitat models, and one common classification method has been to split such models according to whether they are informed by field data (empirical models) and or by expert knowledge (rule-based models) (Guisan and Zimmermann, 2000; Store and Kangas, 2001). Store and Kangas (2001) argued that rule-based models were most suitable for situations in which it would be too expensive or time consuming to gather empirical data.

Rule-based models have been used for species for which good expert knowledge of habitat requirements was available, and the distribution poorly understood (e.g. Lauver et al., 2002). This approach has been considered suitable for those cryptic species whose presence/absence is difficult to map (Gibson et al., 2004). Rule-based models have often used similar techniques to multi-criteria decision making analysis (Pereira and Duckstein, 1993; Lenton et al., 2000), which is a common technique in GIS. In models using this approach, habitat parameters have typically been assigned values based on expert knowledge, and these factors then combined to produce habitat suitability maps (Breininger et al., 1991; Pereira and Duckstein, 1993; Curnutt et al., 2000; Hansen et al., 2001; Store and Kangas, 2001). Areas of suitable habitat have been defined as those areas in which all habitat factors coincide (Kelly et al., 2001). One criticism of this approach is that validation of these models has also often relied on expert knowledge, making the entire process somewhat circular (Pereira and Itami, 1991).

Empirical models take a more objective approach to analysing habitat suitability. Such models have typically attempted to quantify the relationship between observed distributions and habitat parameters. The classification of habitat parameters has generally been independent of the wildlife data (Aspinall and Veitch, 1993). Adélie penguins are suited to empirical modelling because the penguins' present and past distribution is relatively easy to measure. In contrast, there is little expert knowledge available appropriate for forming rules to determine the location of colonies, as such processes are poorly understood (Yeats, 1975; Moczydlowski, 1986 and 1989; Wilson et al., 1990). The conspicuousness of Adélie penguin nests is in contrast to the problems encountered by researchers investigating bird species whose nest sites are cryptic and whose total populations and distributions must be inferred (Guisan and Zimmermann, 2000; Gibson et al., 2004).

Store and Kangas (2001) argued that the accuracy of the results of habitat analysis depends on the quality of the source data and of the analytical techniques. The most important factors affecting the quality of spatial data have been described as currency, completeness, consistency, accessibility, accuracy and precision, as well as the error sources inherent in the data-gathering processes. Error propagation also occurs through the analysis process, particularly as a result of combining data sets with different spatial and temporal scales (Burrough and McDonnell, 2000; Store and Kangas, 2001).

Input data for GIS habitat analyses have typically been derived from three main sources – remotely sensed photographs and images (Breininger et al., 1991; Aspinall and Veitch, 1993; Bian and West, 1997; Hansen et al., 2001; Osborne et al., 2001; Gibson et al., 2004); paper maps (Raphael et al., 1995; Lenton et al., 2000); and DEM derivatives (Raphael et al., 1995; Blackard and Dean, 1999; Guisan and Zimmermann, 2000; Gibson et al., 2004). Input data of animal or plant distribution have been gathered by survey (Aspinall and Veitch, 1993; Osborne et al., 2001; Gibson et al., 2004); or by radio or satellite tracking (Bian and West, 1997).

It has been stated that empirical models tend to decline in accuracy as the number of input variables and environmental complexity increases (Vogiatzakis, 2003). A common source of errors has been in the conversion of the available data to spatial coverages. Guisan and Zimmermann (2000) noted that many of the variables that determine habitat suitability, such as temperature or solar radiation are often interpolated from widely spaced monitoring stations and are hence prone to interpolation errors. In contrast, they argued that available DEMs and their derivatives tend to be highly spatially accurate. An additional benefit of DEM derivatives is that they may be more cheaply and efficiently generated than manually measured directly causative variables (Lenton et al., 2000; Gibson et al., 2004). Guisan and Zimmermann (2000) proposed that DEMs and their basic derivatives – slope, aspect and curvature – are generally the most accurate maps available, though they may not have the highest predictive potential. DEM derivatives may not have direct physiological relevance for a species, but can act indirectly on causative variables, such as temperature (Gibson et al., 2004). However, Vogiatzakis argued that these surrogate parameters may introduce errors into models.

One result of using DEM-derivatives in a model is that the model may not be readily applied to other geographic areas because the same topographic position in a different region may experience a different environmental gradient (Guisan and Zimmermann, 2000; Austin, 2002; Gibson et al., 2004). However, Guisan and Zimmermann (2000) showed that at local scales, DEM derivatives may show strong correlations with species distributions.

Few GIS habitat analyses have accounted for temporal variability in habitat suitability (Curnutt et al., 2000). Guisan and Zimmermann (2000) noted that historical conditions may have a significant influence on the current distribution of organisms, and that most static models have failed to account for this. They urged researchers to incorporate historical data wherever possible. Guisan and Zimmermann also noted that since temporal data of species' responses to environmental change are rarely available, static models are often the only possible approach. Curnutt et al. (2000) developed a model to account for changing hydrological conditions between years and between management scenarios. Baker et al. (1993) noted that selection of nest sites by sandhill cranes may have occurred when vegetation, climate patterns, water management or disturbance levels were different and that the suitability of an individual nest-site may vary from year to year. However, other studies that examined changing distributions of species have not considered temporal changes in the habitat quality (e.g. Glenz et al., 2001).

The majority of GIS-based habitat analyses has used records of species presence/absence to measure habitat suitability (e.g. Pereira and Itami, 1991; Aspinall and Veitch, 1993). Aspinall and Veitch (1993) suggested that presence/absence data is appropriate for species where only a sample survey is available. Breininger et al. (1991) warned that model development and testing based on animal sightings or radio tracking may not reflect actual habitat suitability, because these may detect a small subset of an area used by the population. This was less of a problem in the present study because of the conspicuous nature of Adélie penguin nests, which reduced the risks of error associated with one-off observations of the species in a particular location. One difficulty with calibrating habitat suitability models has been the fact that any species rarely occupies all of its potential range (Curnutt et al., 2000). Fielding and Bell (1997) noted that prediction errors can occur in habitat models when the habitat is unsaturated. They warned that if the species under examination is not using the entire available habitat, this will generate interference in the model. It has also been argued that it often cannot be determined whether an animal has never, or will never use a particular location (Breininger et al., 1991). However, Curnutt et al. (2000) argued that while the species in question may not appear in all areas with suitable habitat indices, it should not appear in sites deemed to be unsuitable. Fielding and Bell (1997) suggested that such "false negatives" are likely to be the result of errors in either the statistical model or due to some relevant ecological process not being mapped. They warned that appropriate data may not be available for some ecological processes. Guisan and Zimmermann (2000) argued that nature is too complex and heterogeneous to be reduced a single predictive model, no matter how complex that model may be. Breininger et al. (1991) argued that long-term studies of population dynamics are needed to accurately quantify habitat suitability, but noted that this is beyond the scope of many mapping applications.

Scale

It has been argued that the scale at which habitat is analysed can have major implications for the conclusions that can be drawn (Maurer, 1994). Baker et al. (1995) argued that ecologists need to be able to understand how habitat requirements change across spatial scales. They argued that choosing the wrong scale can lead to researchers drawing incorrect conclusions or to an inability to draw any conclusions. They found that as the resolution of analysis increased, there was a corresponding decrease in the ability to detect important habitat variables and draw conclusions. However, most GIS-based habitat analyses have been conducted at just one scale. The choice of scale has typically been driven by a compromise between the scales perceived to be important for the species under investigation, the resolution and extent of the available datasets and the available computational power (e.g. Aspinall and Veitch, 1993; Osborne et al., 2001). Pereira and Itami (1991) noted that different variables are likely to be important at different spatial scales. Lauver et al. (2002) found that loggerhead shrikes occurred in sites their model predicted as low quality habitat, and proposed that this was because the birds were making use of habitat features that were too small to be picked up at the 0.1ha resolution of their model. A few studies have investigated landscape variability at more than one scale. Raphael et al. (1995) modelled murrelet nesting habitat requirements at river basin and site-specific scales. Osborne et al. (2001) used datasets that had been acquired at different spatial scales before combining them into a single predictive model.

Most GIS-based studies have worked at broad scales, with resolutions typically larger than 1ha (e.g. Manel et al., 1999; Glenz et al., 2001; Osborne et al., 2001; Gibson et al., 2004). Exceptions to this include Sieg and Becker (1990), who measured landscape variables within an 11.3m radius of merlin (*Falco columbarius*) nests. If the ability to detect important landscape variables decreases as the resolution increases, as argued by Baker et al. (1995) it seems likely there's a need for more studies that use GIS-techniques to examine spatial variability at fine scales.

2.3 Statistics in GIS-based habitat modelling

A wide variety of multivariate statistical tests have been applied to GIS-based habitat modelling. The simplest models have operated entirely within a GIS, combining the data layers by some mathematical function (Guisan and Zimmermann, 2000; Store and Kangas, 2001). However, Guisan and Zimmermann (2000) warned that these tests have typically been inadequate for model-building as they did not allow stepwise selection procedures or graphical tests of model-fitting. More complex models have made use of the more powerful statistical tools available only in specialised statistical packages (e.g. Blackard and Dean, 1999; Debinski et al., 1999; Manel et al., 1999; Osborne et al., 2001; Gibson et al., 2004).

2.3.1 Univariate testing

The first step of statistical modelling of habitat suitability has often involved univariate exploration of the relationship between the dependent variable and each parameter (Manel et al., 1999). It has been argued that if a species selects habitat based on the measured parameters, those parameters should show differences in the mean and variance between sites where the species is present and absent. Larger variation in the distribution of values has typically been expected in sites where the species is absent (Pereira and Itami, 1991).

2.3.2 Multivariate model selection

It has been noted that the large and evolving range of statistical approaches available for ecological modelling can make it difficult for ecologists to choose appropriate methods (Manel et al., 1999).

Multivariate analyses that have been widely used in GIS-based habitat suitability studies include discriminant analysis (Raphael et al., 1995; Manel et al., 1999; Debinski et al., 1999; Kidd and Ritchie, 2000); logistic regression (Sieg and Becker, 1990; Osborne et al., 2001); decision trees (Hansen et al., 2001), artificial neural networks (Blackard and Dean, 1999; Manel et al., 1999); generalised linear models (Gibson et al., 2004); and Bayesian approaches (Aspinall and Veitch, 1993).

One of the challenges in selecting an appropriate testing method is that few studies have compared the performance of different models. Exceptions to this include Blackard and Dean (1999) who compared the performance of two forms of discriminant analysis and artificial neural networks; and Manel et al. (1999) who compared discriminant analysis, artificial neural networks and logistic regression, in predicting the distribution of several Himalayan river bird species. Blackard and Dean found that an artificial neural network was significantly more powerful than either of the discriminant analysis functions, but found no significant difference in the results of the linear- and non-parametric discriminant analyses. They also noted that the artificial neural network took 2500 hours of computer run time, compared with five minutes for the discriminant analysis. Manel et al. (1999) found that the three methods they compared all had strong predictive power. When tested against calibration data, the artificial neural network outperformed the other two models, and discriminant analysis performed slightly better than logistic regression. When tested against datasets from different geographic areas, logistic regression marginally, but significantly outperformed the other two models, with the artificial neural network the worst-performer. They also found that the results from logistic regressions were most variable across species.

Discriminant analysis has been successfully used in ornithological studies to distinguish suitable habitat and the effect of human visitors on breeding populations (Debinski et al., 1999; Manel et al., 1999; Patterson et al., 2003). However, it is limited in its applicability because of the underlying assumptions of normality, equal variance within each group and equal covariance matrices within each group (Flury and Riedwyl, 1988; Hastie et al., 2001).

Many of the available multivariate statistical tests rely on data that is normally distributed, but many ecological datasets do not meet this requirement (Blackard and Dean, 1999). Guisan and Zimmermann (2000) wrote that data sets can be normalised by a variety of functions, but warned that models built on the artificially normalised data can make biological interpretations difficult.

Manel et al. (1999) noted that multivariate normality can be hard to assess, especially with large numbers of predictor variables. They argued that as normal distributions are an underlying assumption of discriminant analysis, the results should always be treated with caution. However, Blackard and Dean (1999) argued that in practice, the assumptions of discriminant analysis are often violated with minimal apparent effect on the result.

Logistic regression has also been widely used in ecological models (Sieg and Becker, 1990; Pereira and Itami, 1991; Bian and West, 1997; Glenz et al., 1999; Kelly et al., 2001; Osborne et al., 2001). This approach has been considered the most suitable when some of the available data were qualitative and did not meet assumptions of multivariate normality (e.g. Pereira and Itami, 1991).

Hansen et al. (2001) used a hybrid decision tree to classify habitat units. They wrote that this approach allowed them to incorporate different data types in the data analysis. Decision trees are non-parametric and hence make no assumptions about the distribution of the data (Quinn and Keough, 2002). Because decision trees attempt to predict each data point exactly, they avoid the need to characterise the model fit (Guisan and Zimmermann, 2000). Such trees have been found to generate almost as many terminal nodes as there are observations, and to therefore not offer any modelling parsimony. Pruning and cross-validation have typically been used to find an "optimal" balance between the number of terminal nodes and the predictive power (Guisan and Zimmermann, 2000).

2.3.3 Model refinement

Once an initial statistical model has been generated, researchers have typically reduced the number of explanatory variables to a "reasonable" number, in order to enhance the model's accuracy and predictive power. Researchers have done this either arbitrarily, automatically by the statistical programs, by methods such as stepwise procedures, by following physiological principles or by following shrinkage rules (Guisan and Zimmermann, 2000). Guisan and Zimmermann suggested that the number of explanatory variables should not exceed 10. A priori decisions about which parameters to include in the model have been found to be useful in studies of rare species, with small numbers of "present" data points (Gibson et al., 2004). For large datasets, it has been considered more appropriate to allow the statistical results determine the choice of datasets to be incorporated, using stepwise procedures (sensu Glenz et al., 1999; Debinski et al., 1999).

2.3.4 Testing validity

The most common method for assessing the performance of a statistical model has been to report the overall percentage of correct predictions (Sieg and Becker, 1990; Raphael et al., 1995; Debinski et al., 1999). Some researchers have tested model performance by cross-validation (Sieg and Becker, 1990). However, it has been argued that cross-validation provides an overly optimistic assessment of the predictive power of the model (Fielding and Bell, 1997; Guisan and Zimmermann, 2000).

Blackard and Dean (1999) and Guisan and Zimmermann (2000) argued that the optimal method for validating a model is to test it on independently collected data. This approach has only been possible in situations where independent datasets were available (Lauver et al., 2002). Guisan and Zimmermann suggested that where a single, large dataset is available, splitting that set into training and testing groups is appropriate. This approach was taken by Blackard and Dean (1999) and Manel et al. (2000). Fielding and Bell suggested that splitting a dataset into training and test sets can cause problems if the original data set is small. This has been a particular problem for studies of rare or cryptic species, with few positive records (e.g. Gibson et al., 2004).

2.4 Snow accumulation modelling

2.4.1 Snow transport mechanisms

Snow drift is an important element in mass-balancing processes in polar and alpine environments because of its potential to move large volumes of snow during strong surface wind events (Kind, 1986; Greene et al., 1999; Bintanja et al., 2001; Doorschot et al., 2004). Snow enters the environment as precipitation, which is not distributed uniformly across the landscape when it occurs in windy conditions (Kind, 1986). Precipitation tends to accumulate on leeward slopes, but this process is poorly understood and rarely incorporated in models of snow accumulation (Lehning et al., 2000).

It is generally understood that snow is ablated from windward surfaces, and deposited in low-velocity zones, such as in the lee of topographic features (Evans et al., 1989; Ishikawa and Sawagaki, 2001; Haehnel et al., 2001). Wind speeds increase on windward and convex slopes, and decrease on leeward and concave slopes (Liston and Sturm, 1998). Exposed areas, such as ridges, may be almost snow-free, while gentle slopes may show a uniform snow distribution (Jaedicke et al., 2000). Deep snowdrifts occur where the areas of ablation are larger than the areas of accumulation (Liston and Sturm, 1998). It has been argued that topography is a crucial driver of snow accumulation patterns because changes in wind direction are more important than changes in wind speed in determining snow deposition (Liston and Sturm, 1998). It has also been found that the heaviest snow drift events occur during periods of roughly constant strong winds, and that short but strong blasts do not produce significant snow drift (Michaux et al., 2002).

The dominant snow transport methods are saltation and turbulent suspension (Kind, 1986; Liston and Sturm, 1998). Creep (the gradual down-slope movement of crystals) also shifts snow, but it is rarely incorporated in flux modelling because of its very small contribution to total flux (Uematsu et al., 1991; Jaedicke et al., 2000). Saltation is the motion of snow crystals bouncing in a flow-layer several centimetres above the snow layer (Kind, 1976; Kind, 1986; Pomeroy et al., 1997; Lehning et al., 2000; Whittow, 2000). Saltation occurs when the wind produces a shear stress that exceeds the amount of stress needed to shatter the bonds of snow surface crystals. It is considered to be responsible for up to 25% of annual snow transport; with the proportion dropping as wind speed increases (Pomeroy et al., 1997; Haehnel et al. 2001). Once snow is moving by saltation, it is available to become suspended in the zone of turbulent flow (Kind, 1976; Kind, 1986; Pomeroy et al., 1997; Lehning et al., 2000). The term turbulent flow describes the net forward movement of air in an irregular, eddying flow, and it stretches tens of metres above the snow surface (Whittow, 2000; Pomeroy et al., 1997).

In addition to snow being redistributed, a significant amount is lost through sublimation. This is the conversion of ice crystals to vapour (Pomeroy et al., 1997). In one study in the Arctic, sublimation losses were calculated at 9-22% of the winter precipitation, with sublimation accounting for up to half the winter precipitation falling on windward slopes (Liston and Sturm, 1998).

2.4.2 Why snow accumulation is modelled

Accurate models of snow accumulation are needed for many applications. These include water-catchment management (Greene et al., 1999; Daly et al., 2000; Walter et al., 2004), avalanche threat abatement (Greene et al., 1999; Lehning et al., 2000; Doorschot et al., 2004), infrastructure planning and maintenance (Kind, 1976; Purves et al., 1998; Jaedicke et al., 2000; Haehnel et al., 2001), ecological studies (Evans et al., 1989, Greene et al., 1999; Ishigawa and Sawagaki, 2001) and management of recreational sites such as rock climbing areas and ski runs (Purves et al., 1998).

2.4.3 Available snow accumulation models

Snow accumulation models vary in both complexity and accuracy, depending on the available input data and required results. Most of the early models operated in two dimensions (Greene et al., 1999), and were designed to predict the evolution of snowdrifts and the distribution of snow along the line of the prevailing wind direction (Liston and Sturm, 1998). Attempts to take into account the three-dimensional spatial variability of snow accumulation patterns are more recent (Daly et al., 2000). One method of classifying models is into those that attempt to provide numerical snow depths (e.g. Pomeroy et al., 1997; Greene et al., 1999; Lehning et al., 2000); and those that provide a relative map of snow distribution (e.g. Purves et al., 1998; Ishigawa and Sawagaki, 2001).

Attempts to model snow accumulation are complicated by the challenges involved in measuring precipitation and snow flux. These challenges include disturbing influences such as low temperatures, high humidity, riming and the difficulty of distinguishing fresh precipitation from drifting snow (Doorschot et al., 2004). Attempts to measure snow flux have used pulse-counting sensors, mechanical traps, acoustic and optic sensors (Bintanja et al., 2001; Doorschot et al., 2004). These measurement difficulties are reflected in the fact that the Australian Bureau of Meteorology does not record precipitation at its Antarctic weather stations (AADC, 2006).

A criticism of many snow accumulation models, such as those described in Pomeroy et al. (1997) and Lehning et al. (2000) has been that they are too complicated for easy assimilation into models such as those used for hydrological management (Walter et al., 2004). There have been attempts in recent years to develop relatively simple snow-distribution models, because of concern that the earlier, mechanistic models were overly complex (Walter et al., 2004). These simpler models typically set a constant value for variables such as snow density and threshold shear velocity, despite the spatial variability of these values (Liston and Sturm, 1998).

Another problem with these physically detailed models is their intensive input requirements. Many of the available models require automatic weather stations to be distributed through the study area (Jaedicke et al., 2000; Daly et al., 2000; Lehning et al., 2004). These requirements are greatest for models designed to cover large expanses of spatially variable terrain, across which the major input parameters (typically solar radiation, wind speed, air temperature, humidity and precipitation) are likely to vary significantly (Daly et al., 2000).

In addition to the meteorological inputs, models typically require an initial snow cover layer. Researchers have used a variety of methods to assess the initial snow cover. These include using radar to measure snow depth (Jaedicke et al., 2000) and manually measuring snow depths along transects (Evans et al., 1989). Such approaches are inappropriate for modelling historic snow cover, or for modelling snow cover in areas that are not readily accessible.

Some of the most prominent snow accumulation models are briefly described here. The Prairie Blowing Snow Model (Pomeroy et al., 1993) was physically based and was found to predict snow accumulation to within 10% of observed snow depths. However, its complexity and data requirements meant that its use was restricted to areas with major climatological stations, and is hence unsuitable for polar environments (Pomeroy et al., 1997). The Distributed Blowing Snow Model was developed for Arctic conditions from the Prairie Blowing Snow Model. This model divided the terrain into homogeneous landscape elements, based on vegetation, terrain, exposure and fetch characteristics. Meteorological observations were used to model the snow transportation processes (Pomeroy et al., 1997). Such an approach ignores the continuous nature of landscape variables, and is considered to be not suitable for adaptation to fine-scale applications such as this study.

Another model, SnowPack (Lehning et al., 2000; Doorschot et al., 2004) was designed for avalanche prediction. Its focus is on modelling the stability of the snow-pack and it therefore required input data related to crystal structure of the snow pack. The model was designed for steep alpine environments, and required input from a network of about 100 weather stations in the Swiss Alps (Lehning et al., 2004). In steep terrain, there is also the potential to use remotely triggered cameras to measure changes in snow distribution, which can then be used as inputs for the development of a statistical model (e.g. Tappeiner et al., 2001).

Daly et al. (2000) used spatially explicit temperature-index, precipitation and snow maps to specify initial snow conditions, before running their model through multiple time steps. Although the authors described this model as "simple" (p. 3269) it relied on 97 air temperature sensors and 287 snow gauges reporting hourly measurements.

The SnowTran-3D model (Liston and Sturm, 1998; Greene et al., 1999) attempted to model three-dimensional snow movement for treeless terrain in the Alaskan Arctic. It was designed to help with water resource management and has been adapted for integration with a GIS (Haehnel et al., 2001). This model accounted for transport variations resulting from accelerating and decelerating flow, using solar radiation, precipitation, wind speed and direction, air temperature, humidity, topography and vegetation snow-holding capacity. It has been argued that SnowTran is only suitable for areas with gentle terrain (Lehning et al., 2000) because it assumes that the wind direction is not affected by the topography.

Scale

One of the principle constraints on snow accumulation modelling is its demand for computational power. Thus, attempts at snow modelling are typically limited in spatial and temporal scale scales (Hägeli and McClung, 2000). Problems associated with scale are common with snow accumulation models. These include the inability of regular meteorological measurements to capture fine-scale processes such as snow drift; the inability of snow profile measurements to capture spatial variability and contradictions between input and output scales.

Temporal and spatial scales can be defined as a combination of extent and resolution (Greenberg et al., 2002).). The temporal extent has typically been one season. Models designed to cover large extents have had commensurately large resolutions – up to $4km^2$ and time-steps up to one month (Evans et al., 1989; Daly et al., 2000; Orndorff and Van Hoesen, 2001). Studies with smaller extents have generally also had finer resolutions – e.g. Greene et al. (1999) used cells of 30x30m, and some models have operated at time-steps as small as one to three hour time-steps (Daly et al., 2000; Haehnel et al., 2001). Liston and Sturm (1998) wrote that more detailed DEMs can resolve more landscape features, finer temporal scales for weather data can better capture short-term weather events and more detailed wind models can improve the accuracy of the model. However, the trade-off is in greater computational complexity. Naaim et al. (1998) wrote that numerical modelling of wind fields over complex terrain was computationally intensive and warned that boundaries must be carefully chosen.

2.4.4 Commonly used physical inputs

Wind Flow Field

Topography alters the speed and direction of the wind flow (Liston and Sturm, 1998; Lehning et al., 2000). Therefore, a robust wind field model has been considered to be an important part of many snow accumulation models (Purves et al., 1998; Liston and Sturm, 1998; Haehnel et al., 2001; Walter et al., 2005). Most of the available algorithms to measure this are based on the slope and aspect of cells. In contrast, Liston and Sturm (1998) elected not to model a wind field because their model was designed for relatively gentle terrain, and excluding a wind field model reduced computational expense.

Wind Shear Stress

The wind speed required to shift snow is dependent on factors such as the age and density of the snow and the temperature and humidity of the air (Doorschot et al., 2004). However, some snowdrift models have used a constant value for this threshold (Purves et al., 1998; Greene et al., 1999). Once the wind shear stress exceeds the threshold, snow begins to move. Modelled wind shear velocity has generally been calculated from the wind flow field and the surface roughness (Liston and Sturm, 1998; Haehnel et al., 2001) and may also take into account the air and snow density (Jaedicke et al., 2000). Liston and Sturm (1998) argued that the threshold changes very slowly in low temperatures, and that a constant value is suitable for environments such as the winter Arctic. They found that attempts to introduce more complex and realistic shear thresholds did not improve the results of their model.

A few models have included no measure of wind shear (Tappeiner et al., 2001; Purves et al., 1998; Orndorff and Van Hoesen, 2001). Tappeiner et al (2001) and Orndorff and Van Hoesen (2001) attempted to explain snow accumulation based on terrain characteristics rather than on weather inputs.

Saltation and Suspension

Most models have taken into account at least a single measure of snow transport through saltation and suspension (e.g. Liston and Sturm, 1998; Haehnel et al., 2001; Walter et al., 2004). Many of the authors used similar algorithms to calculate these factors, and the treatment outlined by Liston and

Sturm (1998) has been followed by others (e.g. Haehnel et al., 2001; Walter et al., 2004; Parajka et al., 2005).

The availability of snow for drifting is largely density dependent. Some models operated on the assumption that old snow is too dense to be easily transported, and that only fresh snow is available for transport (Walter et al., 2004). Purves et al. (1998) argued that during melting periods, no drift would occur, apart from that of precipitating snow. They argued that if a melt-freeze cycle occurs without snow falling during the freeze period; it could be assumed the shear velocity would be so high as to prevent any erosion.

Sublimation

Sublimation is the evaporation of snow to water vapour (Whittow, 2000). The relative importance of sublimation in snow dynamics is driven by factors such as temperature and wind speed, with the sublimation rate increasing as temperatures and wind speeds rise (Pomeroy et al., 1997; Liston and Sturm, 1998). Sublimation calculations require information on the size of the snow crystals (Pomeroy et al., 1993).

Precipitation

Precipitation was a key input for most snow accumulation models (Purves et al., 1998; Lehning et al., 1999; Daly et al., 2000; Orndorff and Van Hoesen, 2001; Tappeiner et al., 2001). However, in polar environments, precipitation may be a less important source of snow than accumulation by horizontal transport (Liston and Sturm, 1998; Seppelt and Connell, 2005).

Initial snow layer

Various methods have been used to produce the initial snow cover inputs for snow models. Purves et al. (1998) initialised their model with a uniform layer of snow over the entire study area. Other approaches have included measurements the snow depth along transects either manually or by using radar (Jaedicke et al., 2000). Ishikawa and Sawagaki (2001) used a fine scale (2m) DEM, and used a pit-filling algorithm to simulate concave areas filling with snow.

3 Chapter 3: Data and Methods

3.1 3.1 Study areas

3.1.1 Windmill Is

The Windmill Is are the islands and coastline covering an area of about 75-80km^2 around Casey (66°17'S, 110°32'E) Wilkes Land, East Antarctica (Fig. 3.1). They comprise four large peninsulas and more than 30 islands (Murray and Luders, 1990; Kirkup et al., 2002). During summer, the Windmill Is contain the only extensive areas of snow-free land for about 800km of coast around Casey (Murray and Luders, 1990; Kent et al., 1998).

The Windmill Is contain extant penguin colonies on fourteen islands and peninsulas. The region's total population was estimated at 93 092 ±9300 pairs in 1990 (Woehler et al., 1991). Historically, colonies have also existed in other parts of the region, such as the Bailey Peninsula (Emslie and Woehler, 2005). Colonies have been monitored on Shirley I, Whitney Pt, Blakeney Pt, the Frazier Is, Odbert I, Ardery I and Peterson I during the period of human habitation.

3.1.2 Geology

The geology of the northern Windmill Is is dominated by metamorphic rocks, in particular schist, gneiss and migmatite (Orton, 1963; Murray and Luders, 1990; Kirkup et al., 2002). At some point during the late Pleistocene-early Holocene, the entire Windmill Is area was glaciated, and has subsequently been subject to fluctuating sea levels (Kirkup et al., 2002). Shirley Is and Whitney Pt are topographically characterised by gentle, rocky outcrops, with a maximum elevation of approximately 30m MSL. The Windmill Is are located to the west of Law Dome, but the inland topography deflects the katabatic winds away from the immediate area (Murray and Luders, 1990).

Fig. 3.1: Location of the study sites.

3.1.3 Weather

The weather around Casey is frigid-Antarctic (Melick et al., 1994). Weather observations between 1989 and 2004 showed that in the warmest month, January, the mean daily maximum temperature was 2.1°C and the mean minimum was -2.6°C. October was the coldest month of the breeding season and also the time when Adélie penguins arrived at the colonies. During October, the mean daily temperature range was -15.3°C to -8.3°C. The area had a modelled mean annual snowfall of 224.6mm (snow water equivalent) (Bureau of Meteorology, 2004). Between 1996 and 2006, monthly mean wind directions ranged between 92.3° and 186.1°, with the prevailing winds coming from ESE. During the Adélie penguin breeding seasons in this decade, the mean wind speed was 12.64 knots, with a mean monthly maximum wind gust of 65.06 knots. During that decade, breeding seasons had a mean 32.5 days in which winds exceeded gale force (37kts) (AADC, 2006).

3.1.4 Human history

Humans have visited and occupied the Windmill Is since the USA Navy's *Operation Windmill* in 1947-48. The USA established the Wilkes research station (66°15.4'S, 110°31.5'E) in 1957. The Wilkes base was handed over to the Australian government in 1960. The Australians inhabited Wilkes until 1969, when they shifted to the Casey Tunnel (66°16.7'S, 110°31.5E), which is located between the current Casey site (66°15.9'S, 110°31.8'E) and the coast. In 1989, the station was shifted to its current site (Woehler et al., 1991; Bureau of Meteorology, 2006). Personnel in both the Australian and American programs undertook scientific research at both Shirley I and Whitney Pt (e.g. Penney, 1968; Kent et al., 1998; Woehler et. al., 1994). In 2005/06, Casey housed 53-60 personnel during summer, and 20 during winter.

3.1.5 Shirley I

Shirley I (66°17'S, 110°29'E) lies about 750m west of Casey, across a 100m-wide channel that is blocked with sea-ice for part of the year. In the 2005/06 summer, the island contained 46 extant Adélie penguin colonies and 22 relic colonies (E.J. Woehler, unpub. data). When Adélie penguins were first counted there in 1968/69, the island was inhabited by approximately 7100 breeding pairs (Woehler et al., 1994). By 2005/06, this had increased to about 11, 000 breeding pairs, an increase of 54.9% (E.J. Woehler, unpub. data).

During the 2005/06 breeding season, nine pairs of south polar skuas *(C. maccormicki)* were breeding on the island (P.K. Bricher, unpub. data), along with an unknown number of Wilson's storm-petrels *(Oceanites oceanicus)* and snow petrels *(Pagodroma nivea)*. The island is frequently visited by Weddell seals *(Leptonychotes weddellii)* and occasionally by male southern elephant seals *(Mirounga leonina)*. Leopard seals *(Hydrurga leptonyx)* patrol the seas around the island, especially during February when the fledged penguin chicks depart for sea. Station personnel have regularly visited the island since Wilkes was established. It is within Casey's extended station limits when the sea ice in the Shirley Channel is safe to cross on foot. This makes it is a popular destination for station personnel during early summer (Woehler et al., 1994; E.J. Woehler, unpub. data). Once the sea-ice has broken out, it is less frequently visited by groups using boats.

In 1963/64, the Adélie penguin population of the island was estimated at 3000 breeding pairs. Formal counts on the island began in 1968 and were conducted on five occasions between 1968 and 1977. The counts lapsed until 1989, apart from a partial count in 1984. Between 1989 and 2005, counts were conducted in 13 years. Most of the studies of penguins and other species on Shirley I have been non-invasive, and relied on observations and samples collected from outside the colonies (e.g. Woehler et al., 1991; Woehler et al., 1994; Petz, 1997; McRae et al., 1999; Emslie and Woehler, 2005). However there are some exceptions to this. In 1968, 140 adult Adélie penguins were banded (Murray and Luders, 1990). In 1992, researchers investigating Adélie penguin diet marked 46 breeding adult pairs, and temporarily banded the flippers of their chicks. The stomachs of 52 adult birds were flushed using the water-flushing technique, and 92 chicks were flipper-banded (Robertson et al., 1994; Kent et al., 1998). A total of 26 birds were fitted with satellite trackers in two studies during the summers of 1995/96 and 1996/97 (Kerry et al., 1997; Wienecke et al., 2000).

3.1.6 Whitney Pt

Whitney Pt (66° 15'S, 110° 32'E) is one of two mainland Adélie penguin breeding localities in the Windmill Islands. It is part of the Clark Peninsula that was designated as a Site of Special Scientific Interest in 1985. In 1996, the Clark Peninsula SSSI was redesignated Antarctic Specially Protected Area No. 136 (AAD, 2006). The site was considered to be of particular value because it is largely undisturbed and supports one of Antarctica's most extensive and best-developed plant communities. The Adélie penguin population in the ASPA was considered to be significant and relatively undisturbed, and so was listed among the values to be protected. The Australian Antarctic Division (2006) stated that these populations provide valuable comparative data for human impacts at Shirley I. Under the rules of the ASPA, access to the site is restricted by a permit system. Permits are only issued for scientific research or for essential management purposes consistent with the site's management plan. Typically, permits to conduct seabird surveys are only issued for two people at a time to enter the ASPA.

In the summer of 2005/06, Whitney Pt was occupied by 43 colonies of Adélie penguins, and contained a further 4 relic colonies (E.J. Woehler, unpub. data). It is also inhabited by breeding South polar skuas (10 pairs in 2005/06 (P.K. Bricher, unpub. data), at least three pairs of snow petrels and approximately 10-20 pairs of Wilsons storm-petrels.

Adélie penguins were first counted at Whitney Pt in 1959/60, when the population comprised approximately 1100 breeding pairs in 14 colonies (Penney, 1968). By 1983/84, this had increased to 4199 pairs in 28 colonies. Five of the new colonies were located on relic sites as identified by Penney (1968). During this period, the eight colonies at the western end increased in population by 33% and four new colonies were established, bringing the total population increase in that area to 58%. In contrast, at the eastern end of the point, 11 new colonies were established and the population increased by 519% (Martin et al., 1990; Woehler et al., 1991). By the summer of 2005/06, the total breeding population at Whitney Pt was 8790 – an increase of 699% since 1959/60 (E.J. Woehler, unpub. data.).

Whitney Pt is approximately 500m from the "Wilkes Hilton" - the radio hut for Wilkes, and now a popular field hut for expeditioners at Casey. Since the SSSI declaration, access to Whitney Pt has been limited to scientists conducting Adélie penguin counts and botanical studies. Before that, it was open to visitation by station personnel. In the summers of 1959/60 and 1960/61, Penney (1968) lived in a small hut (known as "The Wannigan") near colony IV, observing penguin behaviour. In Penney's study, 1528 adults, 66 juveniles and 217 chicks were banded with aluminium flipper-bands, and 25 birds were dissected. Nest locations in colonies I-VI were marked with welding rods. In 1964, a further 100 chicks were banded (Murray and Luders, 1990).

3.2 Data sets

3.2.1 Adélie penguin counts

Adélie penguin count data were available for the colonies at Whitney Pt for 22 seasons between 1959/60 and 2005/06, and for the Shirley I colonies for 18 seasons between 1968/69 and 2005/06. The number of breeding pairs was generally counted between 25 November and 5 December each summer, when the females had laid their eggs and departed to sea, leaving the male to incubate the eggs. In a number of years, bad weather delayed the counts, and in these years the counts were conducted as soon after the planned dates as weather permitted.

Penguins were counted using manual tally counters, while standing outside the colonies. In most years, two people conducted the Whitney Pt count and three people conducted the Shirley I count. Each colony was counted up to six times, and a mean of those counts calculated. To avoid bias in the counting, repeat counts of an individual colony were not conducted consecutively.

3.2.2 Adélie penguin colony maps

In February 2006, the Adélie penguin colonies at Shirley I and Whitney Pt were mapped using a Trimble Pro XH differential GPS (P.K. Bricher, unpub. data). The extents of currently occupied (extant) and abandoned colonies were mapped. The current perimeter was determined by the extent of fresh guano. By February, the neatly defined pebble nests seen at the start of the breeding season were scattered, and hence could not be used to determine the colony extent. Guano-covered areas that were obviously pathways to and from the colonies were excluded. The historic perimeters were determined by the area covered by pebbles of a suitable size for nest building (Emslie and Woehler, 2005). This criterion meant that areas of bare rock, similar to those seen to be occupied by nesting birds in extant colonies, were excluded. Only those areas with clear evidence of past occupation were mapped (Fig. 3.2).

Observations on the ground suggested that the position of these perimeters could be reliably identified to approximately ±0.5 m. In many locations, the position was more precisely located, but an error margin of ±0.5 m was considered conservative and was used here. For Whitney Pt, the GPS had a mean horizontal accuracy of ±0.49 m. For Shirley I, the horizontal accuracy was ±0.69 m.

The mapping was conducted at a time selected to minimise disturbance to nesting birds. In late

February, the chicks had fledged and departed, and most adult birds were feeding at sea. Up to two birds were still within the boundaries of some colonies. If birds were sitting near the colony perimeter, mapping of that colony was delayed until the birds had moved away.

All mapping for this project was conducted in UTM WGS 84, zone 49S.

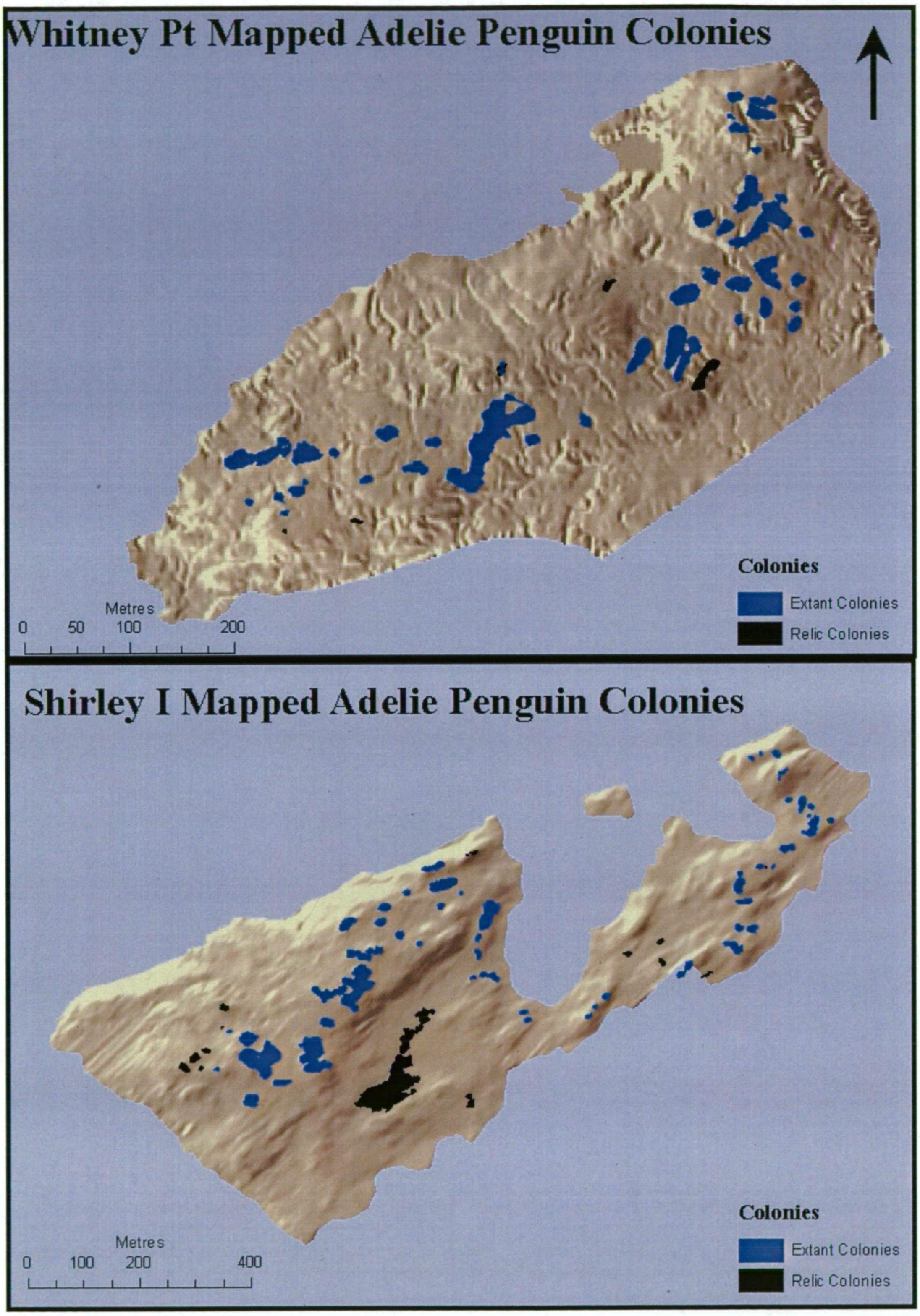


Fig. 3.2 Adélie penguin colony maps for Shirley I and Whitney Pt.

3.2.3 Ethics approval and permits

All fieldwork for this study was conducted as part of Australian Antarctic Science Project 1219. The chief investigator was Dr Eric Woehler. Permit number 05/06-1219 authorised the following:

(i) while on foot, disturb

(a) a concentration of birds; or

(b) a bird that is breeding or moulting;

(ii) enter an Antarctic Specially Protected Area

Ethics approval was granted by the University of Tasmania Animal Ethics Committee, under Animal Use Permit No A0008581.

3.2.4 Aerial photographs

The Australian Antarctic Division has acquired aerial photographs of Shirley I and Whitney Pt on a number of occasions. Photographs were taken of Whitney Pt in 1990, 1994 and 2003. The 2003 aerial photographs were used for the creation of a fine-scale Digital Elevation Model (DEM). The photographs were taken using a Zeiss UMK 1318 photogrammetric camera, from a height of approximately 828 m ASL, with a focal length of 100mm. These photographs showed Whitney Pt in a stereo-pair of photographs, but the photographs were taken shortly after a snowfall event that covered the entire site in a thin layer of snow (AADC, 2003). The snow cover limited the precision of the DEM that could be created because of the difficulty of identifying ground control points and in identifying surface heights (Fig. 3.3).

The 1990 photographs, taken with a Linhof photogrammetric camera, did not include stereo-pairs. They were hence inappropriate for extracting height data. These photographs showed areas of snow and ice in December 1990. The snow cover in these photographs corresponded with observations at Whitney Pt during December 2005 (P.K. Bricher, unpub. data). As 2005/06 was considered a high-melt summer, it was considered that these areas represented permanent snow and ice, and were therefore considered to be unsuitable for Adélie penguin habitat. The Linhof photographs covered all of Whitney Pt except for a small area, approximately 20 x 30 m in the northeast (AADC, 1990) Therefore, permanent snow cover could not be mapped for this area (Fig. 3.4).

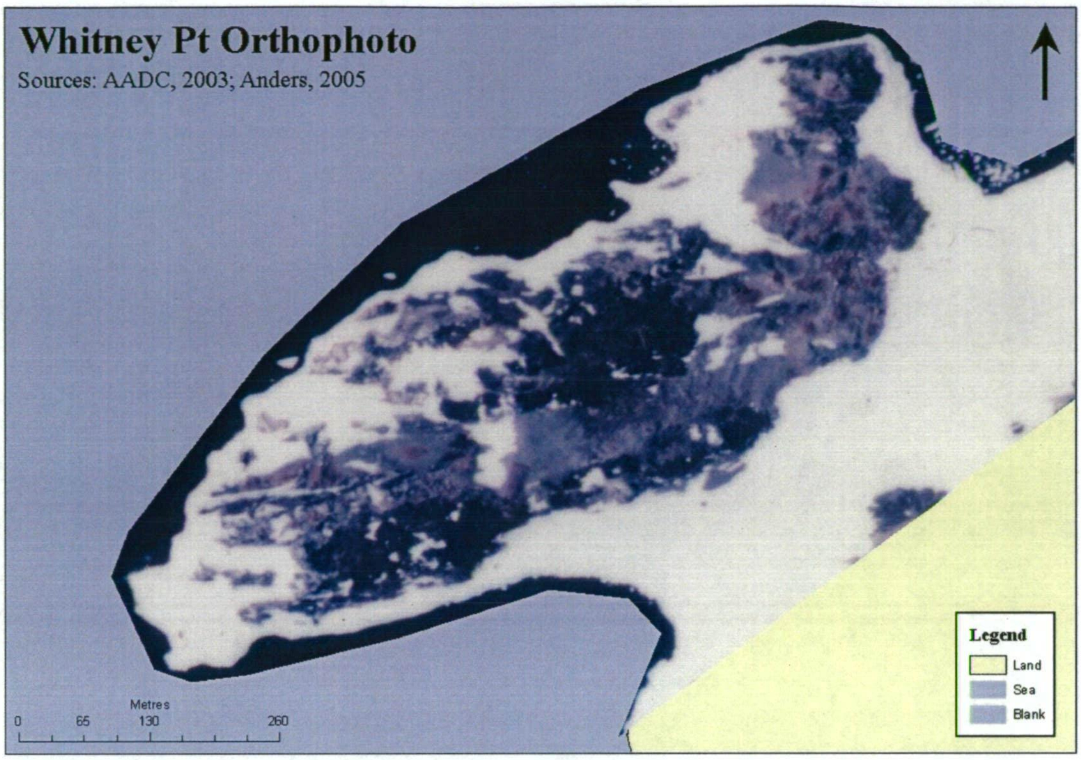


Fig. 3.3: Orthophoto of the aerial photographs of Whitney Pt, used for photogrammetry (AADC, 2003; Anders, 2005).

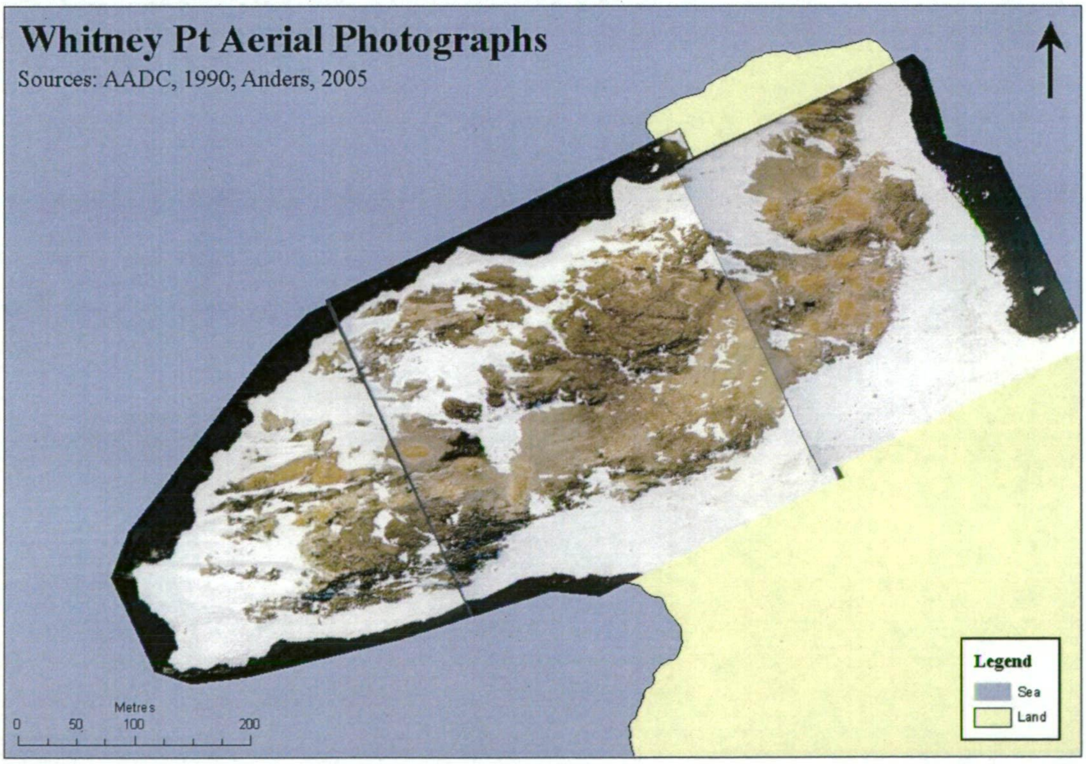


Fig. 3.4: Aerial photographs of Whitney Pt used for mapping permanent snow cover (AADC, 1990)

The 1994 photographs were taken with a Zeiss UMK 1318 photogrammetric camera. Two sets of photographs were taken at heights of approximately 500m ASL and 3000m ASL. The 500m altitude photographs were overexposed and covered in snow (AADC, 1994). They were thus considered unsuitable for DEM creation. The combination of altitude and snow-cover in the photographs taken from 3000m ASL meant they contained less precise height data than the 2003 photographs.

Aerial photographs were taken of Shirley I in 1994, 2001 and 2003. The photographs taken in January 2001 were used for the creation of a photogrammetric DEM with cells 4 m^2. They were taken with a Wild RC8 camera, at an approximate height of 750m and a focal length of 210mm. 10 stereo-pairs of photographs covered the whole island. Similar to the 1990 images of Whitney Pt, these photographs showed the areas of permanent snow cover that were considered unsuitable for Adélie penguin nest-sites (Fig. 3.5). The snow cover in these images corresponded with observations in January 2006 (P.K. Bricher, unpub. data). However, the relatively high number of photographs covering the island made the DEM construction complex.

The 1994 photographs were taken using the same camera and aircraft as those used for the 1994 Whitney Pt mission. These photographs had the same problems with snow and over-exposure as outlined above for Whitney Pt. The 2003 images were taken during the same mission as the 2003 photographs of Whitney Pt, and showed a light snow cover over the entire island. These images could be used to construct a DEM, but the 2001 photographs were considered to be more suitable because of the low altitude, long focal length and minimal snow cover.

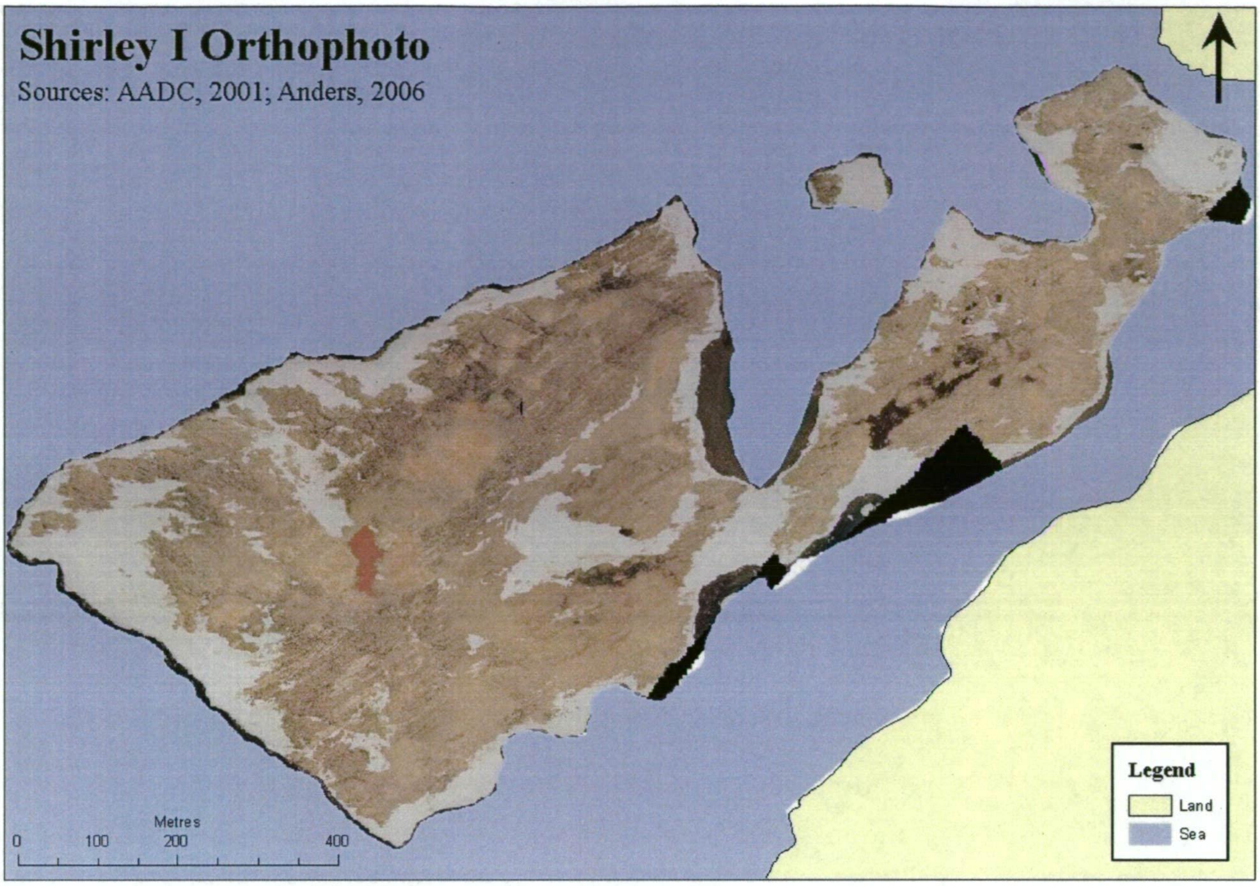


Fig. 3.5: Orthophoto of aerial photographs of Shirley I (AADC, 2001; Anders, 2006).

3.2.5 Digital elevation models

For the present study, high-resolution photogrammetric DEMs were extracted from stereo-pairs of aerial photographs of the study sites (Anders, 2005, 2006) using the software Virtuozo NT (www.supresoft.com). For Whitney Pt, a DEM with 10x10 m cells was derived from the 2003 aerial photographs (see section 3.2), with ground control collected using a differential GPS (Morgan, 2005). The GPS had a mean horizontal accuracy of ±0.8 m, and a mean vertical accuracy of ±0.76 m. There was an additional horizontal error associated with the identification of the ground control points estimated at ±5m. The large size of this error was caused by difficulty in identifying the ground control points due to the extensive snow cover in the aerial photographs. Positional errors were greatest in the areas covered by snow. For areas of exposed rock, the vertical accuracy was ±2m, and the DEM's overall positional accuracy was 6.59 m ±0.58 s.d. (Anders, 2005).

From this DEM, a Topogrid interpolation was used to generate a DEM with 2x2 m cells (Fig. 3.6: Lucieer, 2005a). Topogrid uses an iterative finite difference interpolation method, and is effectively a discretised thin-plate spline technique, where the roughness penalty has been modified to allow the fitted DEM to follow abrupt changes in terrain, like streams and ridges (ESRI, 1999). Spline interpolations are considered to be hydrographically sound, but one of their limitations is that errors cannot be quantified. Splines maintain small-scale features better than other interpolation methods, such as trend surfaces and weighted averages but there is concern that they may produce an unnaturally smooth surface (Burrough and McDonnell, 2000).

The Shirley I DEM was based on the 2001 photography (AAD 2001) using ground control collected with a Trimble Pro XH differential GPS. The vertical accuracy for areas not covered by snow, ice or Adélie penguin faeces was ±2 m (90% certainty) of its true value. The surface heights for such areas could not be accurately determined because non-textured surfaces cannot be viewed in stereo. These areas were corrected for significant gross errors using a linear interpolation algorithm based on surrounding elevation data. The errors in the faeces-covered Adélie penguin colonies are likely to be smaller than those in the snow-covered areas, because of the small relative size of penguin colonies compared to permanent snow coverage. However, as this study was examining the terrain properties of Adélie penguin colonies, errors in these areas were far more critical than errors in the areas covered by permanent snow and ice.

The aerial photographs (AADC, 2001) were of sufficiently fine resolution to allow the identification of well-defined objects less than 0.5 m in diameter. The ground control points had a mean horizontal accuracy of ±0.68 m and a mean vertical accuracy of ±1.33 m. The large number of stereo-models (10) needed to provide coverage of the island meant there were significant areas of overlap between the models. In these areas, the height values for an individual cell were calculated as the mean of the values for that cell in each of the stereo-models. While the resulting cell heights were still within specification for the DEM, they resulted in artefactual "smoothing" in some of the derived data layers. This was especially evident in the surface roughness and curvature layers (see Chapter 4, Figs. 4.5-4.8). These artefacts did not affect the distance data layers, which were not derived from the DEM, or the modelled snow accumulation layers, which were modelled using 5x5m cells. They appeared to have little effect on the aspect, slope, solar radiation and wind exposure layers. To minimise the effects of these artefacts, a DEM was interpolated from the spot height points of the original DEM using a sub-sample of one-third of the original data points. An ordinary kriging algorithm with anisotropy and a large neighbourhood was then applied to the data

(Lucieer, 2006). The resulting DEM produced much less obvious artefacts in the curvature and surface roughness layers. The 4 m^2 cell resolution was considered to be the minimum that could reliably be extracted from the available photography. It was not possible to determine what effect these residual artefacts and the ±2m height error had on the analyses (Fig. 3.6).

For the snow accumulation models, the fine-scale DEMs of Shirley I and Whitney Pt were resampled to 25 m^2 cells, and merged with a DEM of the Windmill Is, that had 100 m^2 cells (Lucieer, 2005b). This DEM was interpolated from an available DEM of the Windmill Is with 625 m^2 cells using Topogrid (AADC, 1999; Lucieer, 2005b). Coastal features on these DEMs had a claimed horizontal and vertical accuracy of ±1m. However, they were derived from aerial photographs taken in several different summers, and at different dates, with varying sea-ice extents. The aerial photographs of Shirley I taken in 2001 showed that the mapped coastline for the island was incorrect – in the photographs used for mapping Shirley I, the Shirley Channel was blocked with sea ice. The 2001, ice-free photographs showed that about half the southern coast of Shirley Is and parts of the coves on the northern coast were mapped incorrectly. It is likely that the mapping errors were concentrated around coastlines where sea ice obscured the coast.

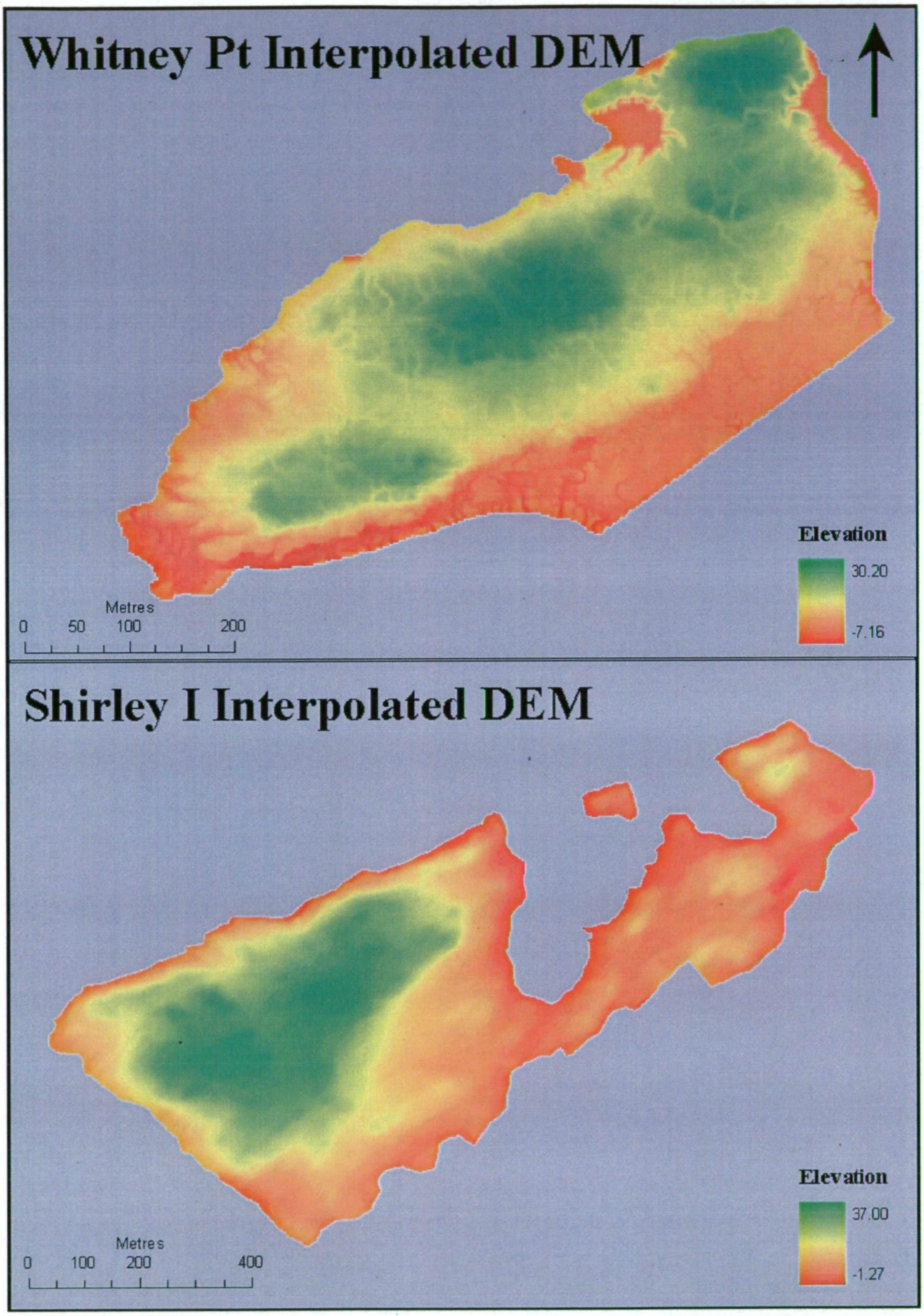


Fig. 3.6: Interpolated DEMs of Whitney Pt and Shirley I (Anders, 2005, 2006; Lucieer, 2005, 2006)

3.2.6 NCEP/NCAR Weather reanalysis data

This study investigated the effects of changes in modelled snow accumulation patterns on Adélie penguin colonies during the past fifty years. Bureau of Meteorology data were available for the Windmill Islands region from 1960 onwards, but it was considered that any changes in weather conditions were likely to have been obscured by changes in the locations of weather observing equipment (N. Adams, pers. comm.). From 1960 until February 1969, the weather station was located at the now-abandoned Wilkes (66 ° 15.3'S, 110° 31.5'E, 12m ASL) on the Clark Peninsula, near Whitney Pt. In February 1969, the weather station was moved to the Casey Tunnel (66°17'S, 110°32'E, 12m ASL) on Bailey Peninsula and in 1989, it was shifted to Casey (66° 16.9'S, 110° 31.8'E, 40m ASL: Bureau of Meteorology, 2006).

The NCEP/NCAR Reanalysis Project was the result of collaboration between the National Centers for Environmental Prediction and the National Center for Atmospheric Research. The project was set up to model global weather data from 1948 onwards, based on the available weather observations. It was designed to resolve problems such as that outlined above, and variations in the quality of weather observations. The project's authors claimed that the reanalysis data eliminated perceived climate jumps associated with operational data assimilation systems, though it was still affected by changes in the observation systems (Kalnay et al., 1996; Kistler et al., 2001).
The NCEP/NCAR reanalysis used observations of upper air temperature, horizontal wind and specific humidity; land surface observations of surface pressure and oceanic reports of surface pressure, temperature, horizontal winds and specific humidity. The values for surface variables, such as those used in the current study, were calculated from a combination of direct observations and the reanalysis model (Kistler et al., 2001).

Hines et al. (2000) noted that the quality of any of the available weather reanalysis outputs is reliant on the quantity and quality of the available data. The Southern Ocean and Antarctica have historically contained fewer weather reporting stations than the northern hemisphere, and Hines et al. (2000) argued that this makes such reanalysis data increasingly important but also increases the risk of errors because of the paucity of input data. In addition, they found that the extreme weather and often sharp topographic changes in Antarctica reduced the accuracy of the NCEP/NCAR results, compared with other parts of the world. They found that the data contained trends in surface pressure at 65°S that were not supported by available observations and warned that this could affect the results of studies involving changes in surface pressure. The spurious trend showed an approximate 0.20hPa yr^{-1} over 50 years. It is unclear whether this same trend in the data is evident at 62°S (the latitude of the current study sites) and what impact this trend may have had on the predictions of snow accumulation patterns, which are partly determined by surface pressure data.

The NCEP/NCAR reanalysis data for the Casey region was extracted for three key periods in 1959/60, 1968/69 and 2005/06 (Kalnay, 1996). The first two periods correspond to the summers in which the first Adélie penguin counts were conducted at Whitney Pt and Shirley I respectively, and 2005/06 with the most recent counts. The data were calculated for a grid-point located 63.14 km north-northeast of Casey (65.7125°S, 110.625°E).

3.3 Methods

3.3.1 GIS processing methods

This study used the GIS package ArcGIS 9.0 (ESRI, 1999-2004) to create, store, manage and display spatial data layers. A freeware GIS package, PCRaster v2 (van Deursen et al., 2006), was also used to generate DEM derivative. The derived layers were then exported back to ArcGIS for further analysis. PCRaster is a dynamic modelling system that is largely used for environmental modelling. Here it was used to develop static layers showing slope, drainage, solar radiation and surface curvature.

Slope

Slope layers were calculated from the DEMs, based on a 3x3 cell neighbourhood. The slope was calculated using a third-order finite difference method. The resulting value for the central cell was given as height difference (vertical distance / horizontal distance). Where a surrounding cell was missing a height value, or the centre cell was at the edge of the DEM, a neighbourhood interpolator was used to fill in the missing values. This interpolator assigned each unvalued cell the mean value of all existing cells in a 3x3 cell window surrounding the cell with the missing value (van Deursen et al., 2006).

Aspect

The PCRaster aspect function operates in a similar way to the slope algorithm. Each cell's aspect value was calculated based on the elevation of its 8 neighbours in a 3x3 cell kernel. The third-order finite difference method was used. The results were presented in 360 degrees, with north assigned 0/360 degrees. Where a cell in the neighbourhood was missing a value, the neighbourhood interpolator, described above, was used to fill in the missing values. The aspect data were only used in decision tree analysis, for which they did not need to be converted to a linear scale.

Wetness Index

In PCRaster, the "wetness index" function calculates the potential drainage of a cell, based on the upstream area. This is a unitless measure as it models the shape of the land rather than actual runoff. From the DEM, a local drain direction network was calculated, using an eight-point pour algorithm. This created a layer describing the direction water flows from each cell to its steepest downslope neighbour. Where a cell had two or more downslope neighbours of equal elevation, the drain direction was assigned randomly. Flat areas were dealt with by calculating flow directions from cells at the edge of the flat area. A repair function was used to remove pits and hence ensure that the drain direction network was hydrologically sound. The decision was made to remove pits from the local drain direction layer; because observations of the shape of the landscape at Shirley I and Whitney Pt suggested that any pits were most likely to be artefacts of the DEM, rather than reflecting that actual shape of the land. From the repaired local drain direction network, an upstream area value was calculated for each cell, showing the total area of all upstream cells. The wetness index was calculated by the following equation:

Wetness = ln (upstream area/slope)

(sensu Burrough and McDonnell, 2000; PCRaster, 2005).

Planar and Profile Curvature

Planar (also known as planform) and profile curvature were calculated for each cell, using the elevation of neighbouring cells in a 3x3 cell neighbourhood. Planar curvature is a unitless measure of the change in slope per distance in horizontal direction, in the direction of the slope (such as that shown by contour lines), where concave slopes (gullies) are negative and convex slopes are positive. Profile curvature is a measure of the shape of the profile of the slope. Positive values occur at sites where the steepness of the slope is increasing such as the tops of hills (convex slopes).

Negative values occur where the steepness of the slope is decreasing, such as at the base of hills (concave slopes).

Solar Radiation

The PCRaster-based solar radiation model, PotRad (van Dam, 2001), was used to calculate the solar radiation in MJ/m^2 for each grid cell, based on the DEM and the latitude of the study sites. The model worked in hourly time-steps, over a year. It did not take into account weather conditions, such as cloud cover, which would reduce the amount of solar radiation. It was thus a measure of maximum potential solar radiation, rather than actual radiation. The lack of weather input was not considered significant as cloud cover over each study site was likely be largely homogeneous and any variation in cloud cover over the study sites was likely to be random.

Surface Roughness

Using ArcGIS, surface roughness was measured by calculating the standard deviation of the elevation of cells in a 3x3 cell neighbourhood. Standard deviation was selected as the measure of roughness, as it was less sensitive to the effects of a single outlying value than other measures, such as range. As surface roughness is correlated with slope, a normalised surface roughness layer was calculated by dividing the surface roughness value of a cell by its slope value. Both the standard deviation and normalised values of surface roughness were incorporated in the statistical analyses.

Adjacency

A data layer was produced showing whether adjacent cells contained Adélie penguin nests. This was done by calculating the mean of a 3x3 cell neighbourhood from a binary penguin colony presence/absence layer, where cells that contained Adélie penguin nests were given a value of 1, and cells without colonies were given a value of zero. This layer was used to exclude the edge cells of colonies from the analysis of all the landscape parameters. Only cells with a mean value of 1 (presence) were incorporated in the analysis. This was done to ensure that the cells used in the analysis actually represented sites that contained Adélie penguins. The precision of the colony maps was ±1 m and excluding the edge cells from the analysis minimised the chance of incorrectly labelling cells as containing penguin nests.

Wind Exposure

The wind exposure layer was based on the prevailing winds for the breeding season as shown in the NCEP/NCAR weather reanalysis data. Wind roses were constructed based on the three-hourly averages for 15th October 1959 - 31st January 1960, 15th October 1968 - 31st January 1969 and 15th October 2005 - 31st January 2006, using the program Grapher 6 (Golden Software, 2005). These dates were selected to reflect the period between the approximate arrival dates for breeding birds and the time when chicks reach the crèche stage, and are capable of some limited movement to avoid serious effects from climatic conditions. The data showed that the majority of wind speeds greater than $10m^{-1s}$ came from east-southeast. It appeared that any changes in prevailing wind direction between the examined time periods were within the annual variability and were hence not considered to be significant (Neil Adams, pers. comm.)

A hillshade model was created from the DEMs, using the wind roses to set the prevailing wind direction. A hillshade model has traditionally been used to model light and shadow based on an illumination source at a set direction and azimuth (ESRI, 1999). In the present study it was used to model exposure to the wind, with the direction set at 122°, and the azimuth at 5°, to imitate wind travelling just above the ground's surface, following the approach of Patterson et al. (2003). This layer could be used as a surrogate for three things – exposure to prevailing winds, snow accumulation (sensu Patterson et al., 2003) and on Shirley I, for exposure to potential airborne pollutants from Casey. In this study, it was used in the analyses of the effect of snow accumulation on population trends of Adélie penguin colonies, as a surrogate for wind exposure, and in the effects of human activities on the Adélie penguin colonies on Shirley I, as a surrogate for exposure to potential emissions from Casey.

Snow Accumulation Model

This study used a GIS-based model to simulate the drifting of snow in the two study areas. The model was developed for Antarctic conditions, and thus accounted for the absence of precipitation data and limited access to the sites (Wallace, 2005). Without direct measures of precipitation or of the starting snow layer, it was impossible to derive numerical results, such as snow-depths. Instead, the model produced a map of relative snow accumulation, similar to the models of Ishikawa and Sawagaki (2001), Orndorff and van Hoesen (2001) and Purves et al. (1998).

Due to the physical complexity of the forces that drive snow transport, models of snow accumulation have generally involved a balance between accuracy and available resources, such as input data and computational power (e.g. Liston and Sturm, 1998) Therefore, factors which were considered less important for a particular application have typically been excluded in order to minimise computational expense (e.g. Walter et al., 2004; Tappeiner et al., 2001). The Windmill Islands model accounted for wind deflection, wind speed, wind shear stresses, saltation, suspension and snow density (Wallace, 2005). Saltation and suspension were treated as one variable, after the approach taken by Kind (1981). Commonly incorporated factors that have been excluded from this model include precipitation, sublimation, vegetation and snow melt.

Wallace's model used equations developed for snow accumulation models in other parts of the world, but adapted for the Windmill Islands.

Wind deflection and speed are crucial to accurate modeling of snow accumulation, and are affected by topography. Algorithms to model deflections in wind direction have typically been based on the slope and aspect of the ground (Purves et al., 1998). Wallace's model used the following wind direction equation, which was developed by Ryan (1977):

$$D= -0.255S \sin(2(A-\theta))$$

Where S= slope (%)

A= aspect of the slope (deg)

θ= initial wind direction (deg)

Wind speeds tend to increase on windward slopes, and decrease in the lee of topographic features, dropping to almost zero just below ridges (Walter et al., 2004). Wallace's model followed Liston and Sturm's (1998) weighting for wind speed based on the slope and curvature of the topography, using the following formula:

$$R = 1.0 + \gamma_s\mu_s + \gamma_c\mu_c$$

Where μ_s = Topographic slope, scaled to be within the range $-0.5<\mu<0.5$

μ_c = Topographic curvature, scaled to be within the range $-0.5<\mu<0.5$

$\gamma_s = 0.6$ = Positive constant to weight the effect of slope on wind speed

$\gamma_c = 0.4$ = Positive constant to weight the effect of curvature on wind speed

Snow transport only begins when the wind shear stress (U_*) exceeds the threshold velocity (U_{*t}) (Gugolj, 2005). The wind shear stress was calculated by the following formula (Liston and Sturm, 1998; Gugolj, 2005)

$$U_* = W_r \frac{K}{\ln(H_r/S_r)}$$

Where W_r = Wind speed at reference height (ms^{-1})

$K = 0.41$ = von Karman's constant

$H_r = 10$ = Reference height (m)

$S_r = 0.1$ = Surface roughness length (m)

H_r is the reference height at which the wind speed is measured. At Casey this height was unknown, but it was assumed that the measurements were taken at 10m (Wallace, 2005). The surface roughness length was set at a value suggested by Linacre and Geerts (1999) for areas with low-lying vegetation and few sharp valleys and peaks (Wallace, 2005).

Saltation and suspension were modelled as one variable, using the following algorithm which was developed by Kind (1981):

$$Q_s = \frac{\rho_a U_*}{g\rho_s}\left(0.25 + \frac{U_s}{3U_T}\right)\left[1 - \frac{U_T}{U_*}\right] \text{ (m}^3\text{s}^{-1}\text{/m perpendicular to the wind direction)}$$

Where U_* = Wind shear stress ms^{-1}

U_T = Threshold velocity ms^{-1}

$U_s = 0.75$ ms-1 = Terminal fall velocity of snow

P_a = Air density (kgm^{-3})

P_s = Snow density (kgm^{-3})

g = Gravitational acceleration (ms^{-2})

Air density was calculated by the following formula:

$$P_a = \frac{p}{R.(T)}$$

Where p = Air pressure (Pa)

$R = 287.05 \frac{J}{kg.K}$ = Gas constant

T = temperature (K)

Snow density is assumed to vary with elevation, and following Walter et al. (2004), this was assumed to be 80kgm-3 for snow at elevations above 10m ASL, and 140kgm^{-3} for elevations below 10m ASL. Snow density is known to vary with age, and modelling to account for this requires daily precipitation observations, which were not available for the Windmill Islands when the model was developed (Walter et al., 2004; AADC, 2006).

Gravitational acceleration was calculated according to the following formula (International Association of Geodesy, 2005):

$$g = g_e(1+\beta_1 \sin^2(\varnothing) - \beta_2 \sin^2(2\varnothing)) - 3.086 \times 10^{-6}\, H$$

Where Ø = Latitude of the point

H = Elevation above sea-level (m)

$g_e = 9.7803184\ ms^{-2}$ = Gravitational acceleration at the equator

$\beta_1 = 0.0053024\ ms^{-2}$ = constant

$\beta_2 = 0.0000059\ ms^{-2}$

As the model was developed for Antarctic conditions, Wallace considered that the generally low annual temperatures would make sublimation and snow melt of limited importance in determining snow accumulation patterns. This assumption is unlikely to hold during summer, when temperatures in the Windmill Islands regularly climb above 0°C, and both snowmelt and sublimation are likely to be significant factors. In this study, the model was run for periods in late winter and spring, when it was considered that Wallace's (2005) assumptions would apply, and so sublimation and melt were not incorporated in the model.

Vegetation in the Windmill Is consists of low-lying mosses and lichens (Seppelt and Connell, 2005). The vegetation's snow-holding capacity is little different to the rock that is typical of the region. It was considered unlikely that it would have a significant effect on the surface roughness, unlike the vegetation in other parts of the world where grasses, shrubs and trees have significantly altered snow transport patterns (e.g. Pomeroy et al., 1993; Pomeroy et al., 1997; Liston and Sturm, 1998; Daly et al., 2000; Evans et al., 1989; Haehnel et al., 2001; Tappeiner et al., 2001). It was therefore considered appropriate to use the constant value of 0.1 m for surface roughness.

Reliable precipitation data has rarely been collected in polar environments, due to the difficulty in determining the difference between fresh precipitation and blowing snow (Bureau of Meteorology, 2006). One study in the low Arctic measured "true" precipitation in a small glade well within an open forest (Pomeroy et al., 1997). Such natural windbreaks do not exist in the Antarctic. The Windmill Islands have been estimated to have an annual precipitation of 175mm, but this has not been recorded directly (Seppelt and Connell, 2005; Bureau of Meteorology, 2006). The relatively small amount of precipitation was considered to be of limited significance in predicting snow distribution. The NCEP/NCAR Weather Reanalysis Project does provide models of precipitation for Antarctica for the past fifty years (NCEP/NCAR, 2006). However, those data were not based on actual observations so it is difficult to know how accurate they are. In addition, they were not available to Wallace (2005) when he designed the model.

Wallace (2005) tested his model against aerial photographs of Shirley I, taken in January 2001 (AADC, 2001). He used a DEM with grid cells of 20x20m, and ran the model over three months leading up to the day of the photography. This test showed strong agreement with the observed snow distribution shown in the photographs. He suggested that increasing the resolution of the DEM would improve this match still further.

The model was applied to a merged DEM that covered the northern Windmill Islands, in order to provide a source for transported snow into the study areas. The 2x2m cell DEMs of Whitney Pt and

Shirley I were resampled to 5x5m cells using a cubic resampling algorithm. The Windmill Island 10x10m DEM was also resampled to 5x5m by cubic resampling. The DEMs were then merged to generate a DEM that provided surrounding land for snow to blow from. The DEMs were resampled to 5x5m cells in order to combine two models of different resolutions. The model was run for three months leading up to 15 November in 1959, 1968 and 2005. This date was chosen as it marks the peak laying period for Adélie penguins and hence the time when snow cover is likely to be most crucial. The model required daily average temperature, wind speed, wind direction and air pressure data, were derived from the NCEP/NCAR weather reanalysis project. The model produced a data layer showing the relative distribution of snow for that day.

Once the snow accumulation data layers had been calculated for the years of first and last Adélie penguin counts, the layers for 1959 and 1968 (first counts for Whitney Pt and Shirley I, respectively) were subtracted from the 2005 layer. This produced layers showing the changes in snow accumulation. Cells where the snow accumulation was the same in both years were given a value of zero. Positive values indicated increasing snow depth and negative values decreasing snow depth.

Permanent Snow and Ice

The snow accumulation model simulated the short-term distribution of snow. Areas of permanent ice and compacted snow were evident in aerial photographs taken of the study sites in mid-summer (AADC, 1990; 2001). These were removed from the analysis. This was done by digitising the extent of snow visible in aerial photographs of both study sites. The extent and depth of snow patches that last through summer varies between years. Ideally, this would be measured in multiple years, and some form of average, maximal and minimal measures calculated. However only one set of suitable images was available for each study site.

The 1990 aerial photographs of Whitney Pt were georeferenced to the photogrammetric DEM (which was based on a single stereo-pair of images), with an affine transformation. Positional discrepancies between the three photographs and the DEM were typically 1-2 m, with errors of up to 4 m in a small area in the northwest. From the transformed photographs, a vector layer was created showing those areas of permanent snow. To minimise the risk of incorrectly describing exposed terrain as snow-covered, a conservative approach was taken to the digitising process. The vector layer showing permanent snow cover was digitised 1 m inside the extent of the snow visible in the photographs. This ensured that sites that represent actual Adélie penguin colonies were not incorrectly described as unsuitable. This was especially important in areas like colonies I -IV on Whitney Pt, which are situated near the base of a cliff, with a narrow bank of snow (about 1 m wide) between the colonies and the cliff.

For Shirley I, an orthophoto generated from the 2001 photographs was available (Anders, 2006; AADC, 2001). From this, Anders (2006a) generated a vector layer showing the areas of permanent ice. As the orthophoto was directly generated from the DEM, positional errors were limited, and the permanent snow layer was more accurate than that for Whitney Pt.

Assigning Population Trends

Each penguin colony was assigned a population trend. The count data for each colony was converted into percentages, with the baseline year assigned the value of 100%. Colonies in which the population increased after the baseline year were given values over 100%, while those that decreased were given values below 100%. The baseline year was typically 1959/60 for Whitney Pt and 1968/69 for Shirley I. For colonies that were established after those years, the baseline year was the first year in which penguins were counted at that site. The trend was based on the percentage difference between the year of first count and 2005. This method eliminated problems caused by the wide disparity in population size between colonies. For example, a population decrease of 10 breeding pairs in a colony of forty has far greater significance on the colony as a whole, than an identically sized decrease in a colony of 2000. This method reflects the compounding effect of population decrease in a small colony, as seen in studies which found that small colonies are less resilient to environmental stressors than larger colonies (Giese, 1996; Patterson et al., 2003).

Assigning percentage values was problematic in newly founded colonies. Such colonies exist at both study sites, and typically began with a few pairs of birds, before increasing. The 100% value reflected a small count (e.g. 10 birds), and the percentage increases were up to 27 400% (Colony L, Shirley I) in 15 years. The population trends were divided into classes (Table 3.1) to eliminate this problem. Decreasing colonies were split into moderate (50-80%) and strong (<50%) decrease classes. Colonies in which the population of 2005 were within 20% of the baseline count were considered to be stable. Those in which the populations were increasing were divided into moderate (120-150%) and strong (<150%) increases. This classification resulted in a small sample size for those that are decreasing moderately, but meant that each class was internally homogeneous, incorporating only colines that exhibited similar trends. Those colonies that have not been inhabited during the period of human occupation in the Windmill Islands were classed as relic. They were included in the analysis of variables which are unlikely to have changed in the past 9000 years, such as elevation and slope, but were excluded from the analyses of population trends.

Table 3.1: Population trend classes for Adélie penguin colonies at Whitney Pt and Shirley I

Population Trend	***Class***	***Number of Colonies***	
		Whitney Pt	**Shirley I**
<50% (Strong Decrease)	1	2	15
50-80% (Moderate Decrease)	2	1	8
80-120% (Stable)	3	5	9
120-150% (Moderate Increase	4	2	2
>150% (Strong Increase)	5	23	10
Total		33	44

Selection of random control plots

Control cells were selected randomly from those areas of exposed rock that had no evidence of having historically contained Adélie penguin colonies. Hawth's random sampling tool (www.spatialecology.com) was used to generate control plots. The tool was used to select random points – 42 points at Whitney Pt and 66 points on Shirley I. On these points, square control plots were generated which contained a number of cells equivalent to the mean number of cells in a colony at that study sites. At Whitney Pt, this number was 25 cells, and at Shirley I it was 49. Some of the control plots overlapped with colonies or areas of permanent snow. These were then shifted the minimum distance required to ensure that the plots represented areas of exposed rock without Adélie penguin colonies (Fig. 3.7).

Export of colony and control data from ArcGIS

The vector layer showing the extent of Adèlie penguin colonies was converted to raster, and this was used to set the analysis extent for the extraction of data from all the raster data layers. A conversion tool (Lucieer, 2005c) was used to convert the cells within the analysis extent to a comma separated file for import to statistical programs. The process was repeated for the control cells.

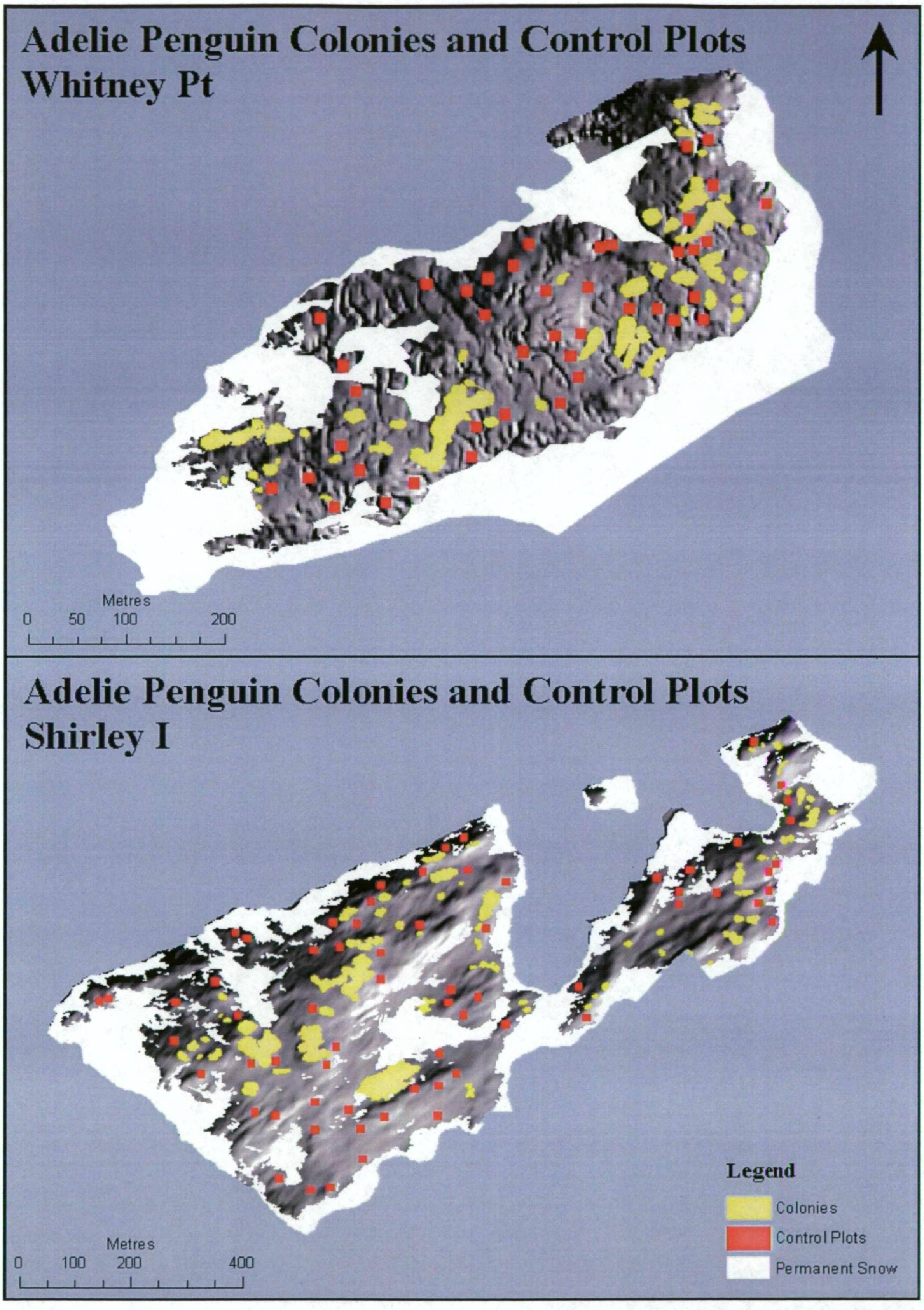


Fig. 3.7: Randomly selected control plots for Shirley I and Whitney Pt.

3.3.2 Statistical Processing Methods

Adélie penguin distribution univariate analysis

The statistical package, JMP 5.1 (SAS Institute, 1989-2003), was used to generate histograms and scatterplots for all the landscape parameters. Visual exploration of histograms of the parameters showed that none was normally distributed. Thus, parametric tests for differences were considered to be unsuitable (Zar, 1999). Instead, Wilcoxon two-group tests for difference were used to explore whether the distribution of colony values was different to the distribution of the control plot values. When applied to two groups of data, the Wilcoxon two-group test is also known as the Mann-Whitney Test (Zar, 1999; Dytham, 2003). The Wilcoxon/Mann-Whitney test has been found to have 95% of the power of the equivalent parametric t-test, when applied to a normally distributed dataset. However, if the data are not normally distributed, as in this study, the Wilcoxon test may be much more powerful than the t-test (Zar, 1999). The Wilcoxon test compares ranks rather than raw values (Dytham, 2003).

In this study, the two groups were Adélie penguin colonies, and control plots. The Wilcoxon test was applied to each of the static landscape parameters being investigated for its ability to explain the presence or absence of Adélie penguin colonies. These parameters were the modelled values for elevation, slope, solar radiation, wetness index, wind exposure, snow accumulation (2005), planar and profile curvature and surface roughness (standard deviation and normalised). The data for Whitney Pt and Shirley I were examined separately.

To minimise the effect of spatial autocorrelation, the tests were conducted on both the individual cell values and the colony/control plot mean values. Spatial autocorrelation can be defined by Tobler's Law, which states "*Everything is related to everything else, but near things are more related than distant things*" (Tobler, 1970). Thus, the values for a single parameter for two cells within a colony or a control plot are likely to be more similar than values for that parameter in two different colonies or control plots. Spatial autocorrelation can either take the form of patches or gradients (Legendre, 1993). In this study, the colonies and the control plots represented clumps, while the landscape data layers represented gradients. Legendre (1993) noted that this phenomenon can cause problems for statistical tests because the data violate the assumption of independence of observations that underlies most statistical techniques. Using colony/control plot mean values in this study allowed the removal of within-colony spatial autocorrelation. However, this approach also meant that much of the variance within colonies and control plots was lost. Therefore, it was considered appropriate to use both the individual cell values and the mean values, despite the limitations of each approach. Between-colony spatial autocorrelation is also likely to be present, but accounting for it was beyond the scope of this study.

Multivariate analysis of Adélie penguin distribution

Univariate tests, such as the Wilcoxon two-group test, can only measure the relationship between the response variable and one parameter at a time. This leads to a number of limitations: first, univariate tests cannot account for interactions between two or more parameters which may affect habitat suitability; second, it is possible that variables that show no mean difference may contribute to multivariate group separation; and third, test statistics generated for individual parameters cannot account for correlations between the parameters (Flury and Riedwyl, 1988). Therefore, in this study, multivariate statistical tests were needed to investigate interactions between parameters that may affect habitat suitability. As noted in section 2.3.2, a wide range of multivariate statistical models is available for predicting species distribution, based on GIS-derived parameters (Guisan and Zimmermann, 2000). Here, discriminant analyses and decision trees were used. The data for Whitney Pt and Shirley I were examined separately, and each test was conducted on both the values for individual cells and on the colony mean values.

Discriminant analysis

Discriminant analysis has often been used in bird ecology studies (e.g. Fraser and Patterson, 1997; Debinski et al., 1999; Manel et al., 1999; Patterson et al., 2003). It is used when observations from predetermined groups are characterised by two or more parameters (Quinn and Keough, 2002). In this study, the predetermined groups were Adélie penguin colonies (present) and control plots (absent). Discriminant analysis generates a linear combination of variables that maximises the probability of correctly assigning observations to their pre-determined groups and can also be used to predict the group-membership of test observations. It is mathematically identical to a single factor MANOVA, and where there are only two response groups, it derives a single discriminant function from a linear combination of the original variables. This function maximises the differences between the groups, and minimises the differences within the groups (Quinn and Keough, 2002).

One of the limitations of discriminant analysis is its assumption that the input data has a normal multivariate distribution. In addition, multivariate normality cannot be inferred from the univariate distribution of a parameter, and it has also been argued that discriminant analysis can still be applied in situations where the normality assumption is violated, though it may no longer be the optimal test (Flury and Riedwyl, 1988; Blackard and Dean, 1999).

The complexity of the discriminant analysis model increases as the number of variables increases. Here, a stepwise variable selection process was used to ensure that only those parameters that improved the proportion of correct classifications were incorporated in the final model. First, the data were split into training and test sets, (see *Model testing*, below) with 80% of the data points used for the model construction. Then all the static landscape parameters were included in a discriminant analysis model, using JMP 5.1. Finally, all the parameters were removed, and added back one-by-one until additional parameters did not further increase the proportion of correct classifications. Once an optimal model had been constructed, it was validated with the test set of data (see *Model testing*, below). The discriminant analysis formulae were then applied to each grid cell to predict the population trend class. This was done by applying the formulae in ArcGIS to generate maps of predicted Adélie penguin colony distributions.

Decision tree analysis

Decision trees are non-parametric and are capable of handling different data types (such as categorical and non-linear data) and also non-normal data (Weka Manual, 2006). The datasets in this study were not normally distributed, and as there is debate about the effect of non-normality on the performance of discriminant analysis, it was considered that decision trees would provide a check on the model performance of the discriminant analyses. Each of the study sites was examined individually, and each test was run on both the values of the individual cells and the colony/control plot mean values, as described above.

Each group in a decision tree is characterised by a typical value for the response variable, the number of observations in the group and the values of the explanatory variables that define it (De'Ath and Fabricius, 2000). This study used the J48 decision tree in Weka 3.4 (Witten and Frank, 2005), which is derived from the C4.5 model developed by J.R. Quinlan (1993). The algorithm chooses an attribute that best differentiates the output values and creates a separate branch for each output value. These subgroups (nodes) are considered to be terminal if all members of that group have the same output value or no further distinguishing features can be found. If the node is not terminal, the process is repeated (Weka Manual, 2006). An example of a decision tree is given in Fig. 3.8.

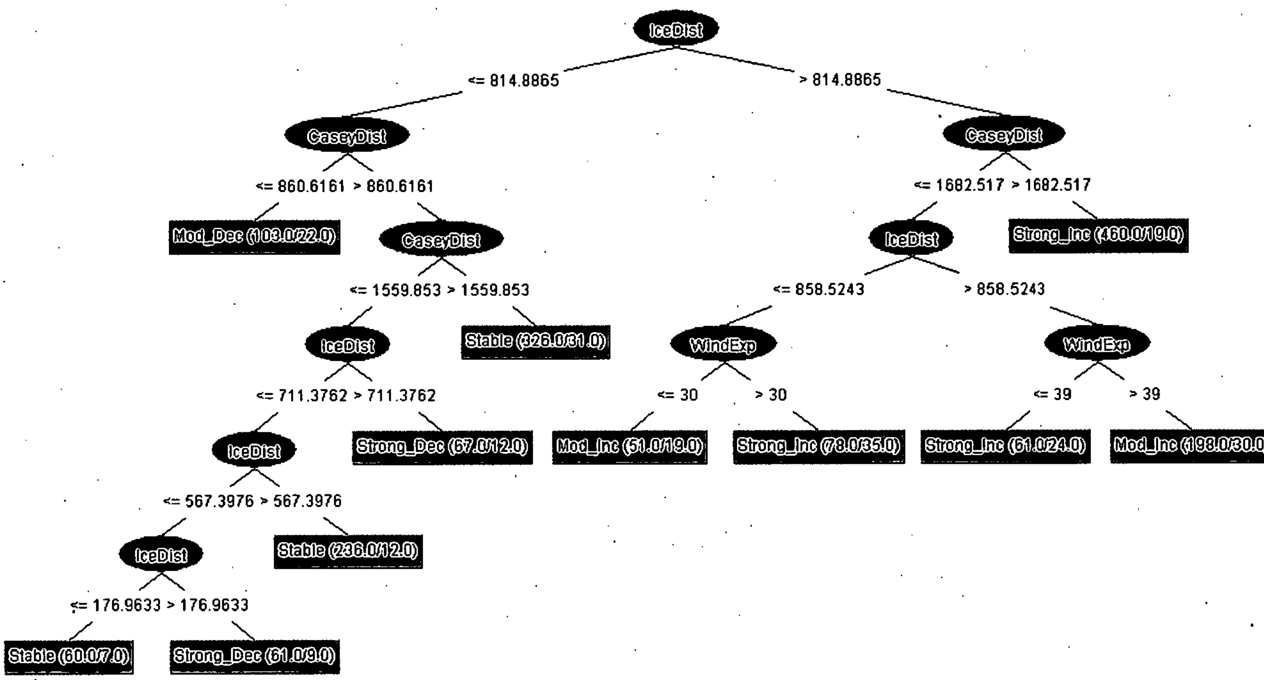


Fig. 3.8: Decision tree for population trends of Adelie penguin colonies at Shirley I, derived from individual cell values for proximity to human activity parameters.

The decision tree process requires a minimum object size to be set. In this case, the minimum object size refers to the minimum number of cells in each terminal node. In this study, the minimum object size was set by experimentally building decision trees with different object sizes to find the model with the smallest minimum object size that produced a tree with fewer than 20 splits in the data (leaves). As with the discriminant analyses, the decision tree models were constructed from training datasets (80% of the original data) and validated with the test sets.

In this study, the decision trees were used for statistical modelling but were not implemented in a GIS to produce predictive maps. Unlike discriminant analysis, a decision tree does not produce mathematical formulae that can be readily implemented in a GIS. Instead, it produces a set of conditional statements. Where a tree has only two or three leaves this may be applied in a GIS using available tools. However, more complex trees result in complex sets of nested conditional statements that would require scripting in order to be implemented efficiently.

Model testing

It has been argued that the optimal method for testing model performance is to validate the model with independently collected data (Blackard and Dean, 1999; Guisan and Zimmermann, 2000). This was not possible in the present study, as independent test data were not available. Instead, each model was tested by both cross-validation and validation. Cross-validation works by taking out an individual value from the training set and predicting its value, and it therefore tends to provide an optimistic measure of a model's accuracy (Burrough and McDonnell, 2000).

In addition to cross-validation, the models in this study were validated by splitting data sets into training and test sets, with 80% of the data points used for training, and 20% for testing. This approach has been recommended for studies in which independently collected data were not available (Blackard and Dean, 1999; Guisan and Zimmermann, 2000; Manel et al., 2000). The models based on colony mean values were validated using the test set for the individual cell value data. This was done because there were too few data points in the mean data sets to generate test sets that could reliably validate the models' performances. In addition, any model applied in a GIS would be applied to individual cell values, not colony means.

A purely random classifier, assigning observations into two groups could be expected to correctly classify 50% of the observations. Therefore, a model here was assumed to have some predictive power if validation showed that it correctly classified significantly more than 50% of the test data.

Univariate tests for Adélie penguin colony population trends and snow cover variables

Three variables associated with snow cover were examined for their ability to explain changes in population trends of Adélie penguin colony trends. These variables were the modelled snow cover for November 15, 2005, changes in the modelled snow cover between the first year of Adélie penguin count data (1959 for Whitney Pt and 1968 for Shirley I) and 2005, and exposure to prevailing winds. Only those colonies in which Adélie penguins have been known to nest during the period covered by count data were included in the population trend analyses.

When a Wilcoxon test is applied to more than two groups of data, it is equivalent to the Kruskal-Wallis non-parametric one-way analysis of variance test (Zar, 1999; Dytham, 2003). The Wilcoxon or Kruskal-Wallis tests can be applied in any situation where the parametric ANOVA test is applicable. If it is applied to normally distributed data, the Wilcoxon test is 95% as powerful as ANOVA, but if it is applied to non-normally distributed data, as in this study, it can be much more powerful (Zar, 1999). Here, Wilcoxon tests were used to assess the parameters for significant differences among the five population trend classes, using JMP. In addition, histograms and scatterplots were generated for visual exploration of the data.

Multivariate tests for Adélie penguin colony population trends and snow cover variables

The snow accumulation and wind exposure variables were tested for their ability to predict population trends for Adélie penguin colonies using the discriminant analysis and decision tree procedures outlined above. As with the models of colony distribution, the two study sites were examined separately and both the individual cell values and colony mean values were tested. The models were validated using the test sets of individual cell values. When classifying data into five groups, a purely random classifier could be expected to correctly predict the class with approximately 20% accuracy. Therefore, models were considered to have some explanatory power if they correctly classified significantly more than 20% of the test data points. As with the distribution investigations, maps of predicted Adélie penguin colony population trends were derived from the discriminant analyses, but not from the decision trees, due to the difficulties outlined above.

Adélie penguin colony population trends and proximity to human activities

The examination of the relationships between Adélie penguin colony population trends and the colonies' proximity and exposure to human activities used similar methods to the examination of the relationships between snow cover parameters and colony population trends. Wilcoxon tests were used to explore univariate differences among colonies in the five population trend classes. Discriminant analyses and decision trees were used to classify colonies into the five observed population trend classes. As with the snow accumulation tests, only those colonies in which Adélie penguins have been known to have nested during the period covered by the count data were incorporated in the analyses. All of these tests were conducted on both the individual cell values and the colony mean values, and validated with the test set of individual cell values.

The investigations into proximity to human activities differed from the distribution and snow accumulation analyses in that it used different parameters for the two study sites. Shirley I is 500m directly downwind from Casey and is regularly visited by station personnel. Therefore, the statistical tests investigated proximity to Casey, proximity to the sea-ice crossing point (used to access the island while the sea-ice is safe to cross on foot) and wind exposure (as a surrogate for exposure to potential airborne emissions from Casey) for their ability to explain population trends. The results of the analyses of wind exposure needed to be interpreted cautiously, as wind exposure was a surrogate for the effects of both wind speeds and of potential human activities on the colony population trends.

Access to Whitney Pt is restricted to scientists with permits (AAD, 2006). Therefore, a measure of proximity to a site access point, equivalent to the sea-ice crossing point on Shirley I, was not appropriate. In addition, the wind exposure layer could give no information about exposure to potential emissions from Casey because Whitney Pt is northeast of Casey, and the prevailing wind blows from the southeast. Therefore, measures of proximity to human activities for Whitney Pt were restricted to the distance from Casey. All the statistical tests and models were applied in an identical way to those for Shirley I.

4 Results

4.1 GIS landscape layers

4.1.1 Slope

The calculated slope data layers (Fig. 4.1) generally describe the terrain in the Windmill Is accurately; they display large areas of gentle slopes, with most of the altitude variation confined to areas of cliffs 10-15 m in height.

4.1.2 Aspect

The aspect data layers (Fig 4.2) displayed a strong agreement with the observed shape of the terrain.

4.1.3 Wetness index

The wetness index layers (Fig. 4.3) show that snow melt run-off for the two study sites is diverted into small gullies and runnels, rather than into large stream flows. The layers show a strong agreement with observed lakes and areas of ephemeral water, which occur in low-lying terrain during summer. The pattern of drainage lines in some parts of Shirley I (Fig. 4.3a) appear to be overly regular, and hence affected by artefacts in the DEM. These areas coincide with the areas of permanent snow as shown in the aerial photographs taken in mid-summer, which resulted in lower accuracy in those parts of the DEM (AADC, 1990; 2001).

4.1.4 Solar radiation

The highest modelled values for solar radiation are similar for Shirley I (4185.44 MJ/m^2) and Whitney Pt (4198.62 MJ/m^2), but the lowest values were different. The lowest value for Shirley I was 1840.91 MJ/m^2; this compared with 1206.81 MJ/m^2 for Whitney Pt. It is likely that this results from the cliffs that run east-west on Whitney Pt; similar cliffs on Shirley I are oriented northeast-southwest. The model results for both sites (Fig. 4.4) displayed strong agreement with calculated slope and aspect.

4.1.5 Planar and profile curvatures

The planar and profile curvature layers generally show a reasonable agreement with the observed shapes of slopes at the two sites. A visual inspection shows no strong patterns between the curvature patterns and the locations of Adélie penguin colonies. For Shirley I, these are the layers most obviously affected by artefacts in the DEM. These artefacts are reduced in the interpolated DEM used for the analyses, but smoothed areas where two or more stereo-models overlapped are still visible in the curvature layers. This problem does not arise for the Whitney Pt models as the Whitney Pt DEM was constructed from one stereo-pair of photographs (Figs. 4.5 and 4.6).

4.1.6 Surface roughness

The calculated surface roughness (standard deviation) data layers show a strong agreement with the calculated slopes, as they measure variability in elevations within a 3x3 cell neighbourhood (Fig. 4.7). The surface roughness (normalised) data layers are dominated by very small values (less than two) with values up to nine generally concentrated in areas where rough terrain was observed. In these layers, smooth but steep slopes have values below two (Fig. 4.8).

4.1.7 Adjacency

The adjacency layers calculate the proportion of cells in a 3x3 cell neighbourhood that contain Adélie penguins (Fig. 4.9). Only the cells that appear red in Fig. 4.9 were labelled as colony cells for the statistical analyses reported here. This process removed all those cells on the edge of the colonies.

4.1.8 Wind exposure

The wind exposure layers (Fig. 4.10) exhibit a negative relationship with solar radiation, as they represent exposure to the direction of the prevailing winds (south-southeast) and the highest solar radiation levels are recorded on north-facing slopes.

4.1.9 Snow accumulation model

The snow accumulation model generates maps of relative snow accumulation across the study sites (Fig. 4.11), rather than numeric values for snow depth, with strong negative associations with wind exposure. The model could only be validated visually, as there is no effective data available to compare it to. The aerial photographs show areas of permanent snow and ice, which are not available for transport. However, areas of snow ablation and accumulation show a general agreement with snow accumulation patterns in the aerial photographs, and with those observed during the summer of 2005/06 (AADC, 1990; 2001). The model does not predict the areas of permanent snow and ice that occur along the eastern and southern coasts of the study sites, but it does predict the snow that accumulates in the valley to the south of Whitney Pt. This suggests that the model cannot account for the effects of sea-ice that builds up around the Antarctic coast. The model was applied for the spring months in 1959 and 2005 for Whitney Pt and in 1968 and 2005 for Shirley I. No area of either study site shows consistent increases or decreases of snow between the two years for which snow cover was modelled. The large areas of modelled snow accumulation or ablation in 2005 are little different to the other years modelled, and other areas showed fine-scale spatial patterns of change (Fig. 4.12).

4.1.10 Proximity to Human Activities

For Whitney Pt, only a distance from Casey layer was generated (Fig. 4.13a).For Shirley I, the layers showing distance from Casey and from the sea-ice crossing point show strong agreement (Fig. 4.13b and 4.14).

4.1.11 Population trends

The population trends for Adélie penguin colonies were calculated as the percentage difference in the number of breeding pairs between the year in which a colony was first counted and 2005. The number of colonies in each trend classes is presented in Table 4.1.

Table 4.1: The number of colonies in the five population trend classes for Whitney Pt and Shirley I.

Population Trend	*Class*	*Number of Colonies*	
		Whitney Pt	**Shirley I**
<50% (Strong Decrease)	1	2	15
50-80% (Moderate Decrease)	2	1	8
80-120% (Stable)	3	5	9
120-150% (Moderate Increase	4	2	2
>150% (Strong Increase)	5	23	10
Total		33	44

4.1.12 Summary Statistics

Summary statistics for each data set are provided in appendix 1. The data are divided into Whitney Pt and Shirley I sets, and further into individual cell values and colony/control plot mean values. Summaries are presented for each population trend class of colony cells. This provides an explanation of the spread of data used in the models.

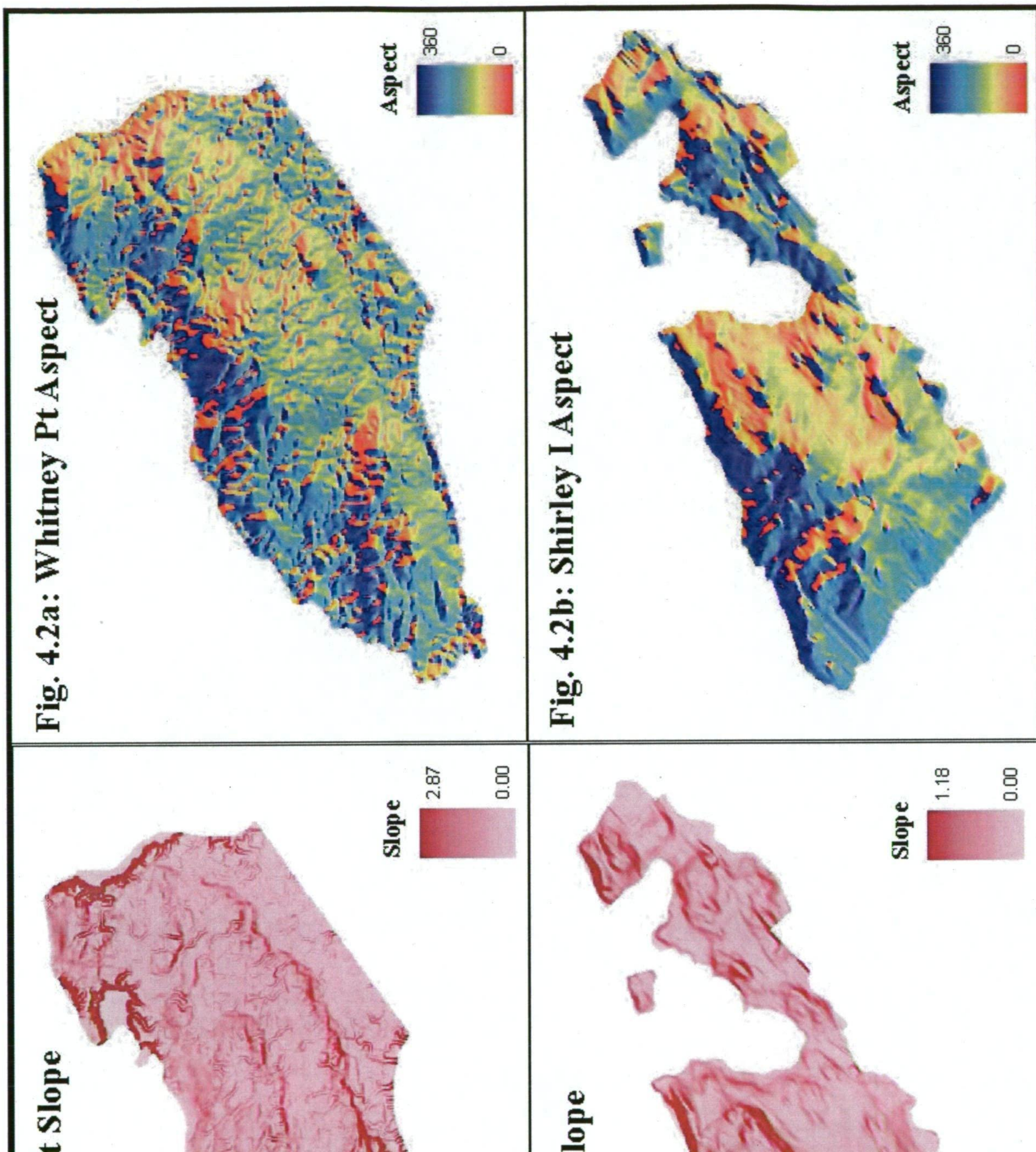


Fig. 4.1a: Whitney Pt Slope

Fig. 4.2a: Whitney Pt Aspect

Fig. 4.1b: Shirley I Slope

Fig. 4.2b: Shirley I Aspect

Fig. 4.3a: Whitney Pt Wetness Index

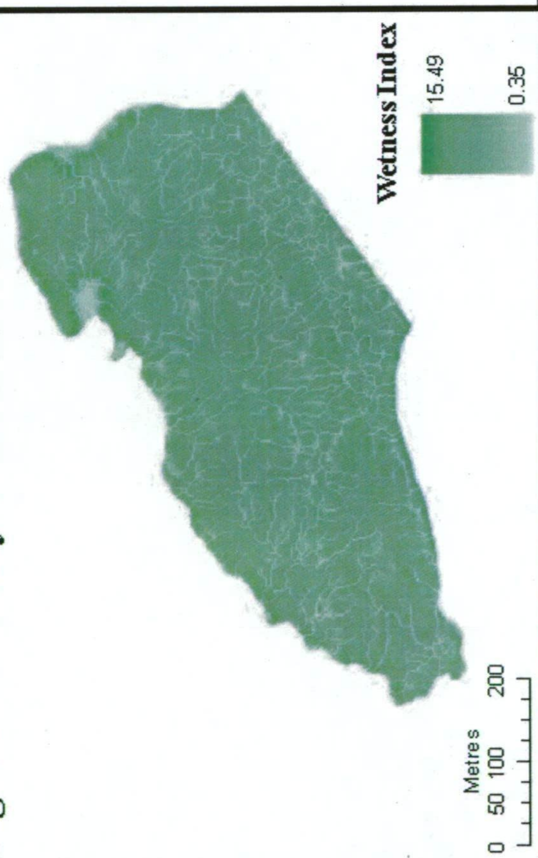


Fig. 4.4a: Whitney Pt Solar Radiation

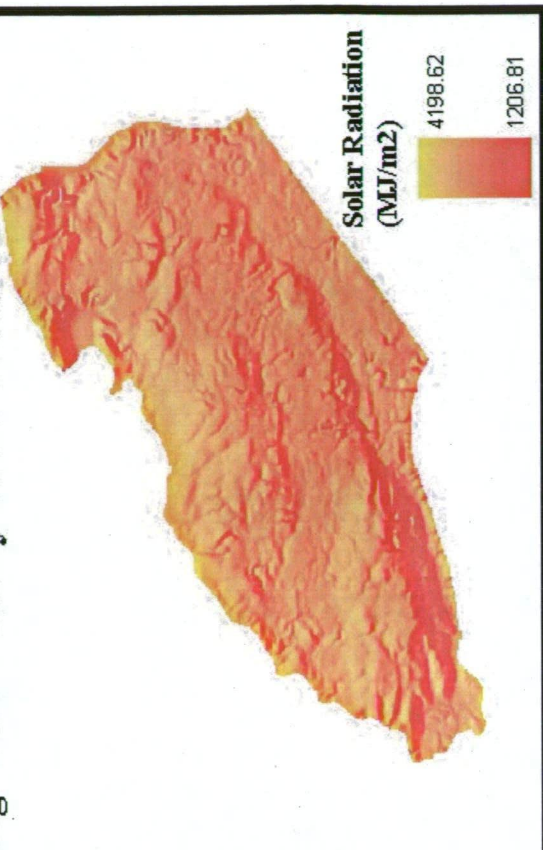


Fig. 4.3b: Shirley I Wetness Index

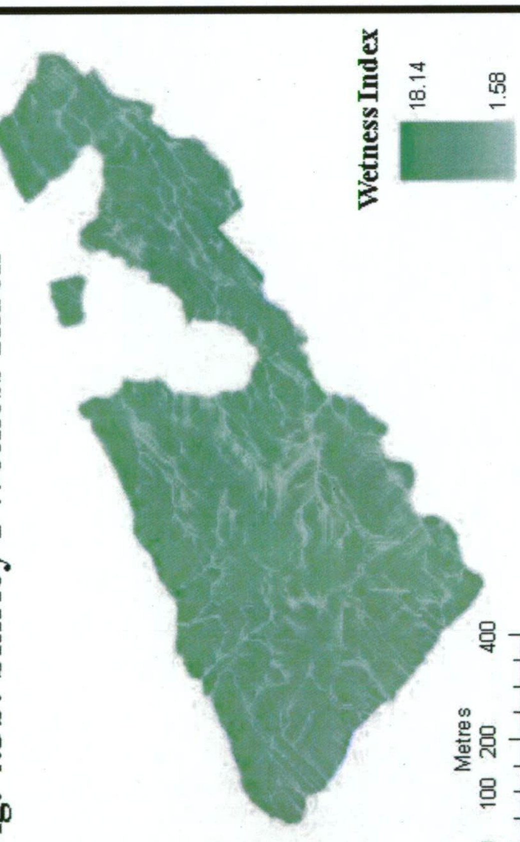
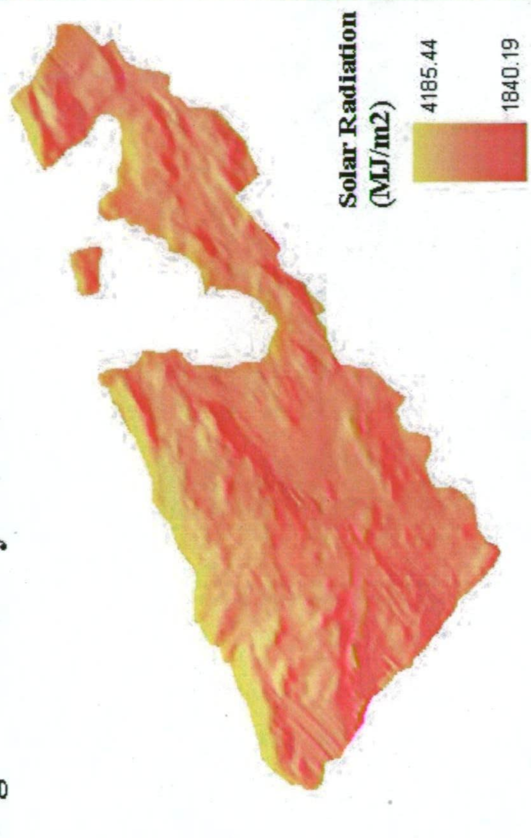


Fig. 4.4b: Shirley I Solar Radiation



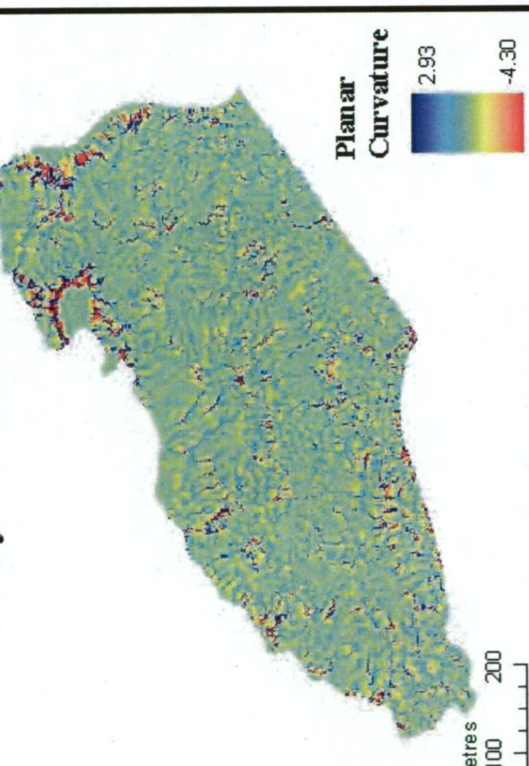


Fig. 4.5a: Whitney Pt Planar Curvature

Fig. 4.6a: Whitney Pt Profile Curvature

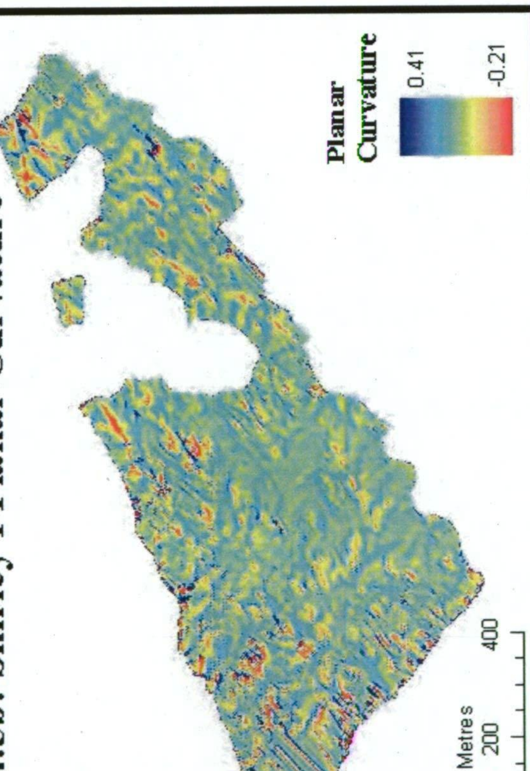


Fig. 4.5b: Shirley I Planar Curvature

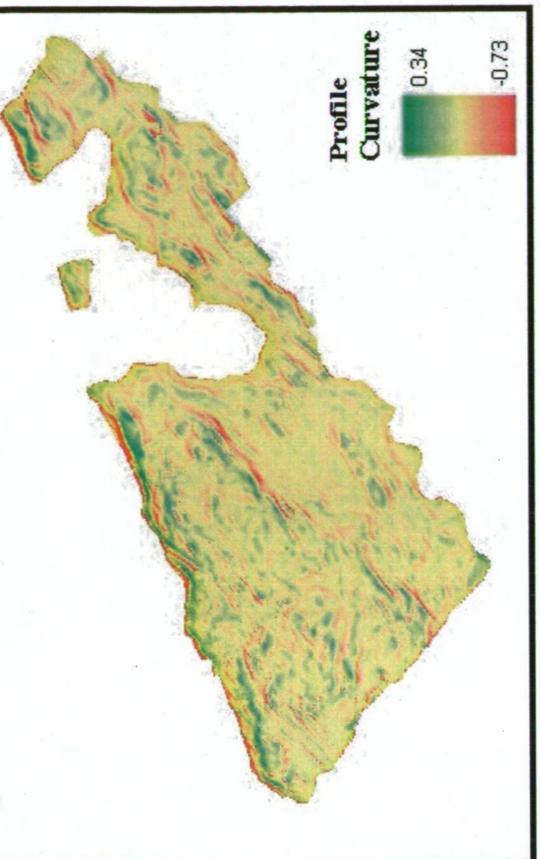


Fig. 4.6b: Shirley I Profile Curvature

Fig. 4.7a: Whitney Pt Surface Roughness (Standard Deviation)

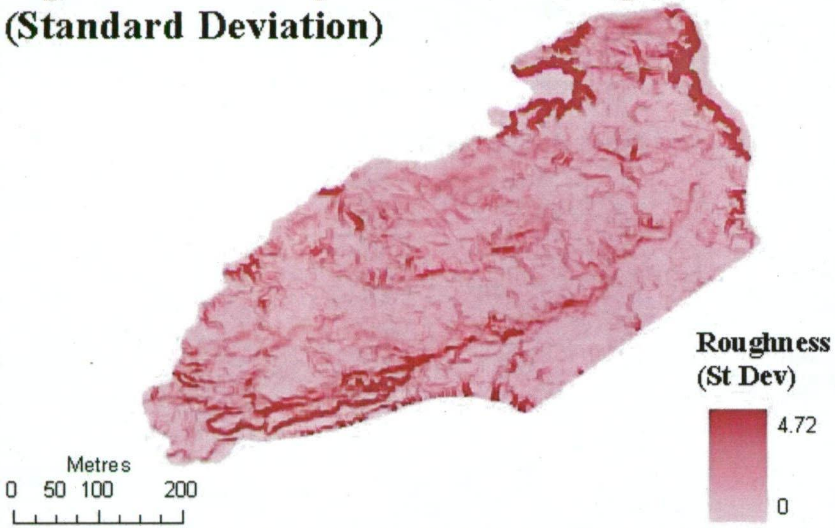


Fig. 4.8a: Whitney Pt Surface Roughness (Normalised)

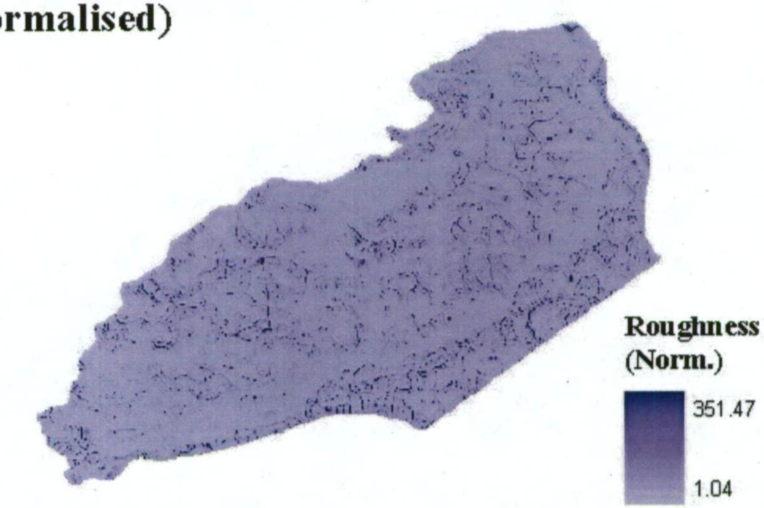


Fig. 4.7b: Shirley I Surface Roughness (Standard Deviation)

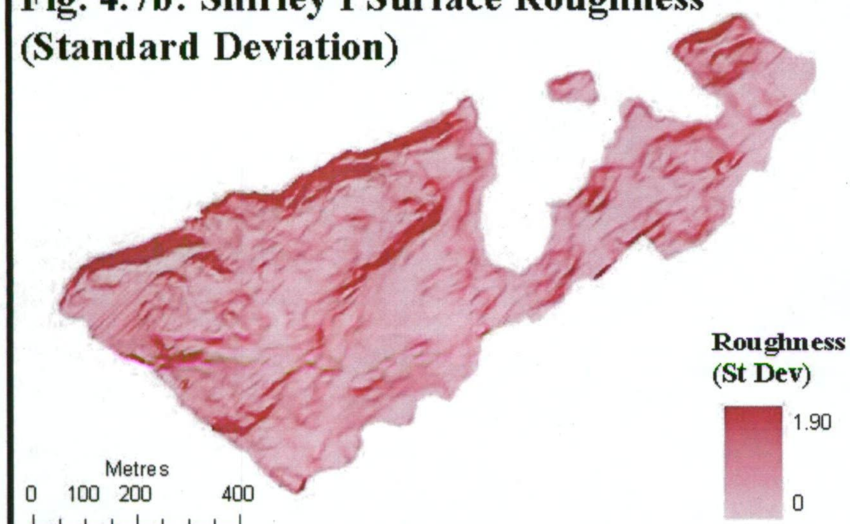
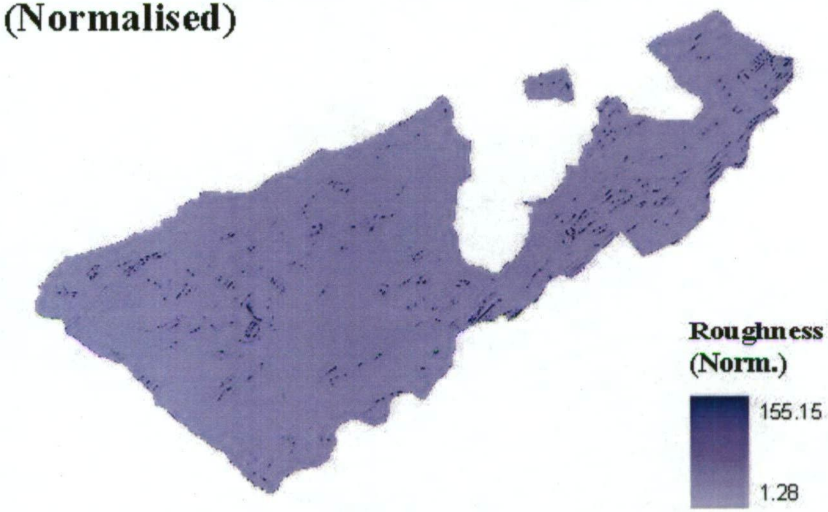


Fig. 4.8b: Shirley I Surface Roughness (Normalised)



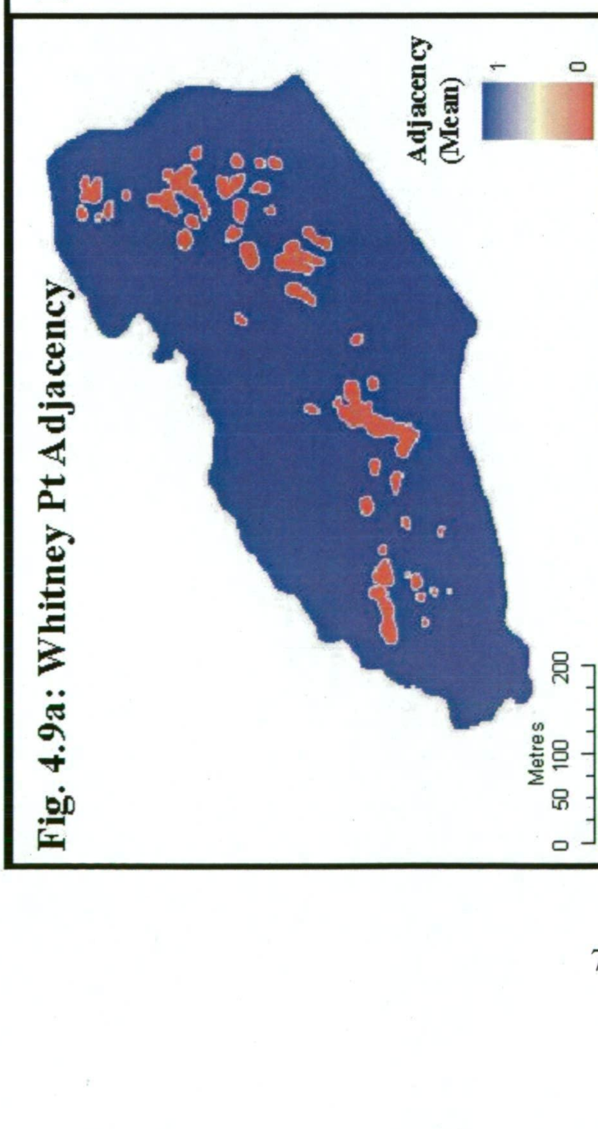


Fig. 4.9a: Whitney Pt Adjacency

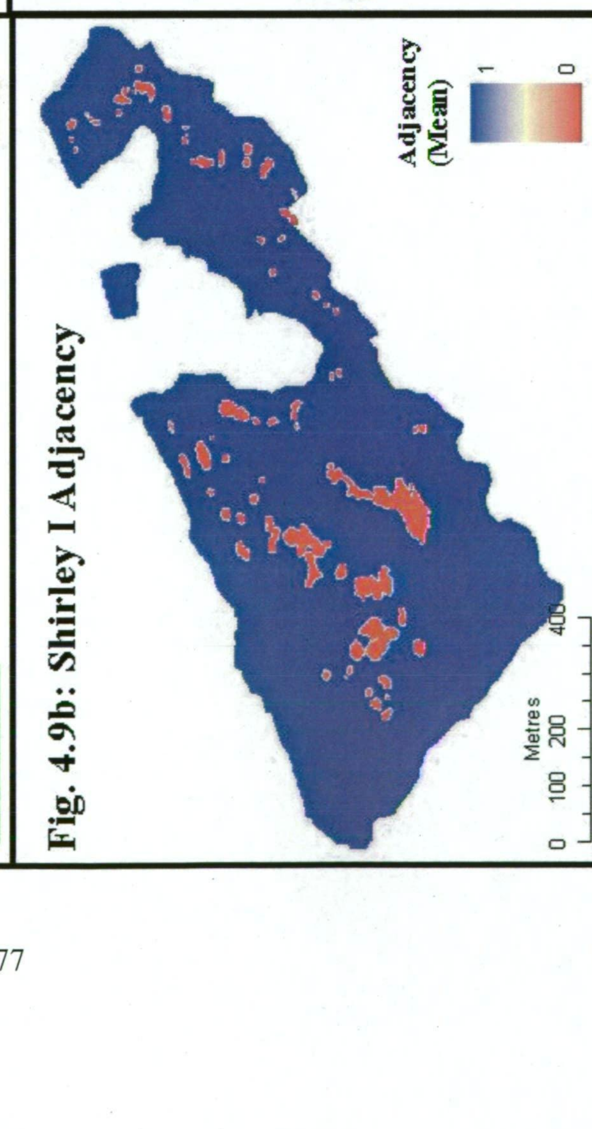


Fig. 4.9b: Shirley I Adjacency

Fig. 4.10a: Whitney Pt Wind Exposure

Fig. 4.10b: Shirley I Wind Exposure

Fig. 4.11a: Whitney Pt Modelled Snow Cover (15 Nov, 2005)

Fig. 4.12a: Whitney Pt Snow Cover Difference (15 Nov 1959 - 15 Nov 2005)

Fig. 4.11b: Shirley I Modelled Snow Cover (15 Nov, 2005)

Fig. 4.12b: Shirley I Snow Cover Difference (15 Nov 1968 - 15 Nov 2005)

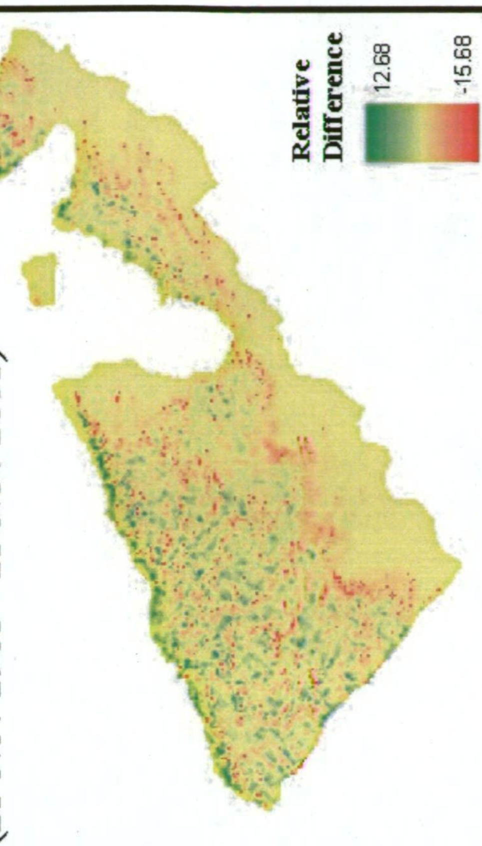


Fig. 4.13a Whitney Pt Distance from Casey

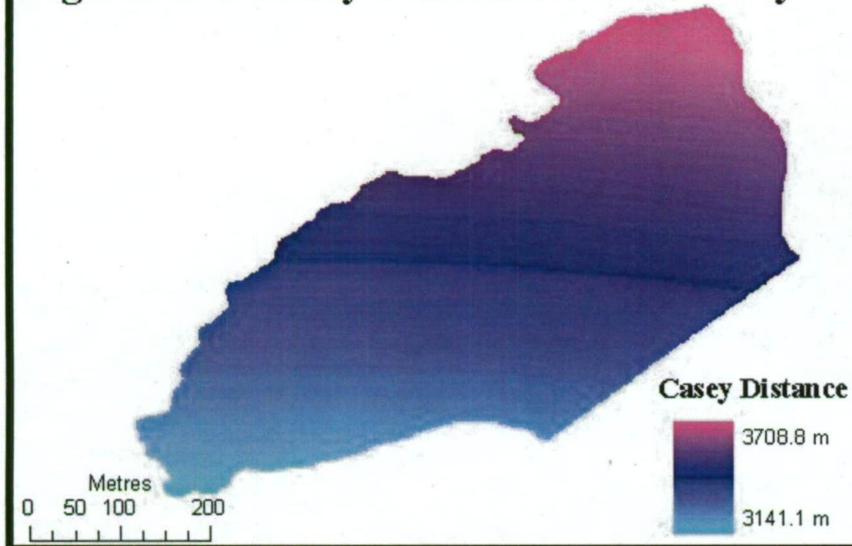


Fig. 4.14 Shirley I Distance from Sea-Ice Crossing Point

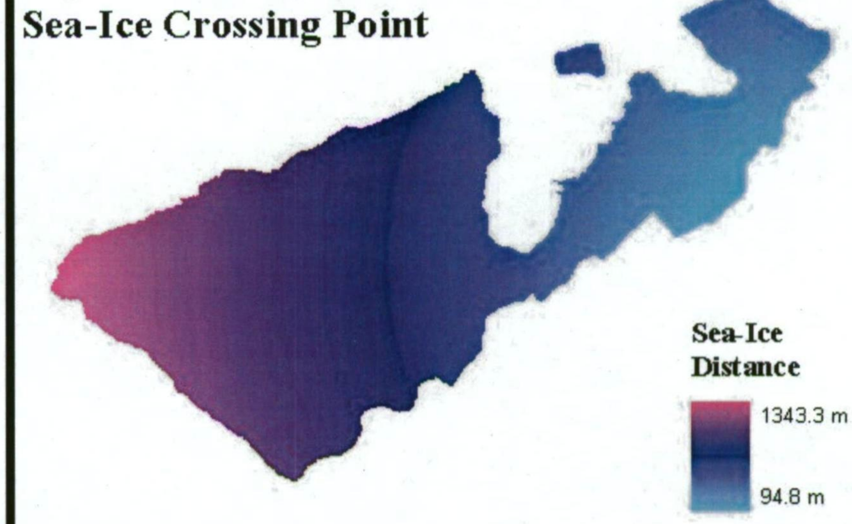
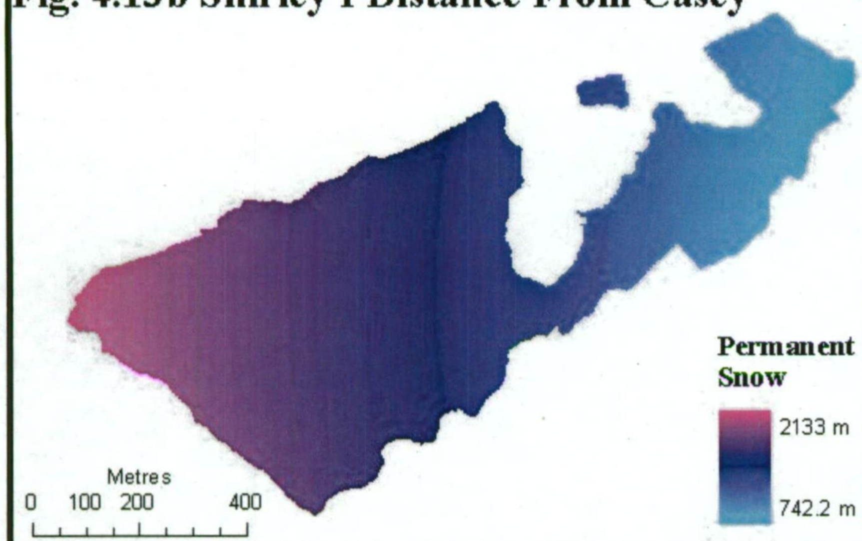


Fig. 4.13b Shirley I Distance From Casey



4.2 Adélie penguin colony distribution

This section presents the results of the tests for the ability of individual landscape parameters, and of the discriminant analyses and decision trees derived from them to explain the distribution of Adélie penguin colonies within the two study sites as expressed in the following null hypothesis.

H_{NULL} 1 *Static landscape variables (slope, drainage, aspect, planar and profile curvature, surface roughness, wind exposure, snow cover and solar radiation) cannot predict the locations of current and relic Adélie penguin colonies at Shirley I and Whitney Pt.*

The results are presented separately for each study site and for the tests conducted using individual cell values and colony mean values.

4.2.1 Univariate Analyses

4.2.1.1 Whitney Pt

Individual Cell Values

Wilcoxon tests show significant differences between those cells in colony sites and those in control plots for the static landscape parameters of surface roughness (standard deviation), slope, elevation, wetness index, surface roughness (normalised) and the modelled snow cover for 15 November 2005 (Table 4.2). Adélie penguin colonies do not occur on the steepest slopes irrespective of aspect, or on the moderately steep, south-facing slopes that have low solar radiation. The penguin colonies occurred in several altitude "zones", separated by 1-2m. It is likely that these zones are the result of the terrain being dominated by a mixture of plateaux and cliffs.

Table 4.2: Wilcoxon/Kruskal-Wallis rank sum tests for differences among individual cell values for static landscape parameters in colonies and control plots at Whitney Pt.

Variable	Z	Prob>Z	Chi-Square	DF	Prob>ChiSq	Significance
Roughness (St Dev)	12.642	0.000	159.823	1	<0.0001	Significant
Slope	11.308	0.000	127.866	1	<0.0001	Significant
Elevation	10.466	0.000	109.546	1	<0.0001	Significant
Wetness Index	-4.578	<0.0001	20.959	1	<0.0001	Significant
Snow 2005	3.779	0.000	14.280	1	0.000	Significant
Roughness (Norm)	-2.301	0.021	5.297	1	0.021	Significant
Profile Curvature	-0.921	0.357	0.849	1	0.357	N.S.
Solar Radiation	-0.463	0.643	0.215	1	0.643	N.S.
Wind Exposure	-0.299	0.765	0.090	1	0.765	N.S.
Plan Curvature	-0.284	0.777	0.008	1	0.777	N.S.

Colony/Control Plot Mean Values

Wilcoxon tests showed significant differences between the mean values for colony sites and control plots for surface roughness (standard deviation) (Table 4.3). One result of using mean values rather than individual cell values is that much of the within colony variance is lost. It is likely that this explains the difference in the number of parameters found to have significant differences in distribution. It is likely that the mean values represent a truer picture of significant differences, because spatial autocorrelation means the values for the individual cells within any one colony share information. Among-colony spatial autocorrelation is still likely to be a factor in all the tests on mean values, but within-colony autocorrelation is removed. At the same time, no colony exists on homogeneous terrain, and the mean values ignore within-colony variance.

Table 4.3: Wilcoxon/Kruskal-Wallis rank sum tests for differences between colony and control plot mean values for static landscape variables on Whitney Pt.

Variable	Z	Prob>Z	Chi-Square	DF	Prob>ChiSq	Significance
Roughness (St Dev)	-1.966	0.049	3.883	1	0.049	Significant
Slope	-1.903	0.057	3.639	1	0.057	N.S.
Snow 2005	-1.494	0.135	2.247	1	0.134	N.S.
Roughness (Norm)	-0.909	0.363	0.836	1	0.361	N.S.
Elevation	-0.702	0.482	0.500	1	0.479	N.S.
Wind Exposure	0.545	0.586	0.303	1	0.582	N.S.
Wetness Index	-0.285	0.776	0.084	1	0.772	N.S.
Solar Radiation	-0.201	0.840	0.043	1	0.837	N.S.
Profile Curvature	0.079	0.937	0.007	1	0.933	N.S.
Planar Curvature	0.000	1.000	0.000	1	0.996	N.S.

4.2.1.2 Shirley I

Individual Cell Values

Wilcoxon tests show significant differences between cells in colony sites and control plots on Shirley I for all static landscape parameters except planar curvature (Table 4.4). Adélie penguin colonies do not occur on steep slopes, rough terrain or in areas with especially high or low solar radiation. These variables tend to covary. That is, rough terrain is associated with slopes, and the highest solar radiation is on north-facing slopes, which on Shirley I tend to be very steep. Likewise, the steepest south-facing slopes have the lowest solar radiation.

Table 4.4: Wilcoxon/Kruskal-Wallis rank sum tests for differences between individual cell values for static landscape parameters in colonies and control plots on Shirley I.

Variable	Z	Prob>Z	ChiSquare	DF	Prob>ChiSq	Significance
Elevation	30.025	0.000	901.490	1	<0.0001	Significant
Roughness (StDev)	-28.173	0.000	793.736	1	<0.0001	Significant
Slope	-27.687	0.000	766.572	1	<0.0001	Significant
Snow 2005	13.321	0.000	177.448	1	<0.0001	Significant
Solar Radiation	13.175	0.000	173.583	1	<0.0001	Significant
Wetness Index	11.524	0.000	132.811	1	<0.0001	Significant
Wind Exposure	4.064	<0.0001	16.517	1	<0.0001	Significant
Profile Curvature	3.079	0.002	9.480	1	0.002	Significant
Roughness (Norm)	2.111	0.035	4.456	1	0.035	Significant
Planar Curvature	-0.927	0.354	0.859	1	0.354	N.S.

Colony/Control Plot Means

Wilcoxon tests show significant differences between the mean values for colony sites and control plots on Shirley I for planar curvature, surface roughness (standard deviation), and slope (Table 4.5). It is likely that a combination of the loss of variance associated with using mean values, and the reduction of the effect of spatial autocorrelation explains the differences in these results compared with the tests on individual cell values. Again, colonies are rare in sites with steep slopes and rough terrain.

Table 4.5: Wilcoxon/Kruskal-Wallis rank sum tests for differences between colony and control plot mean values for static landscape parameters on Shirley I.

Variable	Z	Prob>Z	ChiSquare	DF	Prob>ChiSq	Significance
Slope	-3.842	0.000	14.782	1	0.000	Significant
Roughness (StDev)	-3.810	0.000	14.535	1	0.000	Significant
Planar Curvature	-1.969	0.049	3.889	1	0.049	Significant
SolarRad	1.887	0.059	3.569	1	0.059	N.S.
Profile	1.698	0.090	2.892	1	0.089	N.S.
Elevation	1.483	0.138	2.207	1	0.137	N.S.
Roughness (Norm)	1.415	0.157	2.008	1	0.156	N.S.
Snow 2005	0.803	0.422	0.649	1	0.420	N.S.
Wetness Index	-0.434	0.664	0.191	1	0.662	N.S.
Wind Exposure	-0.196	0.845	0.039	1	0.843	N.S.

4.2.2 Discriminant Analyses

4.2.2.1 Whitney Pt

Individual Cell Values

A stepwise discriminant analysis model was constructed with elevation, surface roughness (standard deviation) and solar radiation as input parameters (App. 2). These parameters increase the predictive power of the model, and are listed in the order of their value in predicting presence or absence. Other parameters were excluded because they did not improve the model performance. Cross-validation shows that this model accurately predicts the presence or absence of nesting Adélie penguins in a cell with 68.1% accuracy. Validation, which is generally considered a more robust assessment of accuracy (Blackard and Dean, 1999; Guisan and Zimmermann, 2000), shows an overall predictive accuracy of 70.5%. The confusion matrix (Table 4.6) shows that this model has a higher proportion of false positives than false negatives. Thus, the model predicts more areas of suitable habitat than are occupied. Visual comparison of the predicted and observed distribution of Adélie penguins (Fig. 4.15) shows agreement between predicted and observed distributions. The use of elevation in this model is considered to be the result of the plateau-dominated terrain, rather than the result of birds preferring particular elevations. Surface roughness covaries with slope, and the distribution map predicts colony occurrence in areas with gentle slopes and at least moderate levels of solar radiation.

Table 4.6: Confusion matrix for validation of the discriminant analysis model based on the individual cell values of colonies and control plots in Whitney Pt. Percentage values in brackets are the percentage of the observed total.

		Observed Distribution		
		Absent	**Present**	**Total**
Predicted Distribution	**Absent**	132 (64.1%)	53 (23.6%)	185 (42.9%)
	Present	74 (35.9%)	172 (76.4%)	246 (57.1%)
	Total	206	225	431

Colony/Control Plot Mean Values

The discriminant analysis model was constructed using wind exposure, surface roughness (normalised), surface roughness (standard deviation) and wetness index as input parameters, entered in the order listed (App. 2). The other static landscape parameters did not increase the predictive power of the model. Cross-validation shows that this model has an overall accuracy of 65.6%. When validated with the test set of individual cell values, the model shows an overall predictive accuracy of 65.4%. The model most accurately predicts the presence of nesting penguins, and has a high proportion of false positive errors (Table 4.7). In the resulting predictive map, areas of predicted suitable and unsuitable habitat are more patchily distributed than in the map derived from individual cell values. It appears that this patchiness results from the inclusion of the wetness index in the model, as many of the breaks in habitat suitability occur along drainage lines (Fig. 4.16). The use of mean values reduces within-group variance, and the effects of spatial autocorrelation. This model produces fewer false absences than the model based on individual cell values, but also produces more false presences. It is likely that collinearities in the data sets account for the differences in the input parameters for the two models. Again, the map predicts penguin colonies in sites with gentle slopes and with moderate degrees of wind exposure.

Table 4.7: Confusion matrix for the validation of the discriminant analysis model of Adélie penguin distribution based on colony and control plot mean values for Whitney Pt. The percentage values in brackets show the proportion of the observed total.

		Observed Distribution		
		Absent	Present	Total
Predicted Distribution	Absent	119 (57.8%)	62 (27.6%)	181 (42%)
	Present	87 (42.2%)	163 (72.4%)	250 (58%)
	Total	206	225	431

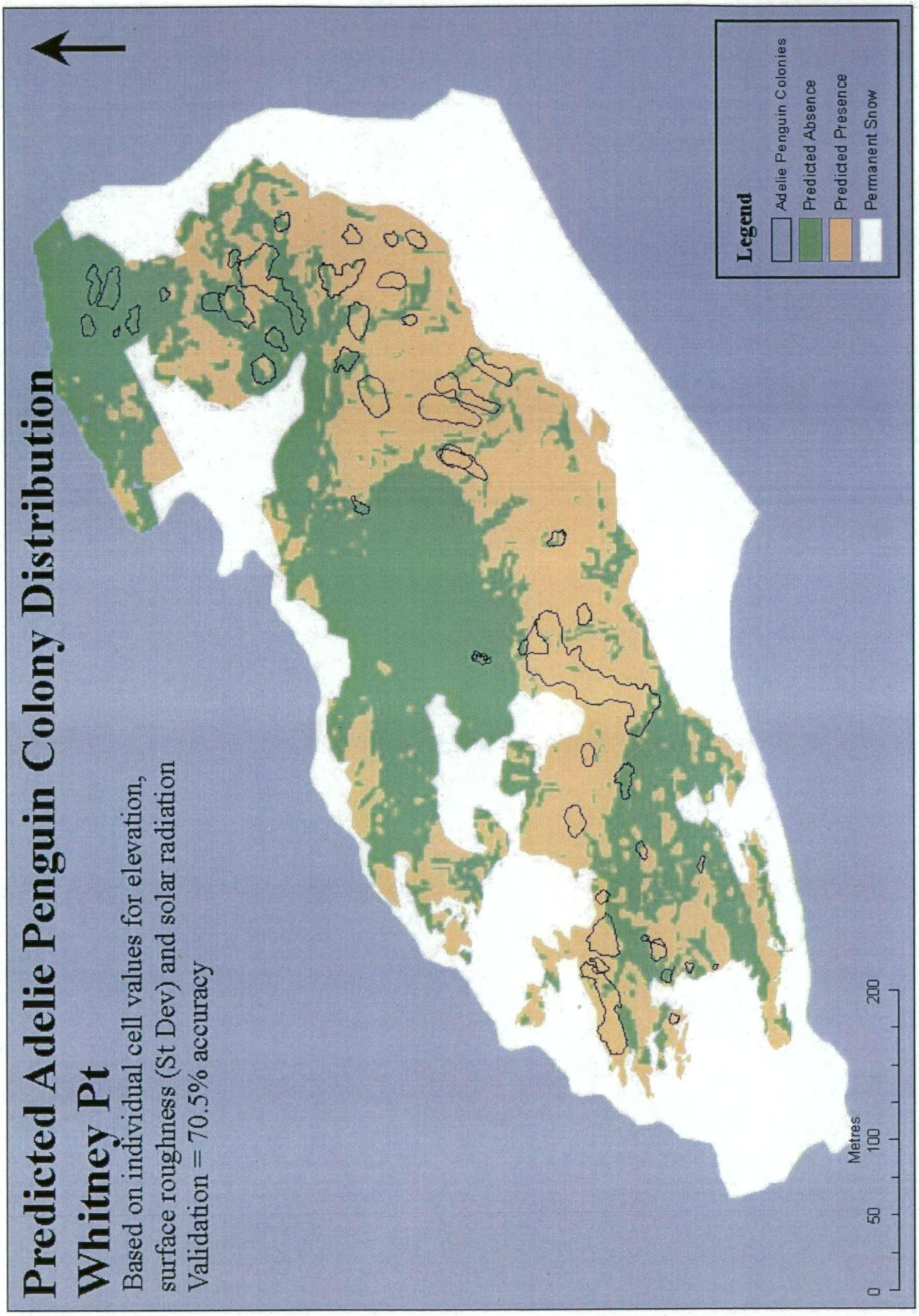


Fig. 4.15: Predicted Adélie penguin colony distribution at Whitney Pt based on discriminant analysis of individual cell values for static landscape parameters.

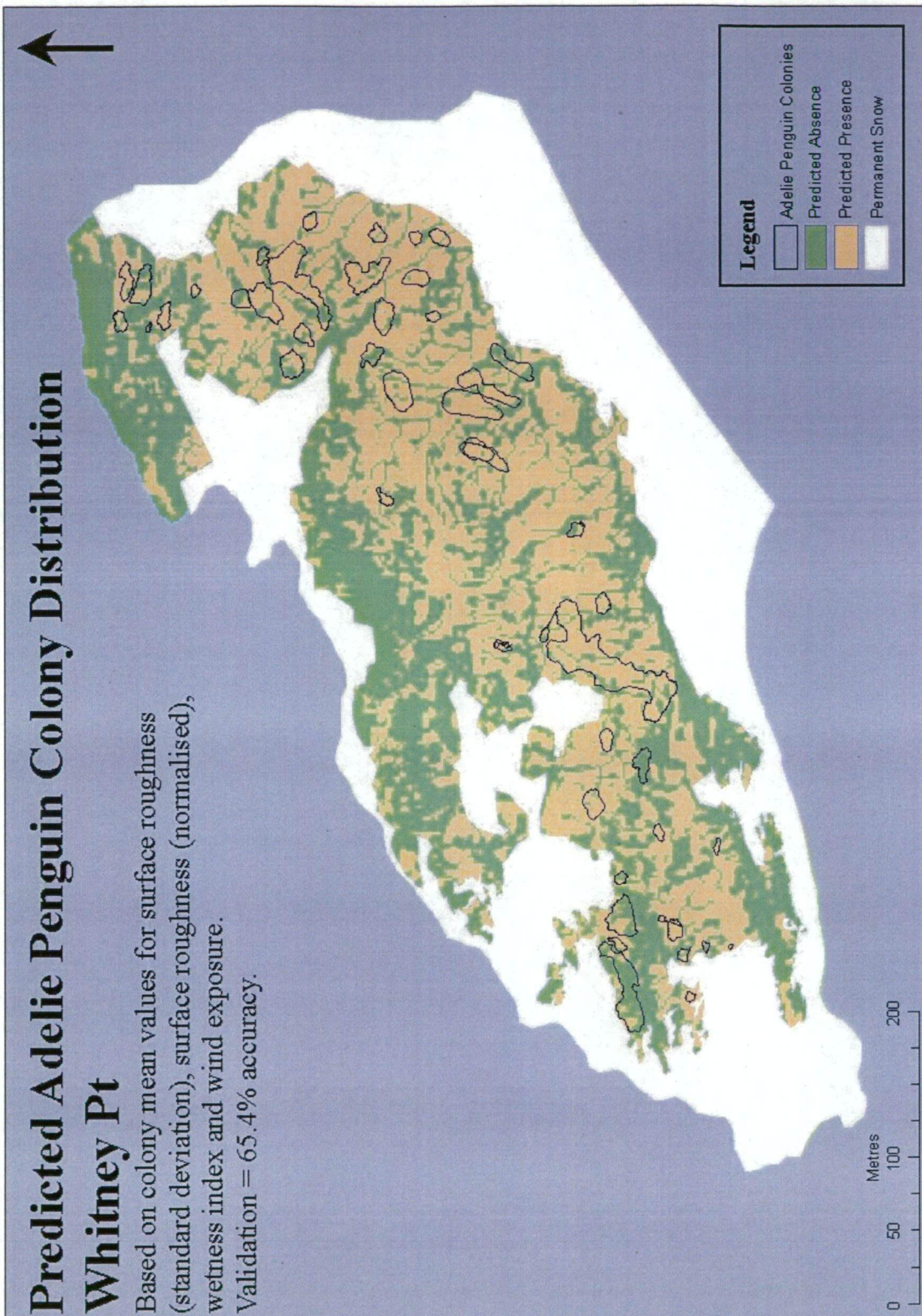


Fig. 4.16: Predicted Adélie penguin colony distribution at Whitney Pt based on discriminant analysis of colony/control plot mean values for static landscape parameters.

4.2.2.2 Shirley I

Individual Cell Values

Surface roughness (standard deviation), elevation, solar radiation, wind exposure, slope and the 2005 snow cover parameters increased the predictive power of the stepwise discriminant analysis model (App. 2). The other static landscape parameters were excluded because they did not increase the predictive power of the model. Cross-validation shows that this model correctly classifies 79.6% of cells. Validation shows the overall model accuracy is 78.9% (Table 4.8). The model predicts the presences of nesting penguins better than absences, with a high proportion of false positives (Fig. 4.17). Both predicted and observed extant presences are concentrated on the large plateau at the western end of the island. Large areas of predicted presences also occur near sea level at the western end of the island, in the area dominated by a relic colony, but partially covered by snow in the available aerial photography.

Table 4.8: Confusion matrix for validation of discriminant analysis model based on individual cell values in colonies and control plots on Shirley I.

		Observed Distribution					
		Absent		**Present**		**Total**	
Predicted Distribution	**Absent**	433	(66.7%)	60	(9.1%)	493	(37.8%)
	Present	216	(33.3%)	597	(90.9%)	813	(62.3%)
	Total	649		657		1306	

Colony/Control Plot Means

A discriminant analysis model to predict Adélie penguin distribution on Shirley I was constructed using the colony and control plot mean values for slope, snow cover in November 2005, wetness index, solar radiation and wind exposure as inputs (App. 2). Cross-validation shows that this model predicts Adélie penguin colony distribution with 76.5% accuracy. Validation with the test set of individual cell values shows the overall model accuracy is 70.8% (Table 4.9). The model predicts colony presences more strongly than absences (Fig. 4.18). Again the wetness index influences the model based on colony means, but not the model derived from individual cell values. It appears that this parameter is affected by spatial autocorrelation, and gains predictive power when within-colony variance is removed. Surface roughness is not used as a parameter in this model, unlike that derived from individual values. This may be because the roughness variables covary with slope.

Table 4.9: Confusion matrix for validation of discriminant analysis model based on colony and control plot mean values of Shirley I.

		Observed Distribution		
		Absent	Present	Total
Predicted Distribution	Absent	412 (63.5%)	114 (21.9%)	556 (42.6%)
	Present	237 (36.5%)	513 (78.1%)	750 (57.4%)
	Total	649	657	1306

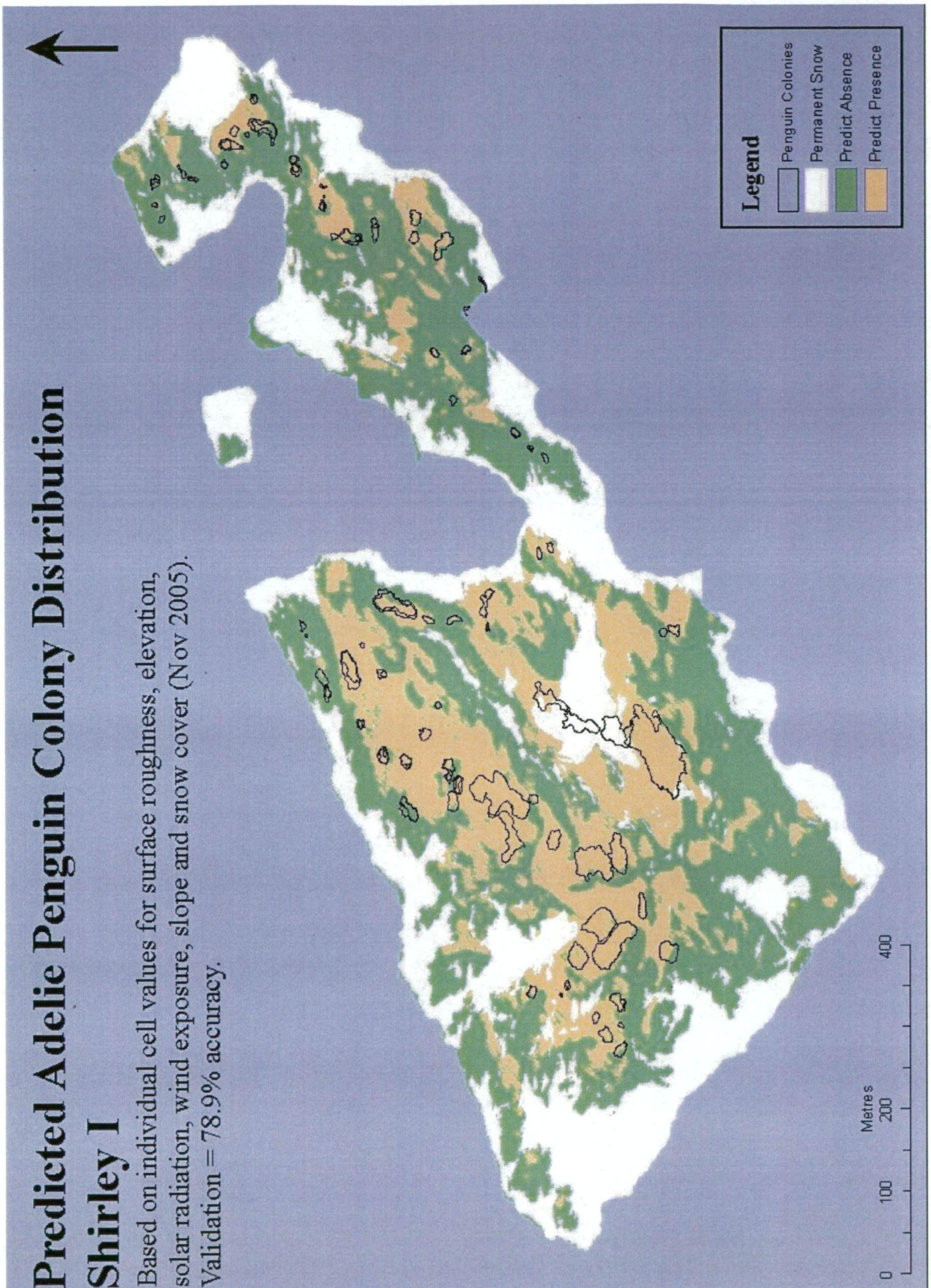


Fig. 4.17: Predicted distribution of Adélie penguin colonies on Shirley I, based on discriminant analysis of individual cell values for static landscape parameters.

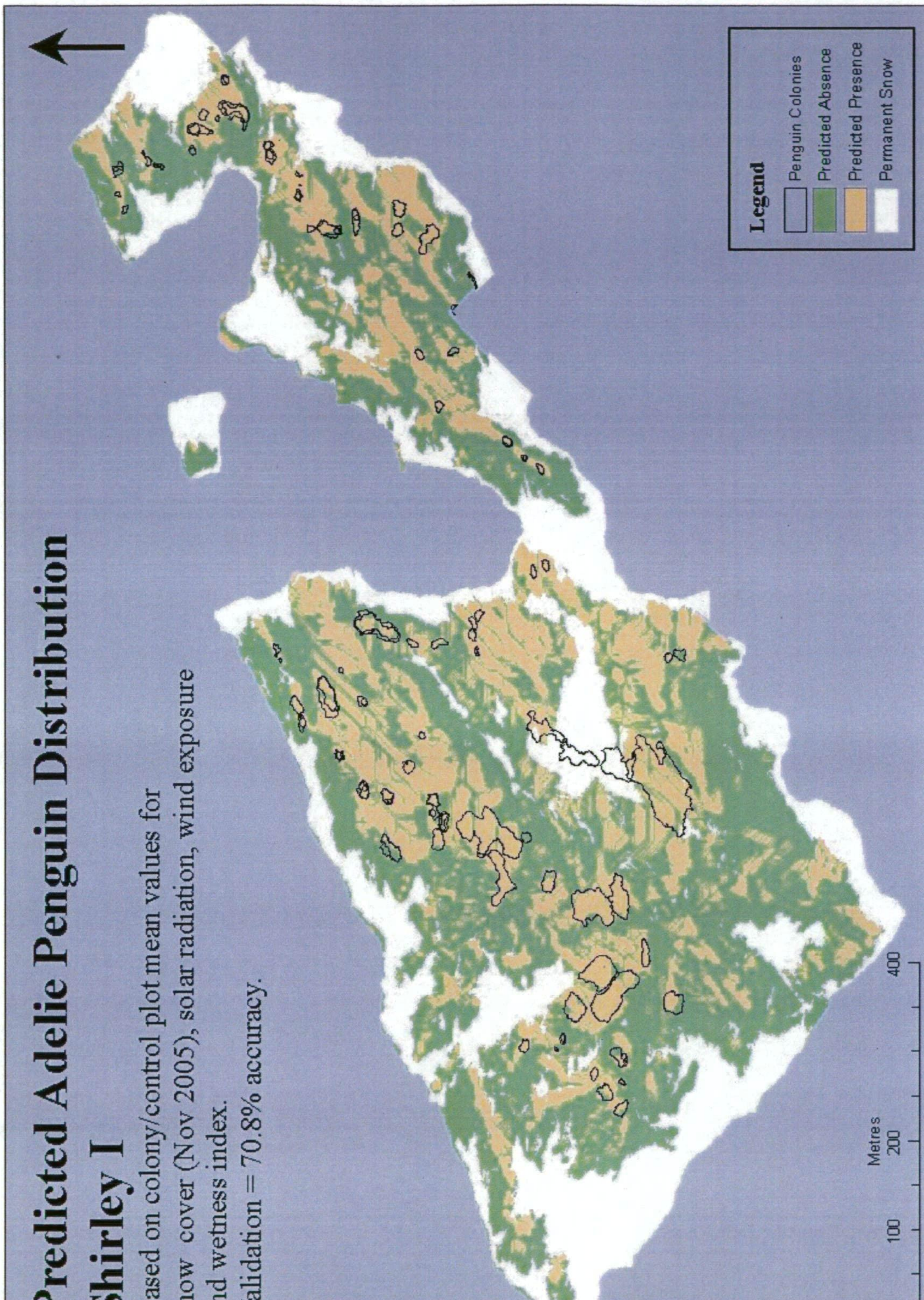


Fig. 4.18: Predicted Adélie penguin colony distribution based on discriminant analysis of colony/control plot mean values for static landscape parameters.

4.2.3 Decision Tree Analyses

4.2.3.1 Whitney Pt

Individual Cell Values

A decision tree was constructed based on individual cell values with a minimum object size of 100 and seven leaves (App. 3). Cross-validation shows that it correctly predicts 73.4% of Adélie penguin colony presences or absences. Validation shows that the model has 74% accuracy. Solar radiation, elevation, aspect, snow cover in 2005 and wind exposure were used as inputs. The other static landscape parameters did not improve the predictive power of the model. The confusion matrix (Table 4.10) shows that the model performance was similar for both presence and absence predictions, though it predicted colony presences slightly better than absences. However, decision trees could not be assessed visually, because of the difficulty in implementing them within a GIS. The tree predicts that Adélie penguins are generally absent where solar radiation is below 2434.05 MJ/m^2., and that where solar radiation is higher than this, Adélie penguins will nest at low altitudes (below 13.4m ASL). After these splits, predictions are based on aspect (greater than or less than 130 degrees – the direction of the prevailing winds), modelled snow cover in November 2005 and elevation (App. 3).

Table 4.10: Confusion matrix for validation of the decision tree model of Adélie penguin distribution based on individual cell values for Whitney Pt.

		Observed Distribution					
		Absent		Present		Total	
Predicted Distribution	**Absent**	151	(73.3%)	57	(25.33%)	208	(48.3%)
	Present	55	(26.7%)	168	(74.67%)	223	(51.7%)
	Total	206		225		431	

Colony/Control Plot Mean Values

A decision tree based on the colony and control plot mean values for static landscape parameters predicts Adélie penguin colony presence or absence with 59.4% accuracy as measured by cross-validation, and with 51.7% accuracy as measured by validation with the test set of individual cell values. The tree has a minimum object size of two and four leaves. Surface roughness (standard deviation) and drainage improved the predictive power of the model, but the other parameters were not selected because they did not increase the predictive power (App. 3). The confusion matrix (Table 4.11) shows that the model predicts colony presences slightly better than absences. The model predicts Adélie penguins to be absent where the surface roughness has a standard deviation greater than 0.85 and to be present where the wetness index is below 3.46 or above 5.06 (App. 3).

Table 4.11: Confusion matrix for validation of the decision tree model of Adélie penguin colony distribution based on colony and control plot mean values for Whitney Pt.

		Observed Distribution		
		Absent	**Present**	**Total**
Predicted Distribution	**Absent**	103 (50%)	105 (46.7%)	208 (48.3%)
	Present	103 (50%)	120 (53.3%)	223 (51.7%)
	Total	206	125	431

4.2.3.2 Shirley I

Individual Cell Values

A decision tree with a minimum object size of 200 and nine leaves, predicts Adélie penguin colony distribution with 82.8% accuracy, as measured by cross-validation and 84.5% accuracy, as measured by validation. Slope, elevation, aspect, the difference in snow cover between 1968 and 2005, surface roughness (standard deviation) and wind exposure were used as inputs (App. 3). The confusion matrix (Table 4.12) shows that the model predicts colony presence slightly more accurately than absence. The model predicts absences where the slope is less than 0.22, and present at the highest possible elevations (>29.39 m). Below this altitude, aspect, changes in snow cover and elevation are used to further differentiate colonies from control plots.

Table 4.12: Confusion matrix for validation of the decision tree analysis of Adélie penguin colony distribution based on individual cell values for colonies and control plots on Shirley I.

		Observed Distribution		
		Absent	**Present**	**Total**
Predicted Distribution	**Absent**	539 (83.1%)	93 (14.2%)	632 (42.9%)
	Present	110 (17%)	564 (85.8%)	674 (57.1%)
	Total	649	657	1306

Colony/Control Plot Means

A decision tree with a minimum object size of two and 12 leaves predicts Adélie penguin colony presence or absence with 64.3%, as measured by cross-validation and 68% accuracy, as measured by validation (Table 4.13). Slope, planar curvature, elevation, snow cover 2005, solar radiation and surface roughness (standard deviation) were used as inputs (App. 3). The other static landscape parameters did not improve the model's predictive performance. The model predicts colony absences better than colony presences (Table 4.13). The model predicts absences where the slope is less than 0.13, a gentler slope than that used for splitting the tree derived from individual cell values. Cells in gullies (planar curvature <=-0.02) are predicted to contain penguins as are all sites above 29 m ASL. Below 29 m, planar curvature, snow cover in November 2005, slope, solar radiation, and surface roughness (standard deviation) are used to differentiate the groups (App. 3).

Table 4.13: Confusion matrix for decision tree analysis of Adélie penguin colony distribution based on colony and control plot mean values for Shirley I.

		Observed Distribution		
		Absent	**Present**	**Total**
Predicted Distribution	**Absent**	494 (76.1%)	287 (43.7%)	781 (59.8%)
	Present	155 (23.9%)	370 (56.3%)	525 (40.2%)
	Total	649	657	1306

The results of the discriminant analyses and decision trees suggest that the null hypothesis should be rejected. Static landscape parameters, as calculated in this study, can be used to predict the presence or absence of Adélie penguin nests in a given cell within the study site with up to 84.5% accuracy.

4.3 Snow Accumulation Patterns and Adélie Penguin Colony Population Trends

This section presents the results of the tests for the ability of individual snow accumulation parameters and the discriminant analyses and decision trees derived from these parameters to explain the observed population trends of Adélie penguin colonies within the two study sites as expressed in the following null hypothesis.

H_{NULL} 2 *Interactions between the shape of the land and the weather conditions that drive snow accumulation patterns cannot predict the population trends of Adélie penguin colonies at Shirley I and Whitney Pt.*

The results are presented separately for each study site and for the tests conducted using individual cell values and colony mean values.

4.3.1 Univariate Analyses

4.3.1.1 Whitney Pt

Individual Cell Values

Wilcoxon tests were used to explore differences among the distributions of individual cell values in colonies in five population trend classes at Whitney Pt. These tests show significant differences for all three data layers related to snow accumulation (Table 4.14). The colonies with strong population increases (>150% of the 1959 population) are associated with the thinnest snow cover, while stable colonies are found in areas with thicker snow cover. For the individual cell values, colonies in all population trend classes are associated with areas with minimal changes in the modelled snow cover between 1959 and 2005.

Table 4.14: Wilcoxon tests for differences among individual cell values for colonies in the five population trend classes on Whitney Pt.

Variable	Chi-Square	DF	Prob>ChiSq	Significance
Wind Exposure	234.772	4	<0.0001	Significant
Snow 2005	28.133	4	<0.0001	Significant
Snow Difference	12.638	4	0.013	Significant

Colony Mean Values

Wilcoxon tests for difference among the distributions of colony mean values show no significant differences among the five population trend classes for any of the data layers related to snow accumulation on Whitney Pt (Table 4.15). It is likely that the loss of variance associated with calculating mean values is responsible for the difference from the results of the Wilcoxon tests on the individual cell values for Whitney Pt.

Table 4.15: Wilcoxon tests for differences among colony mean values in the five population trend classes on Whitney Pt.

Variable	Chi-Square	DF	Prob>ChiSq	Significance
Wind Exposure	3.290	4	0.511	Not Significant
Snow Difference	2.556	4	0.635	Not Significant
Snow 2005	2.316	4	0.678	Not Significant

4.3.1.2 Shirley I

Individual Cell Values

Wilcoxon tests on the individual cell values show significant differences among the five population trend classes for all the parameters associated with snow accumulation patterns (Table 4.16). However, the summary statistics (App. 1) and visual inspection of scatterplots show few obvious trends.

Table 4.16: Wilcoxon tests for difference among the distributions of individual cell values in colonies in five population trend classes on Shirley I.

Variable	Chi-Square	DF	Prob>ChiSq	Significance
Wind Exposure	103.146	4	<0.0001	Significant
Snow 2005	302.697	4	<0.0001	Significant
Snow Difference	33.110	4	<0.0001	Significant

Colony Means

Wilcoxon tests on the colony mean values for colonies in five population trend classes show no significant differences for any of the parameters associated with snow accumulation (Table 4.17). It is considered likely that the loss of variance associated with calculating mean values is responsible for the difference from the results of the Wilcoxon tests on the individual cell values for Shirley I.

Table 4.17: Wilcoxon tests for difference among colony mean values in five population trend classes on Shirley I.

Variable	ChiSquare	DF	Prob>ChiSq	Significance
Snow 2005	8.366	4	0.079	Not Significant
Snow Difference	3.013	4	0.556	Not Significant
Wind Exposure	1.731	4	0.785	Not Significant

4.3.2 Discriminant Analyses

4.3.2.1 Whitney Pt

Individual Cell Values

Cross-validation shows that a discriminant analysis model using wind exposure and snow cover in 2005 as inputs predicted the colony population trend class with 48.6% accuracy (App.2). The change in snow cover between 1959 and 2005 did not improve the predictive power of the model. Validation shows the overall accuracy is 48.3%. The confusion matrix (Table 4.18) shows that the model predicts stable colonies (class 3) with the highest accuracy (58.1%) and moderately declining colonies with the lowest accuracy (0%). However, this class has only one data point in the test set, so this is not statistically significant. The observed colony population trends are shown in Fig. 4.19 and the predicted trends in Fig. 4.23. The predictive map (Fig. 4.20) shows a mix of predicted trend classes within each colony, rather than clear spatial trends. It is likely that the results are biased by the large number of data points within the strongly increasing class 5 (n=847) compared with the other classes (combined n=191). The inequality of sample sizes is the result of different population trends at Shirley I and Whitney Pt, and makes the Whitney Pt results difficult to interpret.

Table 4.18: Confusion matrix for validation of the discriminant analysis model for population trend predictions for Whitney Pt based on individual cell values. The percentage value in brackets represents the percentage of the observed trend class total.

		Observed Trend Class					
		Class 1	Class 2	Class 3	Class 4	Class 5	Total
Predicted Trend Class	Class 1 (Strong Dec	3 (75%)	0	0	0	20 (11.9%)	23 (11%)
	Class 2 (Mod Dec.)	1 (25%)	0	1 (3%)	0	40 (23.8%)	42 (20.1%)
	Class 3 (Stable)	0	0	18 (58.1%)	4 (80%)	25 (14.9%)	47 (22.5%)
	Class 4 (Mod Inc.)	0	0	9 (29%)	1 (20%)	4 (2.4%)	14 (6.7%)
	Class 5 (Strong Inc	0	1 (100%)	3 (9.7%)	0	79 (47%)	83 (39.7%)
	Total	4	1	31	5	168	209

Colony Mean Values

A discriminant analysis model was constructed using the modelled snow cover in 2005, change in snow cover between 1959 and 2005 and wind exposure as inputs (App. 2). Cross-validation shows that this model correctly predicts population trends in 42.4% of cells within colonies on Whitney Pt. Validation with the test set of individual cell values shows that the model has an overall accuracy of 38.8% (Fig. 4.21). The model performs best in predicting strongly increasing colonies (class 5: 41.1%) as shown in the confusion matrix (Table 4.19). This model is even more likely to be affected by the disparity in sample sizes than the model based on individual cell values.

Table 4.19: Confusion matrix for validation of the discriminant analysis model predicting population trends for Adélie penguin colonies at Whitney Pt based on colony mean values for snow accumulation parameters.

		Observed Trend Class					
		Class 1	Class 2	Class 3	Class 4	Class 5	Total
Predicted Trend Class	Class 1 (Strong Dec.)	1 (25%)	0	2 (6.5%)	0	25 (14.9%)	28 (13.4%)
	Class 2 (Mod Dec.)	3 (75%)	0	1 (3.2%)	0	45 (26.8%)	49 (23.4%)
	Class 3 (Stable)	0	0	10 (32.3%)	2 (40%)	5 (3%)	17 (8.1%)
	Class 4 (Mod Inc.)	0	0	7 (22.6%)	1 (20%)	24 (14.3%)	32 (15.3%)
	Class 5 (Strong Inc.)	0	1 (100%)	11 (35.5%)	2 (40%)	69 (41.1%)	83 (39.7%)
	Total	4	1	31	5	168	209

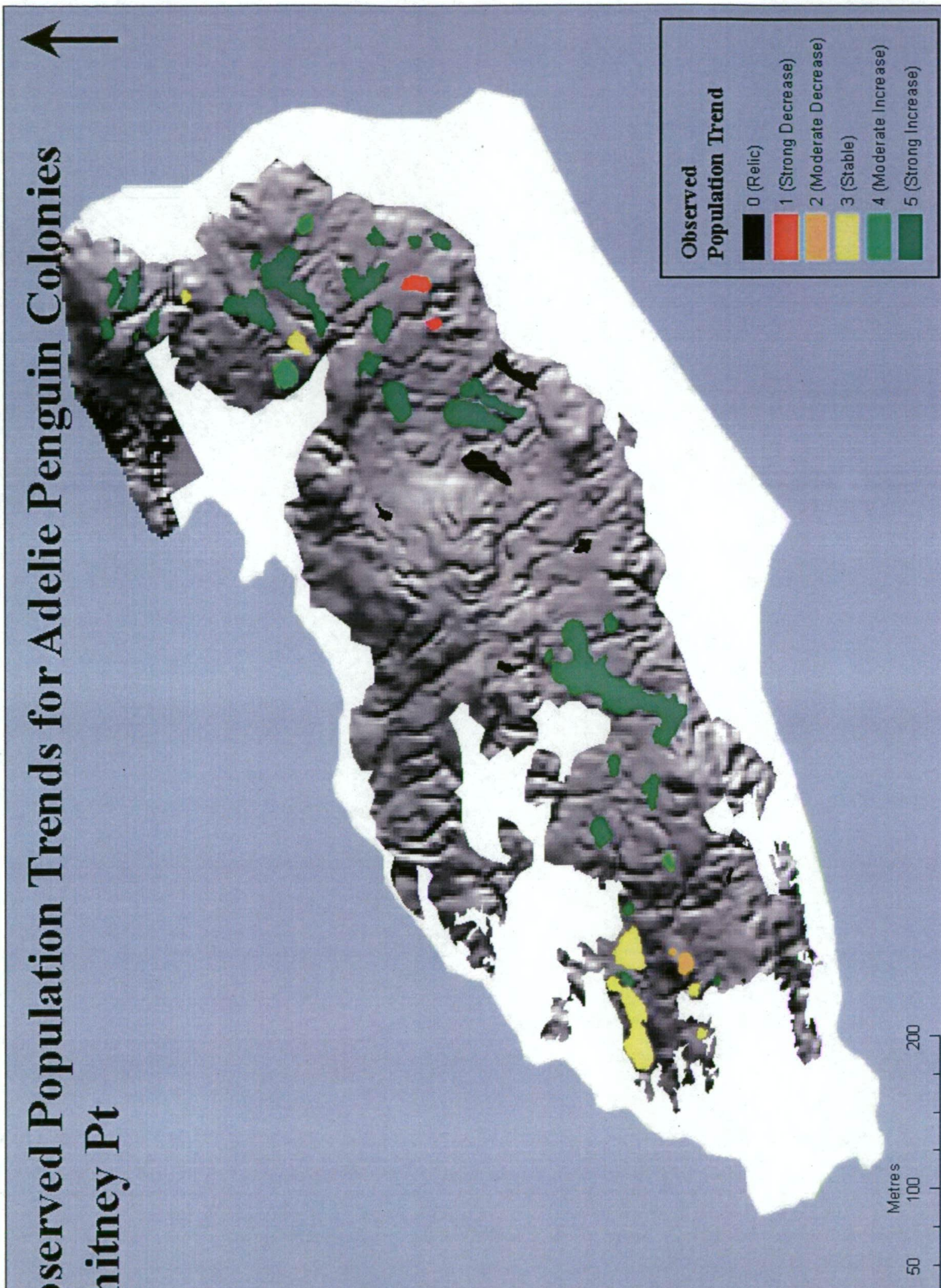


Fig. 4.19: Observed population trend classes of Adélie penguin colonies at Whitney Pt.

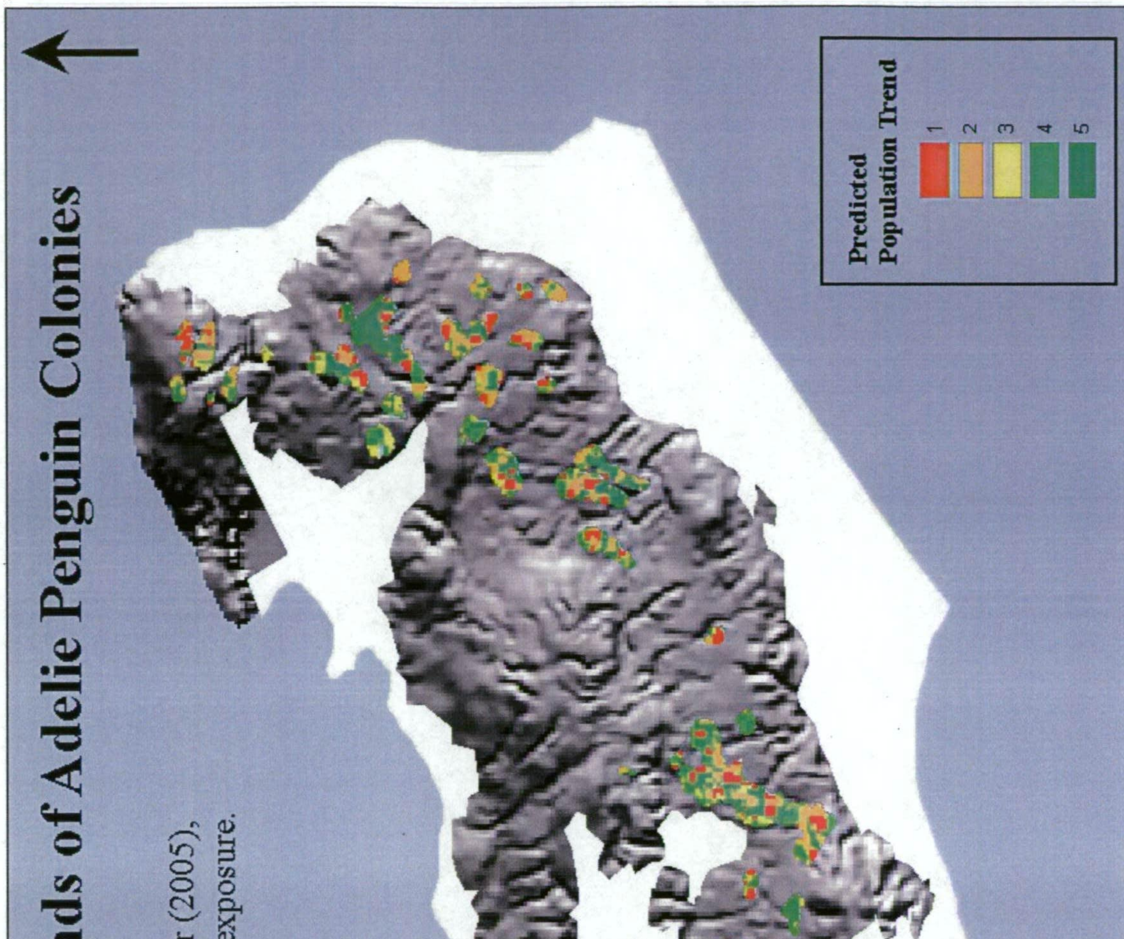


Fig. 4.20: Predicted population trend classes of Adélie penguin colonies at Whitney Pt, based on discriminant analysis of individual cell values for snow accumulation parameters.

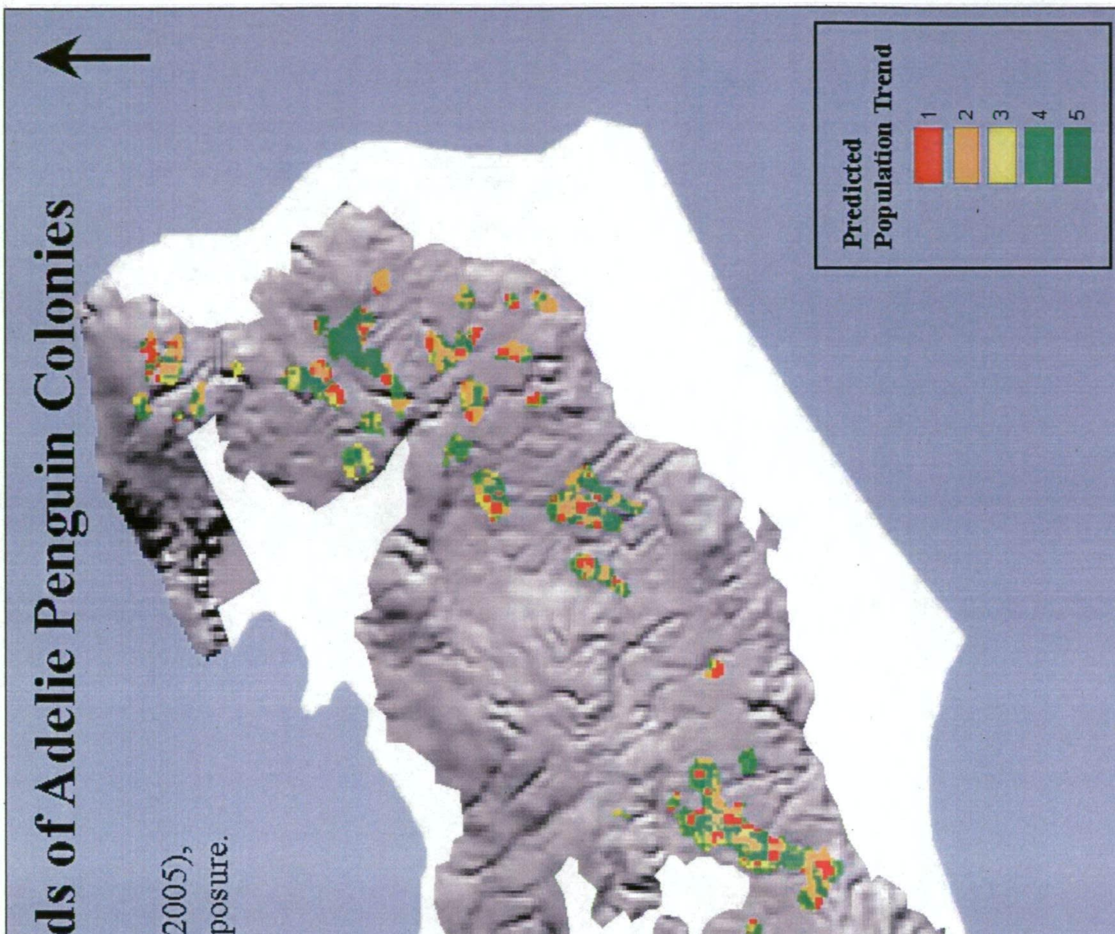


Fig. 4.21: Predicted population trend classes of Adélie penguin colonies at Whitney Pt, based on discriminant analysis of colony mean values for snow accumulation parameters.

4.3.2.2 Shirley I

Individual Cell Values

A discriminant analysis model was constructed using wind exposure, the modelled snow cover in November 2005 and the change in snow cover between 1959 and 2005 (App. 2). Cross-validation suggests this model predicted colony population trends with 31.4% accuracy and validation shows it has an overall accuracy of 25.6%. The model most accurately predicts moderately decreasing (class 2: 63.6%) and stable (class 3: 54.7%) colonies (Table 4.20). The observed population trends for Shirley I are presented in Fig. 4.22 and the predicted population trends in Fig. 4.23.

Table 4.20: Confusion matrix for the discriminant analysis model predicting population trends based on the individual cell values for snow accumulation parameters for Shirley I.

		Observed Trend Class					
		Class 1	Class 2	Class 3	Class 4	Class 5	Total
Predicted Trend Class	Class 1 (Strong Dec.)	2 (4.9%)	1 (9.1%)	7 (5%)	9 (15.8%)	18 (10.1%)	37 (8.7%)
	Class 2 (Mod Dec.)	18 (43.9%)	7 (63.6%)	30 (21.6%)	7 (12.3%)	32 (18%)	94 (22.1%)
	Class 3 (Stable)	15 (36.6%)	2 (18.2%)	76 (54.7%)	28 (49.1%)	88 (49.4%)	209 (49.1%)
	Class 4 (Mod Inc.)	3 (7.3%)	1 (9.1%)	16 (11.5%)	7 (12.3%)	23 (12.9%)	50 (11.7%)
	Class 5 (Strong Inc.)	3 (7.3%)	0	10 (7.2%)	6 (10.5%)	17 (9.6%)	36 (8.5%)
	Total	41	11	139	57	178	426

Colony/Control Plot Means

A discriminant analysis model was constructed using the colony mean values for wind exposure, snow cover in November 2005 and the difference in snow cover between 1959 and 2005 (App. 2). Cross-validation shows that this model has an accuracy of 31.8% and validation with the test set of individual cell values shows it has an overall accuracy of 27.9% (Fig. 4.24). The confusion matrix (Table 4.21) shows that the model most accurately predicts moderately decreasing colonies (class 2: 62.5%)

Table 4.21: Confusion matrix for validation of the discriminant analysis model predicting colony population trends based on colony mean values for snow accumulation parameters for Shirley I.

		Observed Trend Class					
		Class 1	Class 2	Class 3	Class 4	Class 5	Total
Predicted Trend Class	Class 1 (Strong Dec.	3 (7.3%)	2 (18.2%)	13 (9.5%)	9 (15.8%)	23 (13%)	50 (11.8%)
	Class 2 (Mod Dec.)	9 (22%)	5 (45.5%)	20 (14.6%)	4 (7%)	18 (10.2%)	56 (13.2%)
	Class 3 (Stable)	17 (41.5%)	2 (18.2%)	84 (61.3%)	34 (59.7%)	93 (52.5%)	230 (54.4%)
	Class 4 (Mod Inc.)	8 (19.5%)	2 (18.2%)	13 (9.5%)	9 (15.8%)	26 (14.7%)	58 (13.7%)
	Class 5 (Strong Inc.)	4 (9.8%)	0	7 (5.1%)	1 (1.8%)	17 (9.6%)	29 (6.9%)
	Total	41	11	137	57	177	423

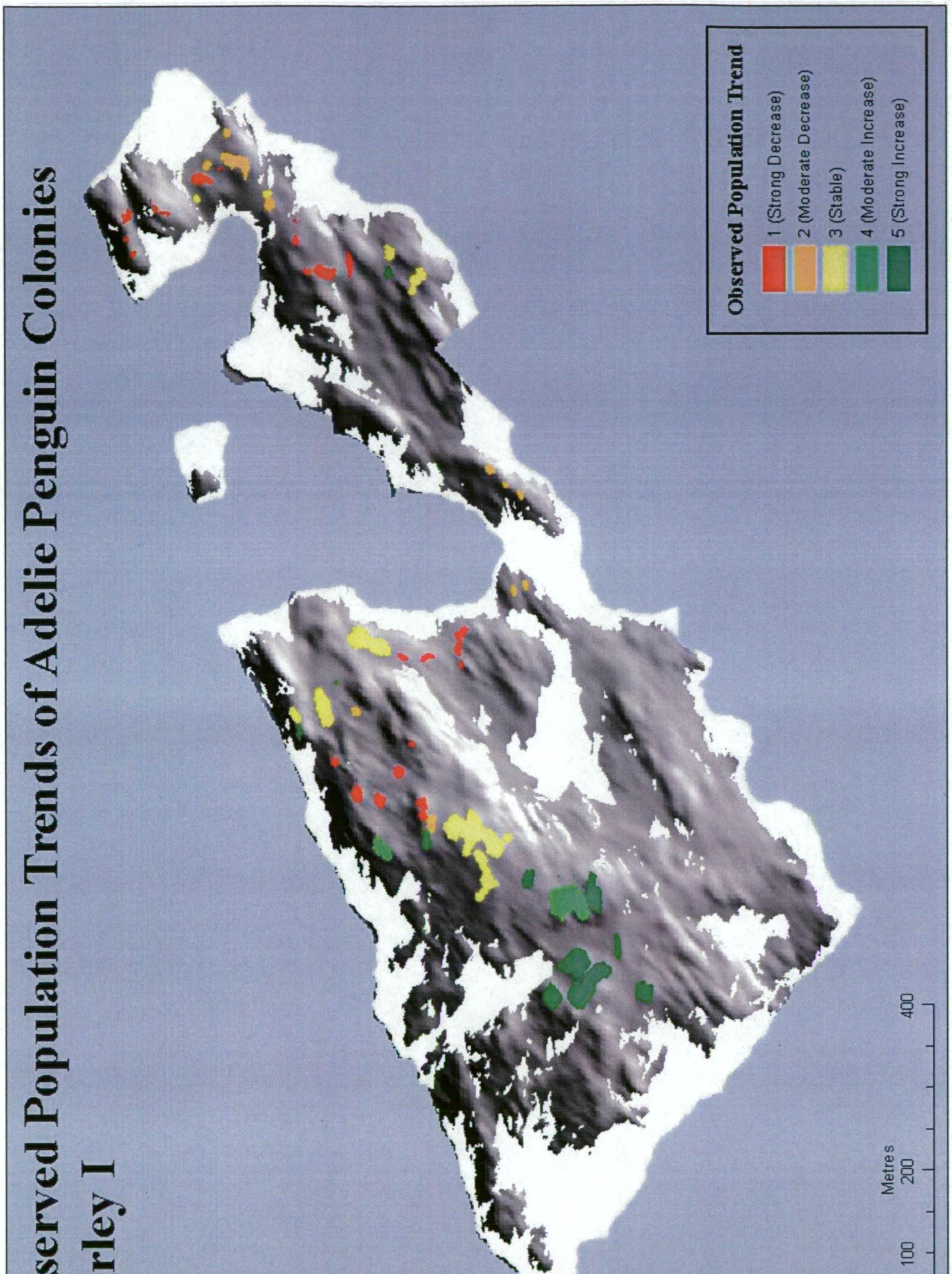


Fig. 4.22 Observed population trend classes of Adélie penguin colonies on Shirley I.

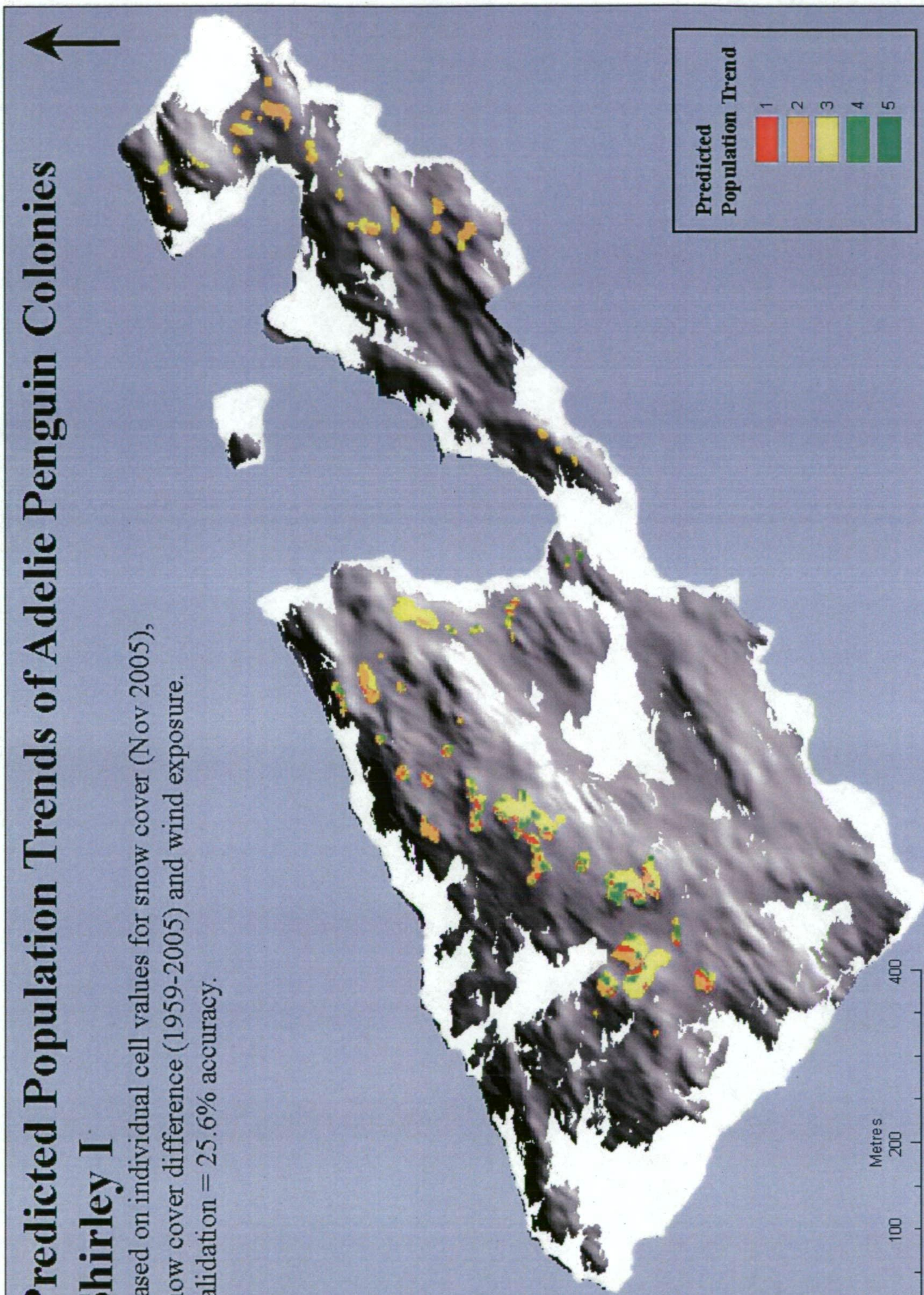


Fig. 4.23: Predicted population trend classes of Adélie penguin colonies on Shirley I based on discriminant analysis of individual cell values for snow accumulation parameters.

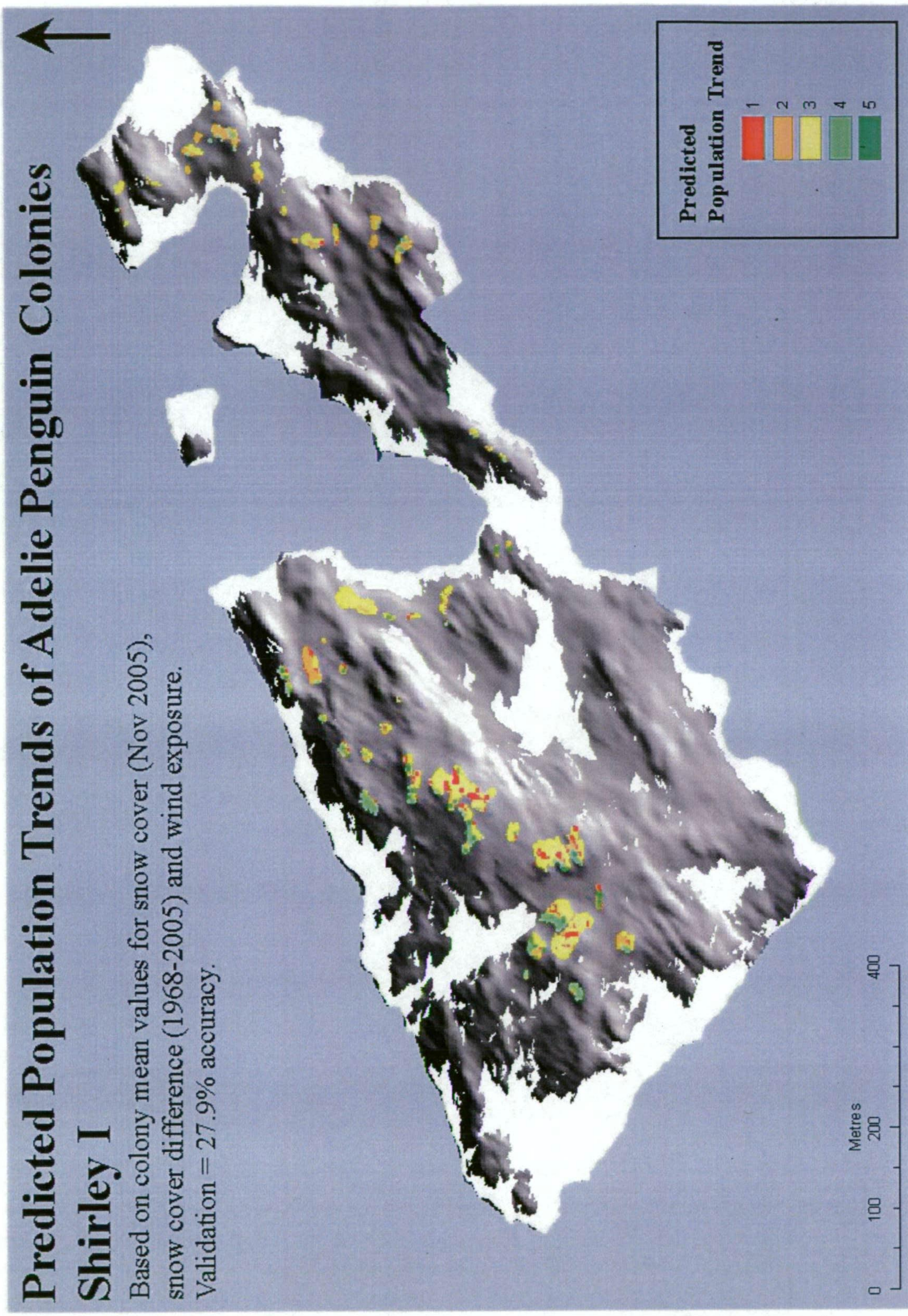


Fig. 4.24: Predicted population trend classes of Adélie penguin colonies on Shirley I, based on discriminant analysis of colony mean values for snow accumulation parameters.

4.3.3 Decision Tree Analysis

4.3.3.1 Whitney Pt

Individual Cell Values

A decision tree analysis, with a minimum object size of 50, and with four leaves was constructed using wind exposure and snow cover in November, 2005 (App. 3). The difference in snow cover between 1959 and 2005 did not improve the predictive power of the model. Cross-validation shows that the model predicts population trends with 83.5% accuracy, while validation shows that the model is 83.7% accurate. The confusion matrix (Table 4.22) shows that the model is most accurate in predicting strongly increasing colonies (class 5: 98.8%). In contrast the model fails to accurately predict any of the decreasing colonies (classes 1 and 2), although the sample size of these in the test set is too small to draw conclusions about this aspect of the model. The tree predicts that those colonies with wind exposure ratings greater than 39 or with snow cover in 2005 of less than 0.07 have strongly increasing population trends. In areas with deeper snow cover, those with wind exposure greater than 5 are also predicted to be increasing, while colonies in other areas are listed as stable (App. 3).

Table 4.22: Confusion matrix of the performance of the decision tree analysis of colony population trends based on individual cell values for snow accumulation parameters.

		Observed Trend Class					
		Class 1	Class 2	Class 3	Class 4	Class 5	Total
Predicted Trend Class	Class 1 (Strong Dec	0	0	0	0	0	0
	Class 2 (Mod Dec.)	0	0	0	0	0	0
	Class 3 (Stable)	0	0	9 (29%)	1 (20%)	2 (1.2%)	12 (5.7%)
	Class 4 (Mod Inc.)	0	0	0	0	0	0
	Class 5 (Strong Inc.)	4 (100%)	1 (100%)	22 (71%)	4 (80%)	166 (98.8%)	197 (94.3%)
	Total	4	1	31	5	168	209

Colony Mean Values

A decision tree analysis with a minimum object size of two and with two leaves, was constructed using the colony mean values for snow cover in November, 2005 (App. 2). Mean values for wind exposure and the difference in snow cover between 1959 and 2005 did not improve the predictive power of the model. Cross-validation of the model suggests that it predicts colony trends with

63.6% accuracy and validation with the test set of individual cell values shows it has an overall accuracy of 81.2% (Table 4.23). The model most accurately predicts strongly increasing colonies (class 5: 95.2%) and fails to accurately predict the trends of any declining or moderately increasing colonies (classes 1, 2 and 4). In this model, those cells with snow cover less than or equal to 1.14 are predicted to be increasing strongly, while all others are predicted as stable.

Table 4.23: Confusion matrix for validation of the decision tree analysis for colony trend predictions based on the colony mean values for snow accumulation parameters.

		Observed Trend Class					
		Class 1	**Class 2**	**Class 3**	**Class 4**	**Class 5**	**Total**
Predicted Trend Class	**Class 1 (Strong Dec.)**	0	0	0	0	0	0
	Class 2 (Mod Dec.)	0	0	0	0	0	0
	Class 3 (Stable)	0	0	10 (32.3%)	1 (20%)	8 (4.8%)	19 (9.1%)
	Class 4 (Mod Inc.)	0	0	0	0	0	0
	Class 5 (Strong Inc.)	4 (100%)	1 (100%)	21 (67.7%)	4 (80%)	160 (95.2%)	190 (90.9%)
	Total	4	1	31	5	168	209

4.3.3.2 Shirley I

Individual Cell Values

A decision tree with a minimum object size of 50 and with 18 leaves was constructed using the November 2005 snow cover, wind exposure and the change in snow cover between 1959 and 2005. Cross-validation shows that the model correctly predicts 53.8% of colony trends. Validation shows that it has an overall accuracy of 57.8%. The confusion matrix (Table 4.24) shows that the model most accurately predicts strongly increasing colonies (class 5: 73%). This model predicts population trends based on a complex interaction between the three variables, with snow cover in November 2005 and wind exposure explaining most of the variance (App. 3).

Table 4.24: Confusion matrix for the decision tree prediction model of penguin colony population trends for Shirley I based on individual cell values for snow accumulation parameters.

		Observed Trend Class					
		Class 1	Class 2	Class 3	Class 4	Class 5	Total
Predicted Trend Class	Class 1 (Strong Dec	12 (29.3%)	0	5 (3.6%)	0	2 (1.1%)	19 (4.5%)
	Class 2 (Mod Dec.)	0	7 (63.6%)	5 (3.6%)	0	1 (0.1%)	13 (3.1%)
	Class 3 (Stable)	13 (31.7%)	0	78 (56.1%)	10 (17.5%)	34 (19.1%)	135 (31.7%)
	Class 4 (Mod Inc.)	2 (4.9%)	2 (18.2%)	9 (6.5%)	19 (33.3%)	11 (6.2%)	43 (10.1%)
	Class 5 (Strong Inc.	14 (34.2%)	2 (18.2%)	42 (30.2%)	28 (49.1%)	### (73%)	216 (50.7%)
	Total	41	11	139	57	178	426

Colony Means

A decision tree with a minimum object size of two and with 13 leaves was constructed using the November 2005 snow cover, the change in snow cover between 1959 and 2005 and wind exposure (App. 3). Cross-validation shows that this model has an overall accuracy of 27.3%. Validation with the test set of individual cell values shows that it correctly predicts 41.6% of colony trends. The confusion matrix (Table 4.25) shows that the model most accurately predicts strongly increasing colonies (class 5: 65%) and fails to correctly predict any cells in colonies with moderately increasing populations (class 4). As with the previous model, snow cover in 2005 was used to make the first splits in the data, followed by wind exposure, and the model uses a complex mix of the three variables to predict population trends.

Table 4.25: Confusion matrix for the decision tree prediction model of penguin colony population trends for Shirley I based on colony mean values for snow accumulation parameters.

		Observed Trend Class					
		Class 1	Class 2	Class 3	Class 4	Class 5	Total
Predicted Trend Class	Class 1 (Strong Dec.)	25 (618%)	2 (18.2%)	28 (20.4%)	17 (29.8%)	56 (31.6%)	128 (30.1%)
	Class 2 (Mod Dec.)	1 (2.4%)	6 (54.6%)	4 (2.9%)	0	7 (4%)	18 (4.2%)
	Class 3 (Stable)	6 (14.6%)	2 (18.2%)	31 (22.6%)	0	0	39 (9.2%)
	Class 4 (Mod Inc.)	0	0	0	0	0	0
	Class 5 (Strong Inc.)	9 (22%)	1 (9.1%)	76 (55.5%)	40 (70.2%)	115 (65%)	241 (56.6%)
	Total	41	11	139	57	178	426

The results of the discriminant analyses and decision trees show that the null hypothesis should be rejected. Purely random classifications could be expected to accurately predict 20% of values across five classes. The predictive models derived from snow accumulation parameters produce significant results for both sites, when based on both individual cell and colony mean values. The results are stronger for Whitney Pt than for Shirley I.

4.4 Proximity to human activities and population trends of Adélie penguin colonies

This section presents the results of the tests for the ability of individual parameters associated with proximity or exposure to human activities, and the discriminant analyses and decision trees derived from these parameters to explain the observed population trends of Adélie penguin colonies within the two study sites as expressed in the following null hypothesis.

H_{NULL} 3 Proximity and exposure to human activities associated with Casey cannot predict the population trends of Adélie penguin colonies at Shirley I and Whitney Pt.

The results are presented separately for each study site and for the tests conducted using individual cell values and colony mean values.

4.4.1 Univariate Analyses

4.4.1.1 Whitney Pt

Individual Cell Values

A Wilcoxon test on individual cell values for colonies in five population trend classes demonstrates a significant difference among the classes for the distance from Casey (Table 4.26). Exploration of the scatterplots and histograms showed that strongly increasing colonies occurred at all distances from Casey, while stable (class 3) and moderately decreasing (class 2) colonies were clustered closest to Casey, and strongly decreasing (class 1) and moderately increasing (class 4) colonies were found near the middle of the study site.

Table 4.26: Wilcoxon test for difference between individual cell values for colonies in five population trend classes.

Variable	Chi-Square	DF	Prob>ChiSq	Significance
Casey Distance	221.307	4	<0.0001	Significant

Colony Mean Values

A Wilcoxon test on colony mean values in five population trend classes shows no significant difference among the population trend classes on Whitney Pt for the distance from Casey (Table 4.27). It is likely that the decreased variance in the colony mean values reduced the significance of the differences among the five classes.

Table 4.27: Wilcoxon test for difference between colony mean values in five population trend classes for Whitney Pt.

Variable	Chi-Square	DF	Prob>ChiSq	Significance
Casey Distanc	5.211	4	0.266	N.S.

4.4.1.2 Shirley I

Individual Cell Values

Wilcoxon tests shows significant differences among the five population trend classes for all the parameters associated with proximity to human activity (Table 4.28). All population trend classes are clustered around the median values for wind exposure. Almost all the strongly increasing colonies are found at the farthest difference from Casey and from the sea-ice crossing point. Most of the stable colonies occur at medium to long distances from Casey and the sea-ice crossing point. Moderately decreasing colonies are clustered very close to Casey, and close to the sea-ice crossing point. Strongly decreasing colonies are bimodal, occurring close to Casey and the sea-ice crossing point, and also at moderately long distances from both.

Table 4.28: Wilcoxon tests for significant difference among individual cell values for the five population trend classes for Adélie penguin colonies on Shirley I.

Variable	ChiSquare	DF	Prob>ChiSq	Significance
Casey Distance	1466.67	4	<0.0001	Significant
Sea-ice Crossing Point Distance	1446.62	4	<0.0001	Significant
Wind Exposure	103.15	4	<0.0001	Significant

Colony Mean Values

Similarly to the individual cell tests, Wilcoxon tests show significant differences among the five population trend classes for all the parameters associated with proximity to human activities, for colony mean values (Table 4.29). They show no significant differences among colonies for wind exposure.

Table 4.29: Wilcoxon tests for difference between colony mean values for the five population trend classes for Adélie penguin colonies on Shirley I.

Variable	Chi-Square	DF	Prob>ChiSq	Significance
Casey Distance	16.990	4	0.002	Significant
Sea-ice Crossing Point Distan	13.868	4	0.008	Significant
Wind Exposure	1.731	4	0.785	N.S.

4.4.2 Discriminant Analyses

4.4.2.1 Whitney Pt

Individual Cell Values

A discriminant analysis model to predict population trends was constructed using the distance from Casey (App. 2). Cross-validation suggests this model has an accuracy of 28.1%, while validation shows the overall accuracy is 26.8% (Fig. 4.25). The confusion matrix (Table 4.30) shows the model predicts colony population trends with a high degree of accuracy for classes 1-4 (ranging from 93.6% to 100%). The small number of data-points in the test set for classes 1, 2 and 4 reduces the reliability of this result. The model accurately predicts 10.1% of the strongly increasing colonies (class 5). This class represents 80.38% of the dataset, and so has a large effect on the overall accuracy.

Table 4.30: Confusion matrix of the discriminant analysis model predicting Adélie penguin colony population trends at Whitney Pt based on the individual cells' distance from Casey.

		Observed Trend Class					
		Class 1	Class 2	Class 3	Class 4	Class 5	Total
Predicted Trend Class	Class 1 (Strong Dec	4 (100%)	0	0	0	31 (18.5%)	35 (16.8%)
	Class 2 (Mod Dec.)	0	1 (100%)	0	0	19 (11.3%)	20 (9.6%)
	Class 3 (Stable)	0	0	29 (93.6%)	0	57 (33.9%)	86 (41.2%)
	Class 4 (Mod Inc.)	0	0	2 (6.5%)	5 (100%)	44 (26.2%)	51 (24.4%)
	Class 5 (Strong Inc.)	0	0	0	0	17 (10.1%)	17 (8.1%)
	Total	4	1	31	5	168	209

Colony Mean Values

A discriminant analysis model predicting Adélie penguin population trends based on the colony mean values for distance from Casey (App. 2) has an overall accuracy of 33.3%, as measured by cross-validation. Validation with the test set of individual cell values shows the model has an overall accuracy 17.2% (Fig. 4.26). The confusion matrix (Table 4.31) shows the model most accurately predicts colony trends for moderately decreasing colonies (class 2: 100%) and moderately increasing colonies (class 4: 100%). However the low number of data points in these classes in the test set reduces the reliability of these results. The model fails to correctly predict any cells in stable colonies.

Table 4.31: Confusion matrix for cross-validation of the discriminant analysis model predicting population trends based on colony mean values for the distance of Whitney Pt colonies from Casey.

		Observed Trend Class					
		Class 1	Class 2	Class 3	Class 4	Class 5	Total
Predicted Trend Class	Class 1 (Strong Dec.)	3 (75%)	0	0	0	8 (4.8%)	11 (5.3%)
	Class 2 (Mod Dec.)	0	1 (100%)	29 (93.6%)	0	47 (28%)	77 (36.8%)
	Class 3 (Stable)	0	0	0	0	44 (26.2%)	44 (21.1%)
	Class 4 (Mod Inc.)	0	0	2 (6.6%)	5 (100%)	42 (25%)	49 (23.4%)
	Class 5 (Strong Inc.)	1 (25%)	0	0	0	27 (16.1%)	28 (13.4%)
	Total	4	1	31	5	168	209

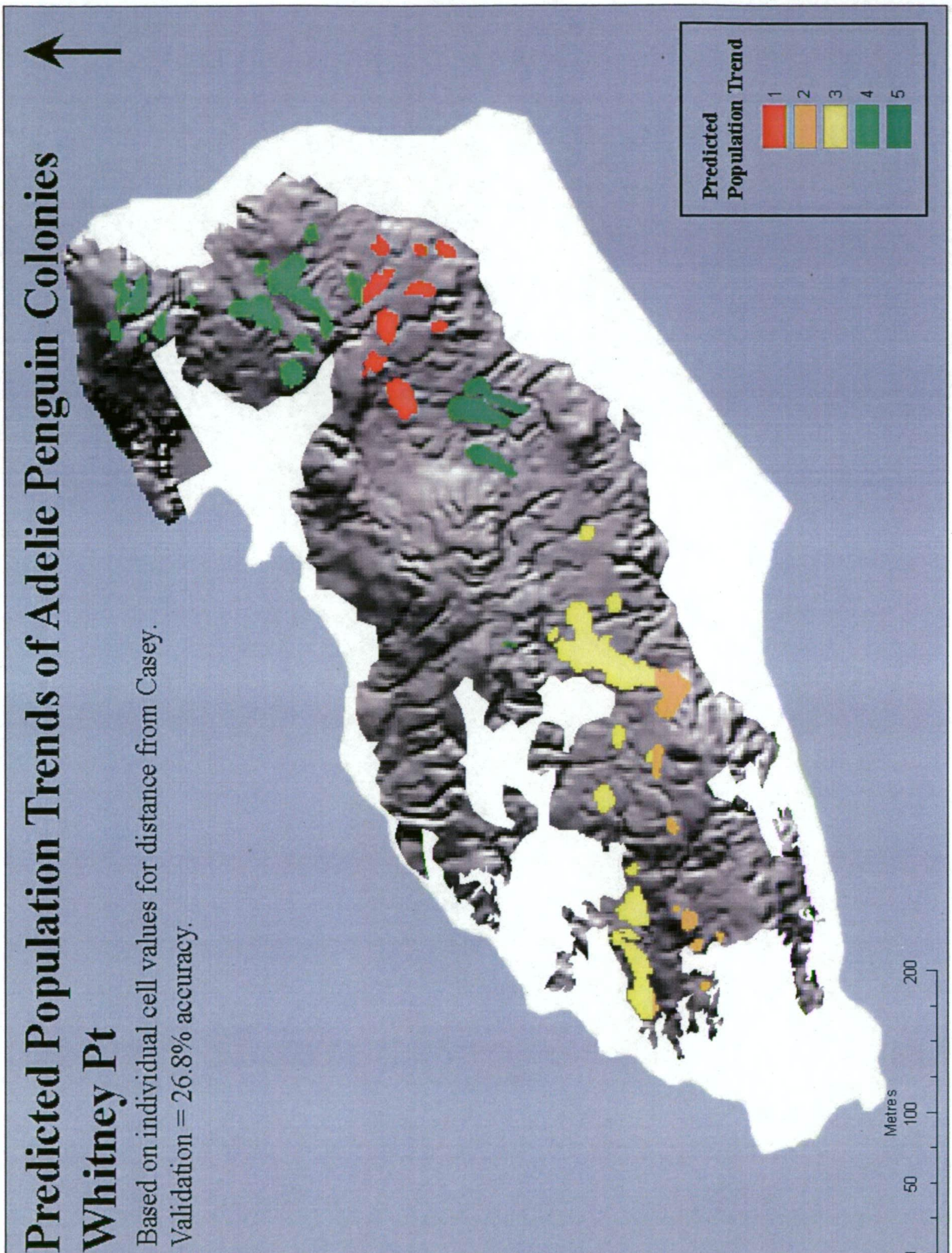


Fig. 4.25: Predicted population trends for Adélie penguin colonies at Whitney Pt, based on discriminant analysis of individual cell values for distance from Casey.

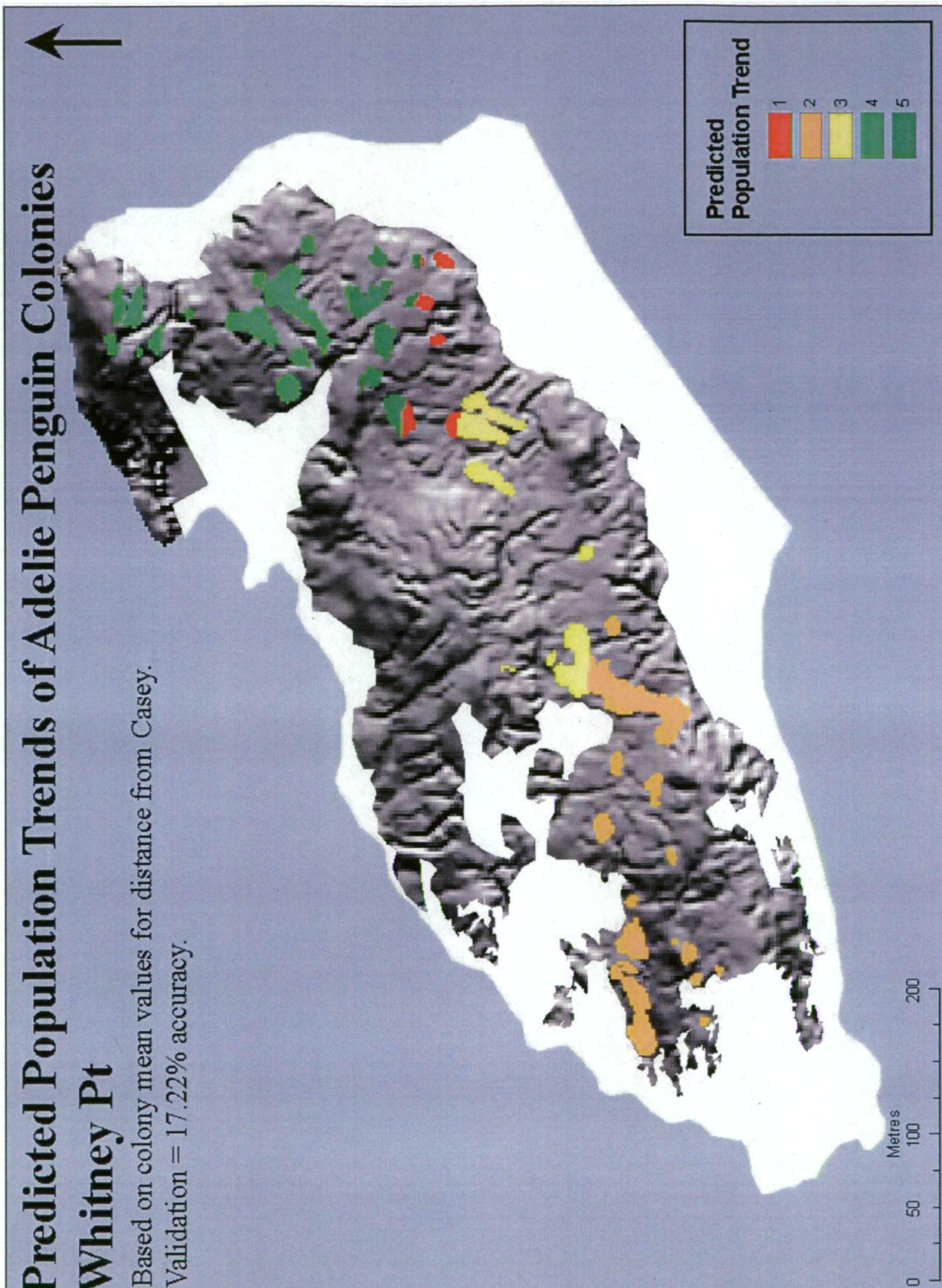


Fig. 4.26: Predicted population trends of Adélie penguin colonies at Whitney Pt, based on discriminant analysis of colony mean values for distance from Casey.

4.4.2.2 Shirley I

Individual Cell Values

A discriminant analysis model was constructed using the individual cell values for distance from Casey, distance from the sea-ice crossing point and the wind exposure (App. 2). The wind exposure layer was included as a surrogate for exposure to any potential airborne emissions from Casey, as Shirley I is directly downwind from Casey. However, wind exposure contributed only a small amount to increasing the model's predictive power. Cross-validation suggests that this model predicts population trends with 72.7% accuracy. Validation shows that the accuracy is 72.1% (Fig. 4.27). Distance from Casey is the most important factor in predicting observed colony population trends. A discriminant analysis model derived from this one parameter, correctly predicts trends for 49.4% of the cells (as shown by cross-validation), compared with 26.8% for Whitney Pt for the same model. The confusion matrix (Table 4.32) shows that the Shirley I model using all three parameters most accurately predicts the trends for moderately increasing colonies (class 4: 100%) and strongly increasing colonies (class 5: 75.8%).

Table 4.32: Confusion matrix for the discriminant analysis model predicting population trends of Shirley I colonies based on individual cells' distance from Casey, distance from the sea-ice crossing point and exposure to prevailing winds.

		Observed Trend Class					
		Class 1	**Class 2**	**Class 3**	**Class 4**	**Class 5**	**Total**
Predicted Trend Class	**Class 1 (Strong Dec.**	11 (26.8%)	0	3 (2.2%)	0	4 (2.3%)	18 (4.2%)
	Class 2 (Mod Dec.)	9 (22%)	7 (63.6%)	1 (0.7%)	0	0	17 (4%)
	Class 3 (Stable)	9 (22%)	2 (18.2%)	97 (69.8%)	0	0	108 (25.4%)
	Class 4 (Mod Inc.)	7 (17.1%)	2 (18.2%)	38 (27.3%)	57 (100%)	39 (21.9%)	143 (33.6%)
	Class 5 (Strong Inc.)	5 (12.2%)	0	0	0	135 (75.8%)	140 (32.9%)
	Total	41	11	139	57	178	426

Colony Mean Values

The discriminant analysis model of Adélie penguin colony population trends based on proximity to human activities based on colony mean values was improved by the distance from the sea-ice crossing point and the wind-exposure (App. 2). Distance from Casey decreased the predictive power of the model and was excluded from the final model. Cross-validation shows that the model correctly predicts the population trend in 36.4% of test cells, while validation with the test set of individual cell values shows that the overall accuracy is 40.6% (Fig. 4.28). The model most accurately predicts strongly increasing colonies (class 5: 75.3%) and least accurately predicts strongly decreasing (class 1: 7.3%) colonies (Table 4.33).

Table 4.33: Confusion matrix for validation of the discriminant analysis model predicting population trends of Shirley I colonies based on the colony mean values for distance from the sea-ice crossing point and exposure to prevailing winds.

		Observed Trend Class					
		Class 1	**Class 2**	**Class 3**	**Class 4**	**Class 5**	**Total**
Predicted Trend Class	**Class 1 (Strong Dec.)**	3 (7.3%)	0	39 (28.1%)	0	0	42 (9.9%)
	Class 2 (Mod Dec.)	13 (31.7%)	7 (63.6%)	5 (3.6%)	0	3 (1.7%)	28 (6.6%)
	Class 3 (Stable)	11 (26.8%)	0	19 (13.7%)	0	0	30 (7%)
	Class 4 (Mod Inc.)	10 (24.4%)	0	16 (11.5%)	10 (17.5%)	41 (23%)	77 (18.1%)
	Class 5 (Strong Inc.)	4 (9.8%)	4 (36.4%)	60 (43.2%)	47 (82.5%)	134 (75.3%)	249 (58.5%)
	Total	41	11	139	57	178	426

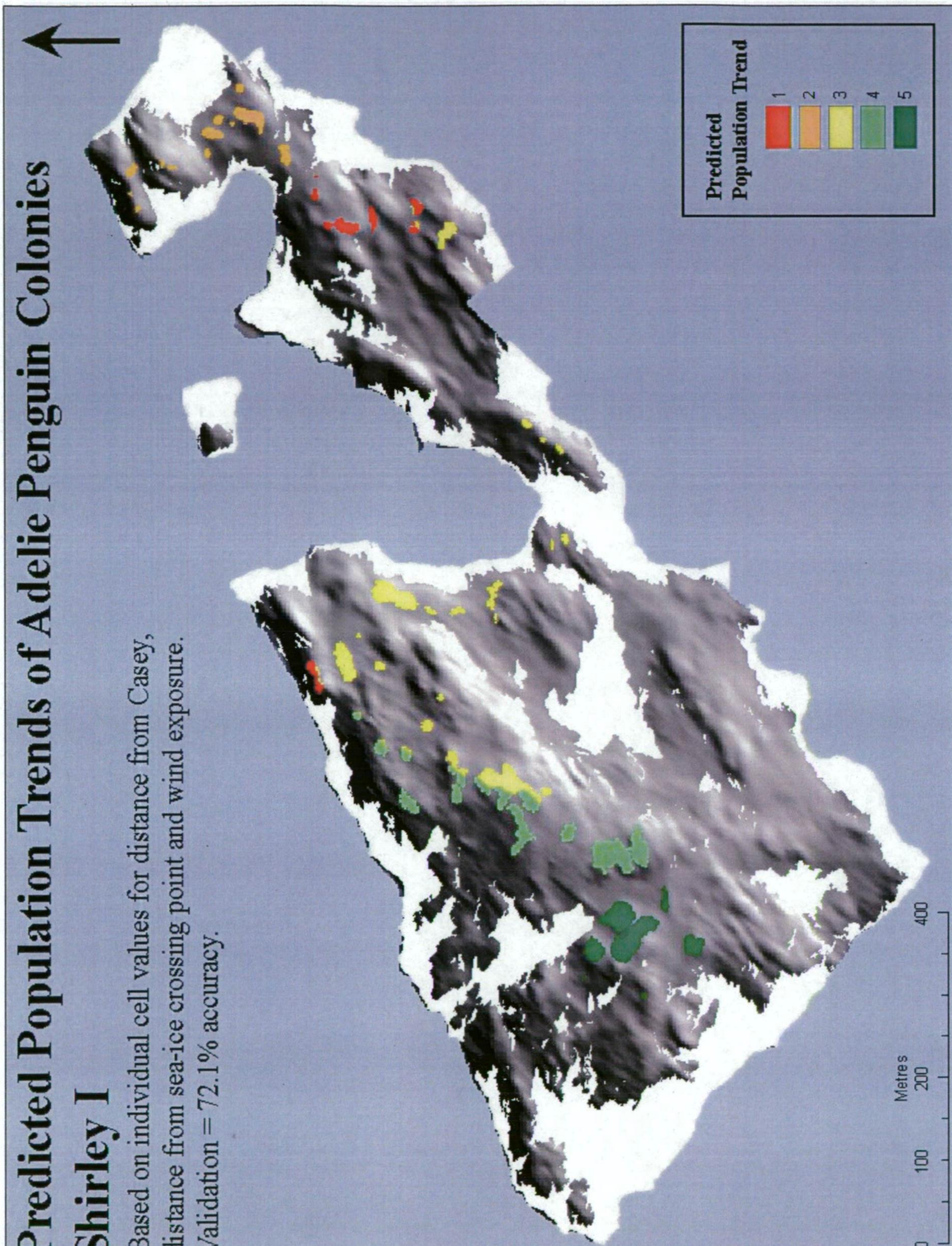


Fig. 4.27: Predicted population trends for Adélie penguin colonies on Shirley I, based on discriminant analysis of individual cell values for proximity to human activities parameters.

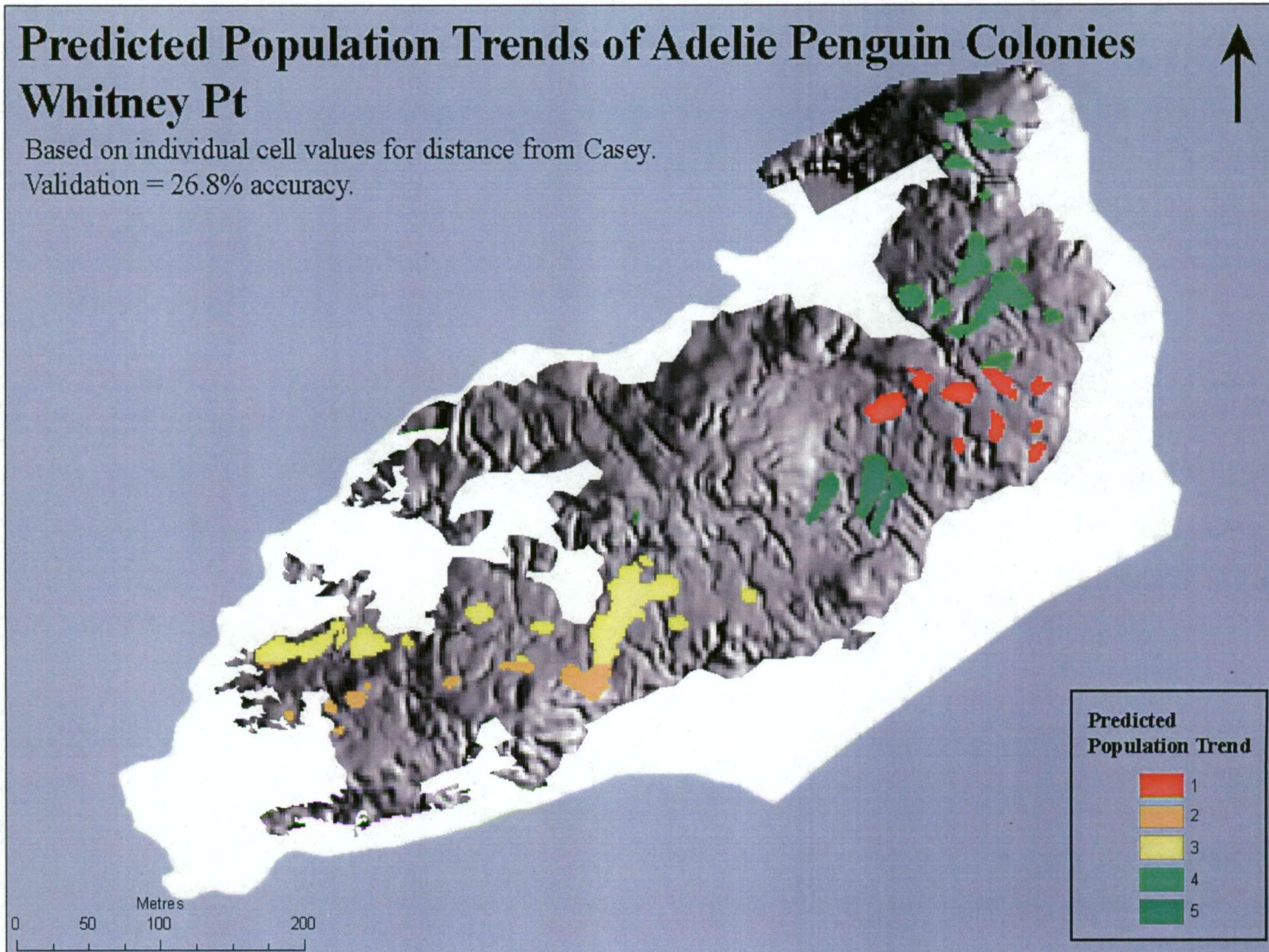


Fig. 4.25: Predicted population trends for Adélie penguin colonies at Whitney Pt, based on discriminant analysis of individual cell values for distance from Casey.

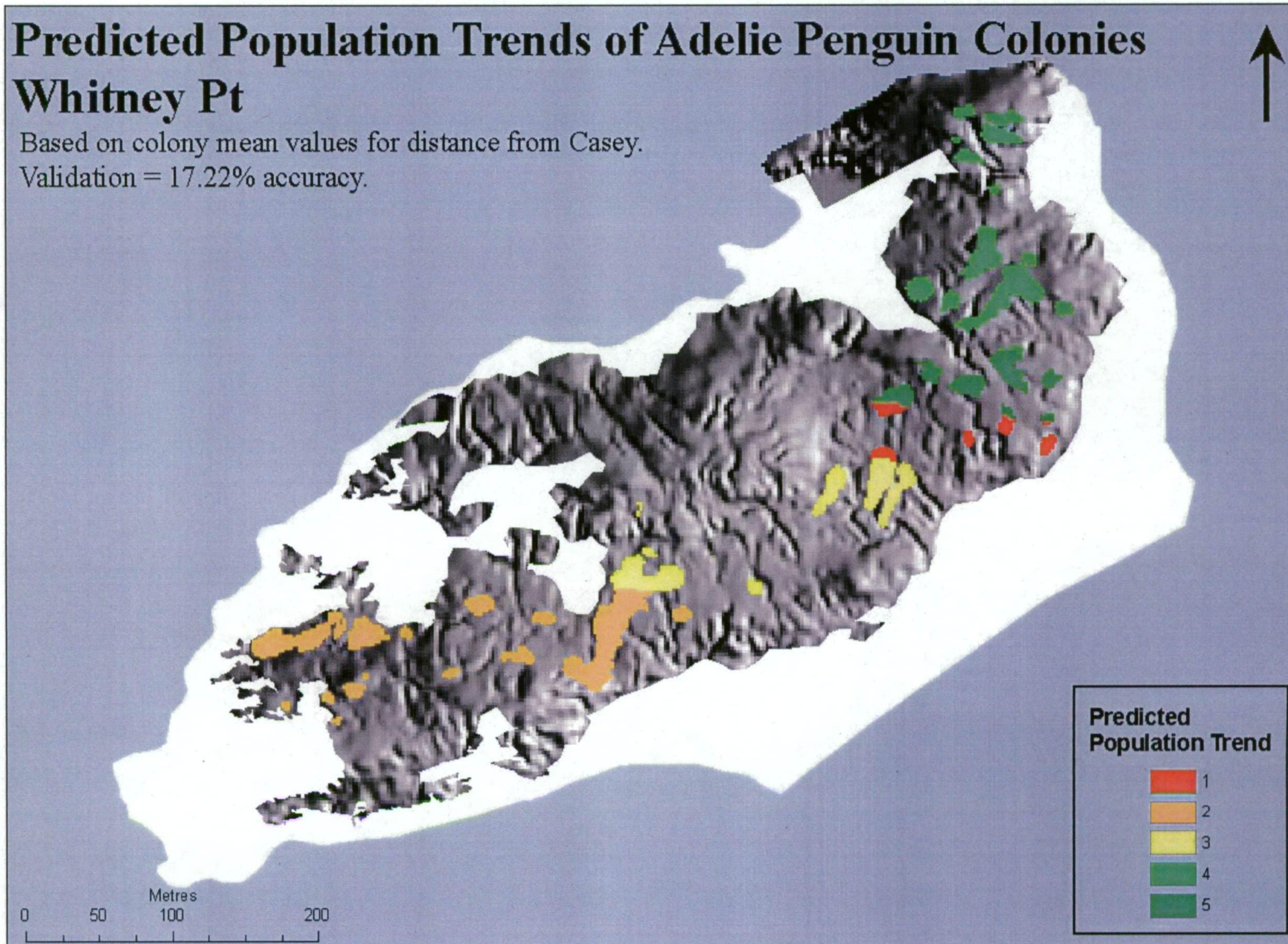


Fig. 4.26: Predicted population trends of Adélie penguin colonies at Whitney Pt, based on discriminant analysis of colony mean values for distance from Casey.

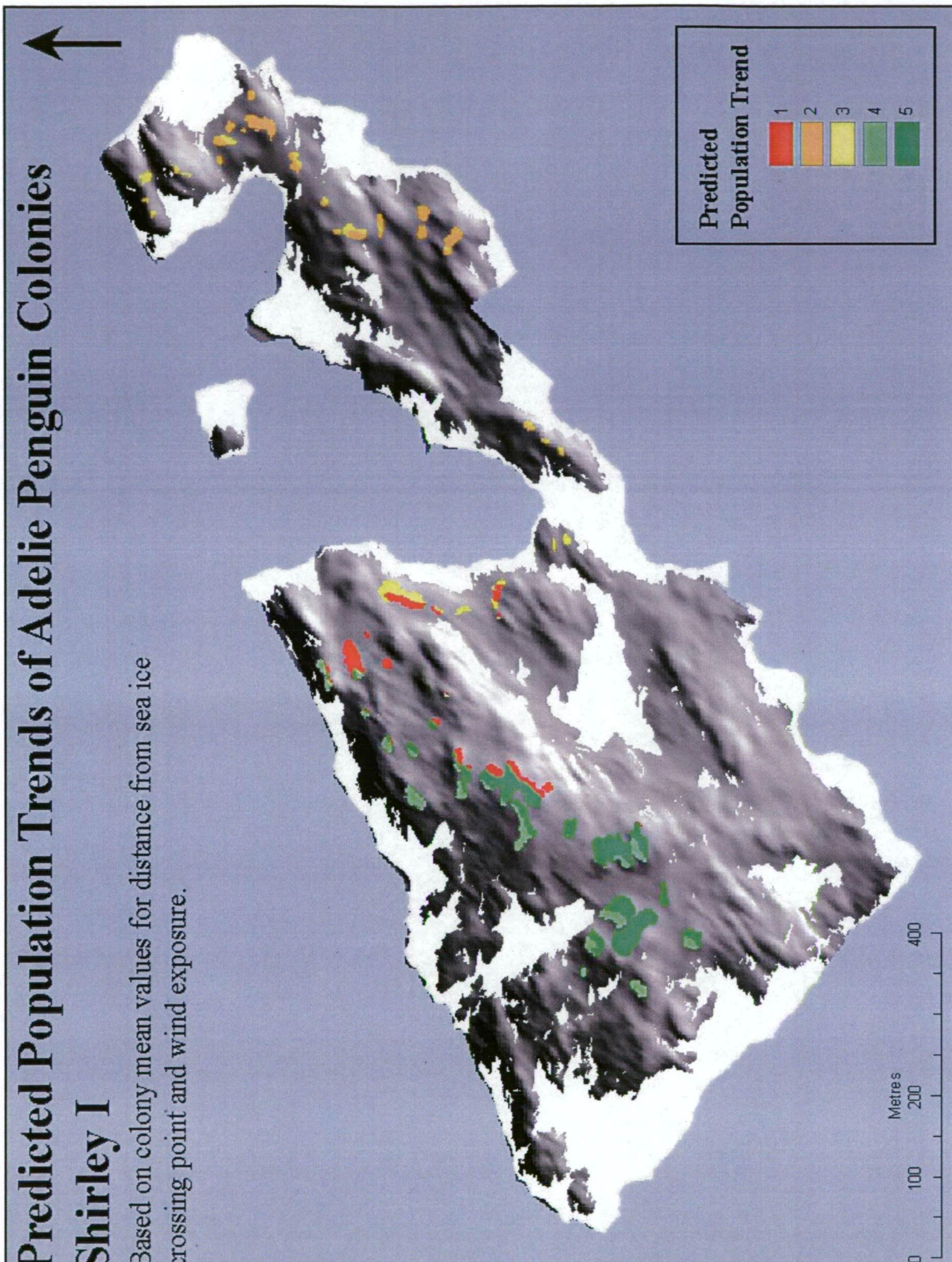


Fig. 4.28: Predicted population trends of Adélie penguin colonies on Shirley I, based on discriminant analysis of colony mean values for proximity to human activities.

4.4.3 Decision Tree Analyses

4.4.3.1 Whitney Pt

Individual Cell Values

A decision tree model was constructed with a minimum object size of 50 and with 5 leaves, using the individual cell values for distance from Casey. Cross-validation shows that the model has an overall accuracy of 88.9% and validation shows that it correctly predicts penguin colony population trends for 90.9% of test cells. An examination of the decision tree shows three distinct bands of cells in class 5 (strong increase), interspersed by a band of cells with stable penguin populations and a band of cells with moderately increasing penguin colony populations (class 4). Both the colonies closest to Casey and farthest away are classified as class 5 (App. 3). The model most accurately predicts increasing colonies (class 4: 100% and class 5: 94.1%). It fails to accurately predict the trends of any decreasing colonies (Table 4.34). However, the small size of the datasets showing population decreases in the test set reduces the reliability of the model for these classes.

Table 4.34: Confusion matrix for the decision tree analysis predicting population trends of Whitney Pt colonies based on the individual cells' distances from Casey.

		Observed Trend Class					
		Class 1	Class 2	Class 3	Class 4	Class 5	Total
Predicted Trend Class	Class 1 (Strong Dec	0	0	0	0	0	0
	Class 2 (Mod Dec.)	0	0	0	0	0	0
	Class 3 (Stable)	0	0	27 (87.1%)	0	5 (3%)	32 (15.3%)
	Class 4 (Mod Inc.)	0	0	2 (6.5%)	5 (100%)	5 (3%)	12 (5.7%)
	Class 5 (Strong Inc.	4 (100%)	1 (100%)	2 (6.5%)	0	158 (94.1%)	165 (79%)
	Total	4	1	31	5	168	209

Colony Mean Values

A decision tree based on the colony mean values for distance from Casey predicted colony trend with 66.7% accuracy, as measured by cross-validation. The model had a minimum object size of two and two leaves (App. 3). Validation with the test set of individual cell values shows that the model has an overall accuracy of 79.9%. The model (Table 4.35) predicts that the populations of all colonies less than 3.3km from Casey are stable (class 3) and the populations of all farther colonies are increasing strongly (class 5).

Table 4.35: Confusion matrix for the decision tree analysis predicting population trends of Whitney Pt colonies based on the colony mean values for distance from Casey.

		Observed Trend Class					
		Class 1	Class 2	Class 3	Class 4	Class 5	Total
Predicted Trend Class	Class 1 (Strong Dec.)	0	0	0	0	0	0
	Class 2 (Mod Dec.)	0	0	0	0	0	0
	Class 3 (Stable)	0	1 (100%)	21 (67.7%)	0	22 (13.1%)	44 (21.1%)
	Class 4 (Mod Inc.)	0	0	0	0	0	0
	Class 5 (Strong Inc.)	4 (100%)	0	10 (32.3%)	5 (100%)	146 (86.9%)	165 (79%)
	Total	4	1	31	5	168	209

4.4.3.2 Shirley I

Individual Cell Values

A decision tree model based on the individual cell values for the parameters related to human proximity predicts Shirley I population trends with 86.5% accuracy, as measured by cross-validation and 83.8% accuracy, as measured by validation. The tree has a minimum object size of 50 and 11 leaves. The distance from the sea-ice crossing point and from Casey are the most predictive parameters, with wind exposure used for two terminal splits in the data. The model uses interactions between the three variables to separate the population trends (App. 3). The confusion matrix (Table 4.36) shows the model most accurately predicted trends for increasing colonies (class 4: 87.7% and class 5: 86.5%). It is difficult to interpret the relative importance of the two distances measures because they covary strongly.

Table 4.36: Confusion matrix for the decision tree analysis predicting population trends of Shirley I colonies based on the individual cell values for distance from the sea-ice crossing point, distance from Casey and wind exposure.

		Observed Trend Class					
		Class 1	Class 2	Class 3	Class 4	Class 5	Total
Predicted Trend Class	Class 1 (Strong Dec.	27 (65.9%)	0	0	0	0	27 (6.3%)
	Class 2 (Mod Dec.)	4 (9.8%)	7 (63.6%)	1 (0.7%)	0	0	12 (2.8%)
	Class 3 (Stable)	5 (12.2%)	4 (36.4%)	119 (85.6%)	1 (1.8%)	12 (6.7%)	141 (33.1%)
	Class 4 (Mod Inc.)	0	0	4 (2.9%)	50 (87.7%)	12 (6.7%)	66 (15.5%)
	Class 5 (Strong Inc.)	5 (12.2%)	0	15 (10.8%)	6 (10.5%)	154 (86.5%)	180 (42.3%)
	Total	41	11	139	57	178	426

Colony Mean Values

A decision tree model derived from the colony mean values related to human activities predicts colony trends with 56.8% accuracy, as measured by cross-validation, and 57.3%, as measured by validation. The tree had a minimum object size of two and nine leaves (App. 3). Distance from Casey is the most important predictor, with wind exposure used for one split in the middle of the tree. Distance from the sea-ice crossing point was not identified as a predictor of observed colony population trends. The model most accurately predicts strongly increasing colonies (82.6%) and strongly decreasing colonies (class 1: 78.1%). It fails to correctly predict any moderately increasing (class 4) colonies (Table 4.37). In this model, those colonies less than 855.58 m from Casey are predicted to be decreasing moderately, while those furthest from Casey (>1778.12 m) are decreasing strongly. However, the latter group has an object size of two and is hence too small to make reliable assumptions from. Those colonies between 1588.42 and 1778.12 m from Casey are predicted to be strongly increasing. For those colonies between 855.58 m and 1588.42 m, a combination of distance from Casey, distance from the sea-ice crossing point and wind exposure is used to differentiate the population trend classes (App. 3).

Table 4.37: Confusion matrix for the decision tree analysis predicting population trends of Shirley I colonies based on the colony mean values for distance from Casey and wind exposure.

		Observed Trend Class					
		Class 1	**Class 2**	**Class 3**	**Class 4**	**Class 5**	**Total**
Predicted Trend Class	**Class 1 (Strong Dec.)**	32 (78.1%)	3 (27.3%)	38 (27.3%)	0	28 (15.7%)	103 (24.2%)
	Class 2 (Mod Dec.)	0	5 (45.5%)	9 (6.5%)	0	0	14 (3.3%)
	Class 3 (Stable)	9 (22%)	3 (27.3%)	60 (43.2%)	0	3 (1.7%)	75 (17.6%)
	Class 4 (Mod Dec.)	0	0	0	0	0	0
	Class 5 (Strong Dec.)	0	0	32 (23%)	57 (100%)	147 (82.6%)	236 (55.4%)
	Total	41	11	139	57	178	426

The results of these models suggest that proximity to human activities partially explains the population trends of Adélie penguin colonies on Shirley I. Distance from Casey appears to have minimal predictive power for the observed population trends of colonies on Whitney Pt. The very high values for the decision tree model based on individual cell values (90.9%) appear to be the result of "banding" in the colonies, with the strongest population trend increases found in the colonies closest to and farthest from Casey. Thus, the null hypothesis should be rejected for Shirley I but not for Whitney Pt.

5 Discussion

5.1 The effect of landscape on Adélie penguin distribution in the Windmill Is

The modelled static landscape parameters have significant predictive power for the presence or absence of Adélie penguin colonies within the study sites at Whitney Pt and Shirley I. A purely random classification for two classes could be expected to predict class membership with approximately 50% accuracy. Both discriminant analysis and decision trees based on both individual cell values and colony mean values predict the presence or absence of Adélie penguins in cells with accuracies far higher than this, suggesting that landscape parameters, as modelled in this study, are important drivers of Adélie penguin colony locations in the Windmill Is. The results are strongest for the individual cell analyses and for the decision trees. The parameters associated with elevation change – slope, surface roughness (standard deviation) and surface roughness (normalised) – are repeatedly shown to be important predictors of colony locations. Surface roughness (standard deviation), solar radiation and wind exposure are the most commonly selected variables in the discriminant analysis and decision tree models. These all improved the predictive power of three of the four models. All of the measured landscape variables were used in at least one of the models, except for the surface curvature layers and aspect, which could not be applied in discriminant analyses or averaged because of the circular nature of the data. The models predict Adélie penguin distribution with higher accuracy for Shirley I than for Whitney Pt by 5.4-16.3%. It is likely that this difference is mostly driven by the difference in the accuracy of the DEMs.

No single static landscape parameter, apart from the presence or absence of permanent snow, can be used to predict the presence or absence of Adélie penguins. Instead, it appears that a complex interaction of several landscape parameters affect habitat suitability. However, a few generalisations can be made. Adélie penguin colonies generally do not occur on the steepest slopes, regardless of aspect, or on moderately steep south-facing slopes with low solar radiation. Similarly, the birds appear to choose colony sites with moderate levels of wind exposure. Beyond that, simple spatial rules determining colony locations could not be discerned. The colonies occur on hilltops, valley bottoms and the sides of hills, facing almost all directions, and with varying scores on the wetness index.

These results at least partially agree with previous studies into the relationships between static landscape parameters and the distribution of Adélie penguin nest sites. Snow accumulation has repeatedly been found to be a significant factor in the distribution of Adélie penguins in other parts of Antarctica (e.g. Levick, 1915; Yeates, 1975; Moczydlowski, 1986, 1989; Trivelpiece and Fraser, 1996; Fraser and Patterson, 1997; Patterson et al., 2003).

This study is the first to compare the observed distribution of Adélie penguin colonies with snow distributions derived from a spatial, physically-based blowing snow model. Historically, studies investigating the relationship between colonies and snow cover have relied on observations or direct measurements of snow depth in or near colonies on a given date. These approaches were unable to account for spatial variability within and among colonies, or to account for temporal changes in snow cover. A more spatially-explicit approach was taken by Patterson et al. (2003), who used a GIS hillshade model as a surrogate for snow accumulation. As has been discussed earlier in this paper, that approach was more directly a surrogate for exposure to prevailing winds. Wind exposure is not the same as snow accumulation, as may be highlighted by considering the case of a concave cliff that faces prevailing winds. The area at the base of the cliff is likely to accumulate snow (this can be seen near colonies I-IV at Whitney Pt, P.K. Bricher, pers. obs.), but a hillshade model would display that area as highly exposed to the prevailing winds, and hence likely to be free of snow. However, there is a strong correlation between wind exposure and modelled snow accumulation, and it is therefore difficult to separate the relative effects of each parameter.

Ainley (2002) proposed that Adélie penguins typically nest on ridges and on higher ground. He did note that where they nested in single-species colonies, they are found closer to sea-level. Similarly, Wilson et al. (1990) found that Adélie penguin colonies occur on well-drained mounds. This study found that Adélie penguins nest at all elevations in the study sites, although they appear to occur in altitude "bands". This may be because the terrain in the Windmill Is is dominated by plateaux and low cliffs. There was no evidence that Adélie penguins nest only on ridges. While many colonies on Shirley I and at Whitney Pt do occur on ridges, others occur in valley bottoms and at the bases of hills. Profile and planar curvature, which are measures of the shape of a slope, were repeatedly found not to improve the predictive power of the distribution models.

Yeates (1975) found that Adélie penguins select nest sites with the highest solar radiation and wind exposure measures. In the present study, the highest modelled solar radiation levels were found on north-facing slopes (Fig. 4.4), and the highest wind exposure on south-east facing slopes (Fig. 4.10). An examination of the scatterplots of solar radiation and wind exposure showed that, Adélie penguin colonies did not occur at sites with either the highest or the lowest solar radiations, especially at Whitney Pt. Instead, they were clustered in sites with modelled annual solar radiation values between 3200 and 3700 MJ/m^2, on a scale of 2500 to 4000 MJ/m^2. In addition, the colonies were clustered on sites with moderate wind exposure. Based on the results obtained in this study, it appears that in the Windmill Is, Adélie penguins select nest sites that balance their requirements for sites with at least moderate levels of solar radiation and that are moderately exposed to prevailing winds.

Moczydlowski (1986; 1989) found that high levels of solar radiation and good drainage were common features of all Adélie penguin colonies on the peri-Antarctic South Shetland Is and that while colonies were located in the sites with the thinnest snow cover, they did not occur in the most exposed sites. In the present study, drainage appeared as a predictive factor in three of the four distribution models based on mean values, but it did not increase the predictive power of the other models. It is possible that drainage is less important as a predictor of Adélie penguin habitat in the Windmill Is, which have a drier continental Antarctic climate, than it is in wetter, maritime peri-Antarctic regions. It is also likely that the wetness index values are more susceptible to the effects of within-colony spatial autocorrelation than other variables. The results of the present study also agreed with Moczydlowski's finding that Adélie penguin colonies do not occur in the sites with the highest wind exposure. An examination of scatterplots and histograms of the wind exposure data for both study sites showed that the colonies were clustered in sites with medium levels of wind exposure.

The techniques applied here allowed an objective analysis of the spatial variability of all the modelled static landscape variables, whereas previous studies have typically relied on subjective descriptions or individual measurements of the variables in a low number of colonies. The multivariate models developed in this study are able to predict the presence or absence of Adélie penguin nests within a 4m^2 grid cell with up to 84.5% accuracy.

5.2 The effect of snow accumulation patterns on Adélie penguin colony population trends in the Windmill Is

On the Antarctic Peninsula, snow accumulation patterns have been found to be important predictors of Adélie penguin colony population trends and of colony distributions (Fraser and Patterson, 1997; Patterson et al., 2003). Results from this study support those findings. At Whitney Pt, the modelled snow accumulation and wind exposure layers explain much of the variation in colony population trends. These layers have less predictive power on Shirley I, where proximity to human activities explain a large proportion of the variation on colony population trends (see section 5.4).

At Whitney Pt, the colonies with strong population increases (>150% of the 1959 population) are associated with the thinnest modelled snow cover, while stable colonies are found in areas with thicker snow cover. For the individual cell values, colonies of all population trend classes are associated with areas with minimal changes in the modelled snow cover between 1959 and 2005. However, the colony mean values of colonies with the strongest decreases (<50% of the 1959 population) are associated with the areas showing the greatest increases in snow accumulation. Given the small sample size for decreasing colonies on Whitney Pt, this result should be interpreted cautiously. The snow accumulation data for Shirley I do not show such obvious trends.

The Whitney Pt results agree with the findings of Patterson et al. (2003) who found that exposure to prevailing winds acted as a primary driver of colony population trends. Their study was conducted in an area where increasing mean temperatures led to increased snowfall. In the Windmill Is, little attention has been paid to potential climate change, and the modelled change in snow cover in this study suggested that there was little broad-scale change in snow accumulation during the period under examination (1959-2005). This suggests that a broader process is driving the overall increase in Adélie penguin numbers for the Windmill Is, but that snow accumulation may mediate that increase in individual colonies, at the site furthest from human activities.

5.3 The effect of proximity to human activities on Adélie penguin colony population trends in the Windmill Is

Distance from Casey has some predictive power for the population trends of colonies at Whitney Pt. The discriminant analysis based on individual cell values predicts colony trends with 26.7% accuracy, a small but significant effect. In contrast, the decision tree based on individual cell values predicts trends with 90.9% accuracy – a highly significant result. This highlights the differing assumptions of discriminant analysis and decision trees. As a parametric test, discriminant analysis assumes that the values of all points in a class will be clustered around a single value. As a non-parametric test, a decision tree is able to cope with multimodal data, such as those observed at

Whitney Pt. The decision tree identified three distinct "bands" of colonies that had strongly increasing populations. These bands were interspersed by a band of colonies with moderately increasing populations and a band of stable colonies. The fact that the colonies both closest to and farthest from Casey have strong population increases suggests that some other environmental factor, that is not accounted for here, is causing bands of terrain with different levels of suitability.

Shirley I is much closer to Casey, is regularly visited by station personnel and is immediately downwind of Casey. At this site, proximity to human activities has much greater predictive power for colony population trends than at Whitney Pt. Wilcoxon tests show significant differences among population trend classes for distance from Casey and from the sea-ice crossing point used by personnel to access the island. These differences are significant for both the individual cell values and the colony mean values. It is difficult to separate the effects of these two variables because of the strong correlation between the distance from Casey and distance from the sea-ice crossing point. Wind exposure, which acts as a surrogate for both exposure to prevailing winds and to possible exposure to noise and particulates from Casey, has little predictive power compared with the distance measures.

The results of the analyses based on individual cell values are stronger than those based on colony mean values for Shirley I. The discriminant analysis based on individual cell values has an overall accuracy of 72.1%, compared with 40.6% for the discriminant analysis based on colony means. However, spatial autocorrelation is likely to have affected the results of the models based on individual cell values. The predictive map (Fig. 4.32) derived from the discriminant analysis based on individual cell values, shows that the model strongly predicts the trends of the large colonies at the western end of the island, but poorly predicts the trends of smaller colonies in other parts of the island. The predictive map (Fig. 4.33) derived from the discriminant analysis based on colony mean values shows that this model more accurately predicts the trends of smaller colonies at the eastern end of the island. The difficulty in applying decision tree results in a GIS precludes a visual assessment of the performance of the decision tree models. Using the more conservative results of the models based on colony mean values, the discriminant analysis has approximately twice the predictive power that a purely random classifier might be expected to have (40.6%, compared with 20%). The decision tree has almost three times the predictive power than it would if proximity to human activities had no effect on Adélie penguin colony population trends (57.3%, compared with 20%).

It is difficult to determine whether the distance from Casey or the distance from the sea-ice crossing point is the most significant factor in explaining penguin colony population trends on Shirley I for the period 1968-2005, because of correlations between these two variables. As a result, their relative importance varied among the models. Therefore, these results should be interpreted with caution. While these results suggest that proximity to human activities is a significant driver of Adélie penguin colony population trends, further investigations are required to separate the effects of station-related activities, such as noise and particulate emission, and the effects of visits to the colonies by station personnel.

Previous studies into the effects of human activities on Adélie penguin population trends have produced site-specific results. The results of the present study at least partially support the proposition of Woehler et al. (1994) that visits by station personnel appear to cause decreases in populations among some Adélie penguin colonies on Shirley I. A similar situation was found at Cape Bird, Ross I (Young, 1990) where Adélie penguin colonies close to the research station underwent significant population decreases at a time when the overall penguin population was increasing.

Giese (1996) reported significantly lower breeding success in Adélie penguin colonies that had been subjected to daily recreational visitors or to regular scientific nest-checks. She concluded that the frequency of disturbance drove the magnitude of the decrease in breeding success. Giese's study was conducted in a breeding locality that had been little disturbed by previous human activities. In contrast, Patterson et al. (2003) investigated colonies on Torgersen I, near Palmer on the Antarctic Peninsula that had been regularly visited by tourists and researchers for many years. They found that tourism had no detectable effect on Adélie penguin breeding population size or breeding success. Similarly, Fraser and Patterson (1997) had found no correlation between Adélie penguin population trends and human-use histories of breeding localities near Palmer.

Studies of the effects of proximity and/or exposure to human activities on the breeding success and population trends of Adélie penguins have also examined the effects of habitat modification on the birds. At Cape Hallett, Wilson et al. (1990) found that Adélie penguin populations decreased during the period in which the station was inhabited, and that they subsequently returned to the numbers present before human occupation, once the station had been abandoned. It is difficult to separate the effects of habitat modification from the effects of visitation or disturbance associated with station activities when habitat modification directly impacts on nesting habitat. At Dumont d'Urville, Terre Adélie, Micol and Jouventin (2000) found that Adélie penguin numbers had increased by 49% in 14 years, despite the destruction of some colonies for the construction of a runway that spanned three islands. They found above-average population increases (154%) on the Ile des Pétrels, where the station is located. The largest population increase (826%) over the same time period was found in the breeding locality farthest from the station, at Cap Géodésie. However, extensive habitat modification on islands around Dumont d'Urville makes it difficult to draw conclusions about the effects of proximity to human activities on these penguin population trends.

From these studies, it appears that the effects of proximity and/or exposure to human activities are determined by a combination of the types of activities involved and the history of interactions between humans and Adélie penguins at a given site. They may also be confounded by environmental trends at regional scales. The results of this study suggest that in the Windmill Is, proximity and/or exposure to human activities may play a significant role in mediating the observed long-term increases in Adélie penguin numbers. Of the colonies for which long-term census data are available, the majority of colonies with decreasing population trends are located on Shirley I, which is the closest breeding locality to Casey (Woehler et al., 1991). On Shirley I, discriminant analysis and decision tree models based on proximity and exposure to human activity data predict population trend classes with a high degree of accuracy (40.6%-86.5%).

5.4 How this study compares with other GIS-based habitat analyses

The majority of GIS-based habitat analyses have focused on mapping or predicting the distribution of a species (Guisan and Zimmermann, 2000; Lenton et al., 2000; Osborne, Alonso and Bryant, 2001; Lauver, Busby and Whistler, 2002; Gibson et al., 2004). Typically they have not investigated temporal changes in habitat suitability (Curnutt et al., 2000). The present study was able to do this because of the relative ease with which current and relic Adélie penguin colonies can be mapped, and because of the existence of long-term population data for Adélie penguin colonies in the Windmill Is.

A common problem in habitat suitability analyses is that it is rarely possible to determine that a site has never or will never be used by the species under examination (Breininger et al., 1991). In the present study, it was possible to determine which sites have been used in the past because relic Adélie penguin colonies are relatively easy to map. However, this does not preclude the possibility that other suitable sites are available that have not yet been exploited by the birds. Fielding and Bell (1997) warned that interference will occur in a model if a species is not using the entire available suitable habitat. It is likely that this is the case in the Windmill Is, as evidenced by the high proportion of recently established colonies at Whitney Pt on sites that had no evidence of previous habitation (Martin et al., 1990). The strength of the predictive power of the distribution models in the present study was similar to the results of other GIS-based habitat analyses (e.g. Aspinall and Veitch, 1993; Debinski et al., 1999; Guisan and Zimmermann, 2000; Osborne et al., 2001).

Most habitat analysis studies have used just one multivariate statistical technique (e.g. Debinski et al., 1999; Patterson et al., 2003; Gibson et al., 2004). This study followed the example of Blackard and Dean (1999), Manel et al. (1999) and Guisan and Zimmermann (2000) and applied two different statistical techniques, with differing underlying assumptions. The discriminant analyses and decision trees produced similar results for each of the models, except in the cases outlined earlier in this discussion. The fact that different inputs were selected for the different models suggests collinearities between some of the data, and the use of two different tests enabled confirmation of the results and the identification of weaknesses in the models. The data used in this study violate the assumption of normality, as well as assumptions about equality of variance and of covariance matrices (Flury and Riedwyl, 1988) that underpin discriminant analysis. Although it has been argued that violations of the normality assumptions of discriminant analysis have apparent minimal effect on results (Blackard and Dean, 1999), it was considered appropriate to compare the results of disciminant analysis models with the non-parametric decision tree analysis. The results show that the non-parametric decision trees have stronger predictive power than the parametric discriminant analyses.

This study used advanced GIS habitat analysis techniques, including complex models and multivariate statistical tests. It used two different multivariate modelling methods, and used colony mean values to eliminate the effects of within-colony spatial autocorrelation. This study therefore accounted for potential artefacts of the methods that have often been ignored in previously published studies (Legendre, 1993; Guisan and Zimmermann, 2000). However, it does not eliminate among-colony spatial autocorrelation, and this is an area for future investigation (Legendre, 1993).

5.5 Limitations on the study

The results of the analyses based on DEM derivatives for Whitney Pt are less reliable than those for Shirley I, because the accuracy of the Whitney Pt DEM was severely limited by the snow cover in the available aerial photograph. The DEM had a mean height accuracy of ±2m for areas of exposed rock and ±6.6m for snow-covered areas. This, in turn, affected the reliability of the DEM derivatives. It is likely that the large differences in the predictive power of the models of Adélie penguin colony distribution between Shirley I and Whitney Pt are caused by the greater errors in the Whitney Pt DEM. The Shirley I DEM was much more accurate, but was affected by the "smoothing" effect where stereo-models overlapped. This effect was reduced in the interpolated DEM that was used for the analyses. The effects of these errors on the analyses were unquantifiable; however, it was considered that they were less severe than the errors in the Whitney Pt DEM. As outlined in section 6.2, better aerial photography would greatly reduce the errors resulting from positional uncertainty in the DEMs.

The Adélie penguin population trends were calculated from the changes between counts conducted in two breeding seasons (1959/60 and 2005/06 for Whitney Pt and between 1968/69 and 2005/06 for Shirley I). These calculated trends were potentially affected by counting errors and by data aliasing issues associated with interannual fluctuations in breeding pair numbers. However, an examination of the plotted long-term trends showed a strong agreement with the calculated trends for these colonies. In addition, the general trends for the population trends are well-known (Woehler et al., 1991; Woehler et al., 1994). Trends in colony populations were categorised into five classes. Doing this increased the power of the resulting statistical analyses, but reduced the sample size and concomitant degrees of freedom. Classifying the data into regularly-spaced classes allowed comparison of "like" population trends, but it resulted in small sample sizes for some classes of data. This was especially so for the moderately increasing class (class 4), and made interpretation of the multivariate models difficult for these classes.

The boundary cells of current and relic Adélie penguin colonies were removed from the analyses because of the positional uncertainty of the colony boundaries. This ensured that all cells described as having Adélie penguins present represented actual penguin colony habitat. However, it reduced the number of data points in the analyses and resulted in some very small colonies being entirely excluded from the analyses. It is possible that the landscape properties of these small colonies are different to larger colonies. It is also possible that the effects of environmental or human-related stressors are more severe for these small colonies, as has been shown in other parts of Antarctica (Giese, 1996; Patterson et al., 2003).

The solar radiation model was a measure of potential, rather than actual solar radiation, and was hence unable to account for the effect of cloud cover on actual radiation. As the PotRad model used here was designed for use in the tropics, it did not account for the high albedo of snow and ice present in the Antarctic (van Dam, 2001). Thus, it is possible that the model underestimated the total solar radiation for south-facing slopes. However, the results of the models here agree with findings in other parts of Antarctica, where Adélie penguin colonies were found not to occur on the sites with the lowest solar radiation (Moczydlowski, 1986; 1989). The solar radiation model used in this study proved to be an important predictor of the distribution of Adélie penguin colonies, even if it did not show actual solar radiation levels. The model could be tested by comparing its modelled results with pyranometer observations along transects at the study sites.

The NCEP/NCAR weather reanalysis data were produced for a grid point located 63.1km to sea off Casey. Thus, the effects of local topographic features in mediating the weather conditions at the study sites could not be examined in this study. It has also been argued elsewhere that the paucity of weather observations in the Southern Ocean reduced the accuracy of the reanalysis data for Antarctica (Hines et al., 2000). The sensitivity of the snow accumulation model to different weather data input could be tested in the future by comparing results of a snow model based on NCEP/NCAR data with one based on Bureau of Meteorology observations for years where both sets of data are available.

The snow accumulation model produced maps of relative spatial patterns of snow accumulation, rather than numerical results of snow depths. It was not validated with ground-truth data, and so the modelled snow distribution cannot be assumed to represent actual snow distribution. Snow accumulation patterns have long been known to be important drivers of the distribution and population trends of Adélie penguin colonies in the Antarctic (e.g. Levick, 1915; Moczydlowski 1986, 1989; Fraser and Patterson, 1997; Patterson et al., 2003). However, GIS-based snow accumulation models are rare and often rely on input data that is not available for Antarctic environments (Liston and Sturm, 1998). Although no snow accumulation model can expect to capture all of the physical processes associated with snow transport, the model used in this study showed strong visual agreement with observed snow patterns and with Adélie penguin colony distribution and population trends (Greene et al., 1999).

The discriminant analyses were likely to be adversely affected because some of the data violated the assumptions of normality, independence of variables and equal variance that underpin discriminant analysis (Flury and Riedwyl, 1988). Other multivariate statistical techniques, such as logistic regression, that do not make the same assumptions could have been used more effectively here. Because of the data violations of the discriminant analysis, it was considered appropriate to repeat the analyses using decision tree models. Decision trees are capable of handling data that are not normally distributed. The resulting models consistently had stronger predictive power than the discriminant analyses, but could not be readily applied in a GIS. Both of the study sites consisted of several hundred thousand cells, and it was unwieldy to produce predictions for each cell in Weka. To implement the models in ArcGIS would have involved complex sets of nested conditional statements. This was a major limitation to the study, as the models with the greatest predictive power could not be explored spatially. However, there is potential for a tool to be developed to enable the implementation of decision trees in a GIS environment.

6 Conclusions

6.1 Conclusions of the study

This study used advanced GIS-based analysis and multivariate statistical techniques to investigate factors that affect the distribution and population trends of Adélie penguin colonies within breeding localities. Landscape parameters were derived from fine-scale DEMs, and a physically-based snow accumulation model was used to simulate the patterns of snow cover. Discriminant analysis and decision tree analysis were used to construct predictive models of distribution based on static landscape parameters, and of population trends based on parameters associated with snow accumulation and proximity to human activities.

This study showed that landscape parameters can explain much of the distribution of Adélie penguin colonies within breeding localities in the Windmill Is. In particular, slope, surface roughness, wind exposure and solar radiation were found to have the greatest predictive power for Adélie penguin colony distribution. Further, the study showed that at Whitney Pt, which is 3km upwind of Casey, parameters associated with snow accumulation patterns can explain much of the variation in population trends. At Shirley I, proximity to activities associated with the station explained much of the variation in population trends, with snow accumulation having reduced predictive power.

The distribution analysis displayed a general agreement with the findings of previous Adélie penguin habitat analyses, in that the penguins chose to nest in areas with modelled high solar radiation, thin snow cover and with moderate exposure to prevailing winds. However, it also found that the distribution of Adélie penguin colonies at Whitney Pt and on Shirley I could not be easily explained by general rules or by the values of a few static landscape parameters, as has been previously suggested (e.g. Moczydlowski, 1986; 1989). The colonies occurred on slopes facing every compass direction; in the bottoms of gullies and on the tops of ridges; in sites with all but the lowest modelled solar radiation and on all but the steepest or roughest terrain. The distribution of colonies appeared to be governed by a complex interaction of landscape parameters.

The analyses of the effects of snow accumulation parameters on Adélie penguin colony population trends for Whitney Pt showed strong agreement with the findings of Patterson et al. (2003). At Whitney Pt, as on the Antarctic Peninsula the colonies with the highest modelled levels of snow accumulation and lowest wind exposure had the strongest decrease in population. However, at Shirley I, the snow accumulation layers had much less predictive power for Adélie penguin colony trends. There, the layers showing proximity to human activity had much more power to explain observed population trends for colonies than the snow accumulation layers. It appears that at Shirley I, local effects associated with Casey explain a large amount of the variability in population trends of colonies. Further research is needed to investigate potential causative factors, as it is unclear whether the observed effects are the result of human visits to colonies, emissions from Casey, some combination of these or another cause. However, this study demonstrated that the effects of natural climate variability can be mediated at a local scale by proximity to human activities.

6.2 Future directions for research

Aerial photography

One of the major limitations on the accuracy and reliability of the data layers used in this study was imposed by the accuracy of the DEMs that could be constructed from the available aerial photography. It is believed that the difference in the predictive power of the distribution models for Shirley I and Whitney Pt was caused by the difference in the accuracy of the DEMs. Better aerial photography for both sites – with the same spatial resolution and taken on days with minimal snow cover – would improve the accuracy of the resulting DEMs and the DEM derivatives. In addition, the existing aerial photographs of Shirley I could be used to produce very fine resolution DEMs for small areas of the island, which would enable the investigation of the effect of microtopography on Adélie penguin colonies.

Sample size

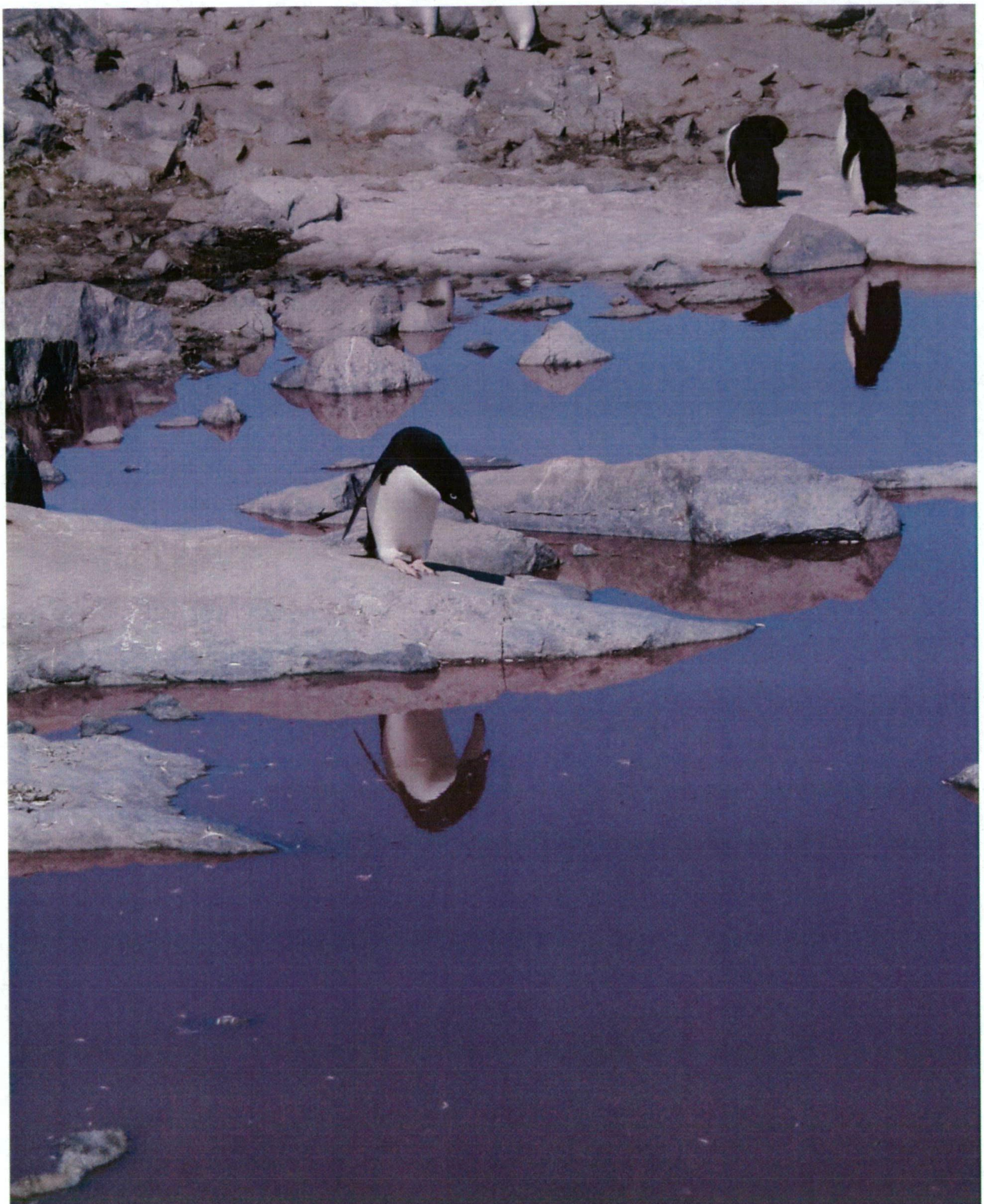
This study examined distribution and population trend data for approximately 80 colonies in two breeding localities. The number of colonies included in the study enabled the data points to be replicated to account for local effects. However, the analyses could be repeated for other locations in the Windmill Is and elsewhere to test the applicability and generality of these results. In addition, examination of the effects of the landscape, snow accumulation and proximity to human activity parameters at different spatial and temporal scales could improve understanding of the processes involved.

Statistical improvements

This study attempted to account for the effect of spatial autocorrelation by building predictive models on colony mean values as well as individual cell values. Future research could investigate the use of more sophisticated statistical methods to account for spatial autocorrelation. Different classification methods for population trends could produce more equal class-sizes, and hence avoid creating classes which could not be adequately examined in some of the models. The use of logistic regression, rather than discriminant analysis, would make fewer assumptions about the underlying structure of the data and would hence be more robust. Decision trees produced statistically-robust models, but cannot be readily applied in a GIS, without writing specific scripts, a task that is beyond the scope of this study. It would be worthwhile to develop a tool to readily translate the results of a decision tree into a format that could be applied in a GIS to produce predictive maps.

Human impacts

Given the increasing human presence in Antarctica, it is important to identify potential impacts on Adélie penguin populations. In particular, there is a need to separate the effects of natural variability from anthropogenic variability. This study identified significant relationships between proximity to human activities and Adélie penguin population trends. Further investigation of these results may involve examination of the full spectrum of station activities, including visits to Adélie penguin colonies by station personnel and station emissions such as noise, particulates and sewage, in a bid to establish causative relationships.



Narcissus: Adélie penguin on Shirley I, January 2005.

References:

Ainley, DG, Ballard, G, Barton, KJ, Karl, BJ, Rau, GH, Ribic, CA and Wilson, PR, (2003) 'Spatial and temporal variation of diet within a presumed metapopulation of Adélie penguins', *The Condor*, 105 **(1)**, pp. 95-106.

Ainley, DG and Le Resche, RE (1973) 'The effects of weather and ice conditions on breeding in Adélie penguins', *The Condor*, **75**, pp. 235-255.

Ainley, DG, Nur, N and Woehler, EJ (1995) 'Factors affecting the distribution and size of pygoscelid penguin colonies in the Antarctic', *The Auk*, **112(1)**, pp. 171-182.

Anders R (2005) *Photogrammetric Digital Elevation Model of Whitney Point, Windmill Islands, East Antarctica*, University of Tasmania.

Anders R (2006) *Photogrammetric Digital Elevation Model of Shirley Island, Windmill Islands, East Antarctica*, University of Tasmania.

Aspinall, R and Veitch, N (1993) 'Habitat mapping from satellite imagery and wildlife survey data using a Bayesian modeling procedure in a GIS', *Photogrammetric Engineering and Remote Sensing*, **59(4)**, pp. 537-543.

Austin, MP (2002) 'Spatial prediction of species distribution: an interface between ecological theory and statistical modelling', *Ecological Modelling*, **157**, pp. 101-118.

Australian Antarctic Data Centre (1947) Antarctica Map, Australia 1:11 250 000 map, (Commonwealth of Australia), www.aad.gov.au.

Australian Antarctic Data Centre (1990) Aerial photographs of Whitney Point, Film ANT 142, Australian Antarctic Division (Commonwealth of Australia), www.aad.gov.au.

Australian Antarctic Data Centre (1994) Aerial photographs of Whitney Point, Film ANT 1007, Run 19, Australian Antarctic Division (Commonwealth of Australia), www.aad.gov.au.

Australian Antarctic Data Centre (1994) Aerial photographs of Whitney Point, Film ANTC 1000, Australian Antarctic Division (Commonwealth of Australia), www.aad.gov.au.

Australian Antarctic Data Centre (1994) Aerial photographs of Shirley Island, Film ANTC 1004, Run 26, Australian Antarctic Division (Commonwealth of Australia), www.aad.gov.au.

Australian Antarctic Data Centre (2001) Aerial photographs of Shirley Island, Film ANTC 1099, Run, 1-3, Australian Antarctic Division (Commonwealth of Australia), www.aad.gov.au.

Australian Antarctic Data Centre (2002) Map 12939 - "North-east Bailey Peninsula, Antarctic Specially Protected Area No.135, Geology", Australian Antarctic Division, (Commonwealth of Australia), www.aad.gov.au.

Australian Antarctic Data Centre (2003) Aerial photographs of Windmill Islands, film ANTC 1216, Runs 15 and 24, Australian Antarctic Division (Commonwealth of Australia), www.aad.gov.au.

Australian Antarctic Data Centre (2006) Automatic Weather Stations at Bases, www.aadc-maps.aad.gov.au/aadc/aws/, accessed 8/9/2006.

Australian Antarctic Data Centre (2006) 'Automatic Weather Station Data from Casey', http://aadc-db.aad.gov.au, last updated 25/5/2006, Accessed 28/8/2006.

Australian Antarctic Division (2006) Clark Peninsula ASPA No 136, www.aad.gov.au, accessed 7/09/2006, last updated 01/09/2004.

Baker, BW, Cade, BS, Mangus, WL, McMillen, JL (1995) 'Spatial analysis of sandhill crane nesting habitat', *Journal of Wildlife Management*, **59(4)**, pp. 752-758.

Beyer, HL (2004) Hawth's Analysis Tools for ArcGIS, http://www.spatialecology.com/htools, Accessed 10/07/2006.

Bian, L and West, E (1997) 'GIS modeling of elk calving habitat in a prairie environment with statistics', *Photogrammetric Engineering and Remote Sensing*, **63(2)**, pp. 161-167.

Bintanja, R, Lilienthal, H and Tüg, H (2001) 'Observations of snowdrift over Antarctic snow and blue-ice surfaces', *Annals of Glaciology*, **32**, pp. 168-173.

Birt, VL, Birt, TP, Goulet, D, Cairns, DK, Montevecchi, WA (1987) 'Ashmole's Halo: Direct evidence for prey depletion by a seabird', *Marine Ecology Progress Series*, **40**, pp. 205-208.

Blackard, JA and Dean, DJ (1999) 'Comparative accuracies of artificial neural networks and discriminant analysis in predicting forest cover types from cartographic variables', *Computers and Electronics in Agriculture*, **24**, pp. 131-151.

Breininger, DR, Provancha, MJ and Smith, RB (1991) 'Mapping Florida scrub jay habitat for purposes of land-use management', *Photogrammetric Engineering and Remote Sensing*, **57(11)**, pp. 1467-1474.

Bureau of Meteorology (2004): Climate averages for Australian sites, www.bom.gov.au, last updated 16/8/2004, accessed 5/09/2006.

Bureau of Meteorology (2006) 'History of Casey meteorological office', www.bom.gov.au, accessed 25/8/2006.

Burrough, PA and McDonnell, RA (2000): *Principles of Geographical Information Systems*, Oxford University Press, Oxford.

Clarke, J and Kerry, K (1994) 'The effects of monitoring procedures on Adélie penguins', *CCAMLR Science*, **1**, pp. 155-164.

COMNAP (2006) 'Antarctic Facilities in Operation', http://www.comnap.aq/operations/facilities/, Council of Managers of National Antarctic Programs, accessed 28/10/2006.

Croxall, JP, Trathan, PN and Murphy, EJ (2002) 'Environmental change and Antarctic Seabird Populations', *Science*, **297**, pp. 1510-1514.

Culik, B, Adelung, D and Woakes, AJ (1990) 'The effect of disturbance on the heart rate and behaviour of Adélie penguins (*Pygoscelis adeliae*) during the breeding season', in Kerry, K and Hempel, G (eds.) *Antarctic Ecosystems. Ecological Change and Conservation*, Springer-Verlag, Berlin Heidelberg, 1990.

Curnutt, JL, Comiskey, J, Nott, MP and Gross, LJ (2000): "Landscape-based spatially explicit species index models for Everglades restoration", *Ecological Applications*, **10(6)**, 1849-1860.

Daly, SF, Davis, R, Ochs, E and Pangburn, T (2000) 'An approach to spatially distributed snow modelling of the Sacramento and San Joaquin basins, California', *Hydrological Processes*, **14**, pp.3257-3271.

De'Ath, G and Fabricius, KE (2000) 'Classification and regression trees: A powerful yet simple technique for ecological data analysis, *Ecology*, **81(11)**, pp. 3178-3192.

Debinski, DM, Kindscher, K and Jakubauskas, ME (1999) 'A remote sensing and GIS-based model of habitats and biodiversity in the Greater Yellowstone Ecosystem', *International Journal of Remote Sensing*, **20(17)**, pp. 3281-3291.

Doorschot, JJ, Lehning, M and Vrouwe, A (2004) 'Field Measurements of Snow-Drift Threshold and Mass Fluxes, and Related Model Simulations', *Boundary-Layer Meteorology*, **00**, p. 1-22.

Dytham, C (2003) *Choosing and Using Statistics A Biologist's Guide*, Blackwell, Science, Malden, MA.

Emslie, SD and Woehler, EJ (2005) 'A 9000-Year Record of Adélie Penguin Occupation and Diet in the Windmill Islands, East Antarctica', *Antarctic Science*, **17(1)**, pp. 57-66.

ESRI (1999-2004): ArcGIS 9.0, http://www.esri.com, accessed 25/07/2006.

ESRI (2006) 'GIS Dictionary', http://support.esri.com, accessed 20/11/2006.

Evans, BM, Walker, DA, Benson, CS, Nordstrand, EA and Peterson, GW (1989) 'Spatial interrelationships between terrain, snow distribution and vegetation patterns at an arctic foothills site in Alaska', *Holarctic Ecology*, **12**, pp. 270-278.

Evans, K (1991): *Population assessment, colony distribution and nest location of Adélie penguins (utilising computer scanning techniques on aerial photographs) in the Windmill Islands, Antarctica*, Honours thesis, University of Tasmania, Hobart.

Falla, RA (1937) 'Birds', *B.A.N.Z. Antarctic Research Expedition 1929-1931 Reports*, **Series B, vol. 2**, B.A.N.Z.A.R. Expedition Committee, Hassell Press, Adelaide.

Fielding, AH and Bell, JF (1997) 'A review of methods for the assessment of prediction errors in conservation presence/absence models', *Environmental Conservation*, **24(1)**, pp. 38-49.

Flury, B and Riedwyl, H (1988) *Multivariate Statistics: A Practical Approach*, Chapman and Hall, London.

Forcada, J, Trathan, PN, Reid, K, Murphy, EJ and Croxall, JP (2006) 'Contrasting population changes in sympatric penguin species in association with climate warming', *Global Change Biology*, **12**, pp. 411-423.

Fraser, WR and Patterson, DL (1997) 'Human disturbance and long-term changes in Adélie penguin populations: a natural experiment at Palmer Station, Antarctic Peninsula' in Battaglia, B, Valencia, J and Walton, D (eds.) *Antarctic Communities: Species, Structure and Survival*, Cambridge University Press, Cambridge.

Fraser, WR, Trivelpiece, WZ, Ainley, DG and Trivelpiece, SG (1992) 'Increases in Antarctic penguin populations: reduced competition with whales or a loss of sea ice due to environmental warming?', *Polar Biology*, **11**, pp. 525-531.

Gibson, LA, Wilson, BA, Cahill, DM and Hill, J (2004): "Spatial prediction of rufous bristlebird habitat in a coastal heathland: a GIS-based approach", *Journal of Applied Ecology*, **41**, 213-223

Giese, M (1996) 'Effects of human activity on Adélie penguin *Pygoscelis adeliae* breeding success, *Biological Conservation*, **75**, pp. 157-164.

Giese, M (1998) 'Guidelines for people approaching breeding groups of Adélie penguins *(Pygoscelis adeliae)'*, *Polar Record*, **34(191)**, pp. 287-292.

Glenz, C, Massolo, A, Kuonen, D and Schlaepfer, R (2001): "A wolf habitat suitability prediction study in Valais (Switzerland)", *Landscape and Urban Planning*, **55**, pp. 55-65.

Golden Software (2005) Grapher6.2.24 Graphing System, http://www.goldensoftware.com, Accessed 15/06/2006.

Greenberg, JD, Gergel, SE and Turner, MG (2002) 'Understanding landscape metrics II: effects of changes in scale', in Gergel, SE and Turner, MG (eds.) *Learning Landscape Ecology: A Practical Guide to Concepts and Techniques,* Springer, New York, pp. 101-111.

Greene, EM, Liston, GE and Pielke, RA Snr (1999) 'Simulation of above treeline snowdrift formation using a numerical snow-transport model', *Cold Regions Science and Technology*, **30**, pp. 135-144.

Gugolj, D., 2005, 'Three Dimension Snow Distribution Modeling in a GIS', University of Calgary, http://homepages.ucalgary.ca/~dgugolj/, accessed 25/7/2005.

Guisan, A and Zimmermann, NE (2000): "Predictive habitat distribution models in ecology", *Ecological Modelling*, **135**, 147-186.

Haehnel, R, Liston, G, Sturm, M and Lever, J (2001) 'GIS-based three-dimensional snowdrift computer model', *58th Eastern Snow Conference*, Ottawa, Ontario, Canada.

Hägeli, P and McClung, DM (2000) 'A new perspective on computer-aided avalanche forecasting: scale and scale Issues', *International Snow Science Workshop*, Big Sky, Montana, USA, October 2-6, 2000.

Hansen, MJ, Franklin, SE, Woudsma, CG and Peterson, M (2001) 'Caribou habitat mapping and fragmentation analysis using Landsat MSS, TM, and GIS data in the North Columbia Mountains, British Columbia, Canada', *Remote Sensing of Environment*, **77**, pp. 50-65.

Hastie, T, Tibshirani, R and Friedman, J (2001) *The Elements of Statistical Learning, Data Mining, Inference, and Prediction*, Springer-Verlag, New York, NY.

Hines, KM, Bromwich, DH and Marshall GJ (2000) 'Artificial surface pressure trends in the NCEP–NCAR reanalysis over the Southern Ocean and Antarctica', *Journal of Climate*, **13(22)**, pp. 3940-3952.

Holmes, N, Giese, M, Achurch, H, Robinson, S, Kriwoken, LK (2006) 'Behaviour and breeding success of Gentoo penguins *Pygoscelis papua* in areas of low and high human activity', *Polar Biology*, **29**, pp. 399-412.

International Association of Antarctic Tourism Operators (IAATO) (2006), 'www.iaato.org.

International Association of Geodesy (IAG), 2006, 'International Association of Geodesy', http://www.iag-aig.org, accessed 6/6/2006.

Ishikawa, M and Sawagaki, T (2001) 'GIS-simulation of the spatial distribution of snow cover and observed ground temperatures in the Daisetsu Mountains, Japan', *Norsk Geografisk Tidsskrift*, **55**, 212-218.

Jaedicke, C, Thiis, T, Sandvik, AD and Gjessing, Y (2000) 'Drifting snow in complex terrain – Comparison of measured snow distribution and simulated wind field', in Hjorth-Hansen, Holand, Løset and Norem (eds), *Snow Engineering*, Balkema, Rotterdam, pp. 65-73.

Kalnay, E, Kanamitsu, M, Kistler, R, Collins, W, Deavin, D, Gandin, L, Iredell, M, Saha, S, White, G, Woollen, J, Zhu, Y, Leetmaa, A, Reynolds, R, Chelliah, M, Ebisuzaki, W, Higgins, W, Janowiak, J, Mo, KC, Ropelewski, C, Wang, J, Jenne, R and Joseph, D (1996) 'The NCEP/NCAR 40-year reanalysis project', *Bulletin American Meteorological Society*, **77**, pp. 437-470.

Kato, A, Ropert-Coudert, Y and Naito, Y (2002) 'Changes in Adélie penguin breeding populations in Lützow-Holm Bay, Antarctica, in relation to sea-ice conditions', *Polar Biology*, **25**, pp. 934-938.

Kato, A, Watanabe, K and Naito, Y (2004) 'Population changes of Adélie and Emperor penguins along the Prince Olav Coast and on the Riiser-Larsen Peninsula', *Polar Bioscience*, **17**, pp. 117-122.

Keage, PL (1982) 'The location of Adélie penguin colonies, Windmill Islands, Antarctica', *Notornis* **24(4)**, pp. 339-340.

Kelly, NM, Fonseca, M and Whitfield, P (2001) 'Predictive mapping for management and conservation of seagrass beds in North Carolina', *Aquatic Conservation: Marine and Freshwater Ecosystems,* **11**, pp. 437-451.

Kent, S, Seddon, J, Robertson, G, and Wienecke, B (1998): Diet of Adélie Penguins *Pygoscelis adeliae* at Shirley Island, East Antarctica, January 1992, Marine Ornithology, **26**, pp. 7-10.

Kerry, K, Clarke, J, Eberhard, S., Gardner, H, Corsolini, S, Focardi, S, Franchi, E, Rodary, D and Thomson, R (1997) 'The Foraging Range of Adélie Penguins – Implications for CEMP and Interactions with the Krill Fishery', *CCAMLR Science,* **4**, pp. 75-87.

Kerry, KR, Clarke, JR and Else, GD (1995) 'The foraging range of Adélie penguins at Béchervaise

Island, MacRobertson Land, Antarctica as determined by satellite telemetry', in Dann, P, Norman, I and Reilly, P (eds.), *The Penguins,* Surrey Beatty and Sons, Sydney, pp. 216–243.

Kidd, DM and Ritchie, MG (2000) 'Inferring the patterns and causes of geographic variation in *Ephippiger ephippiger* (Orthoptera, Tettigoniidae) using geographical information systems (GIS)', *Biological Journal of the Linnean Society*, **71**, pp. 269-295.

Kind, RJ (1976) A critical examination of the requirements for model simulation of wind-induced erosion/deposition phenomena such as snow drifting, *Atmospheric Environment*, **10**, pp. 219-227.

Kind, R. J., 1981, 'Snow Drifting', *Handbook of Snow*, D. M. Gray and D. H. Male, Pergamon, Tarrytown, N. Y., pp 338 – 359.

Kind, RJ (1986) Snowdrifting: A Review of Modelling Methods, *Cold Regions Science and Technology*, **12**, pp. 217-228.

Kirkup, H, Melles, M and Gore, D. (2002): Late Quaternary environment of southern Windmill Islands, East Antarctica, *Antarctic Science*, **14**, pp. 385-394.

Kistler, R, Kalnay, E, Collins, W, Saha, S, White, G, Woollen, J, Chelliah, M, Ebisuzaki, W, Kanamitsu, M, Kousky, V, van den Dool, H, Jenne, R, and Fiorino, M, (2001) 'The NCEP/NCAR 50-Year Reanalysis', *Bulletin of the American Meteorological Society*, **82(2)**, pp. 247-267.

Lauver, CL, Busby, WH and Whistler, JL (2002): "Testing a GIS model of habitat suitability for a declining grassland bird", *Environmental Management*, **30(1)**, pp. 88-97.

Legendre, P (1993) 'Spatial autocorrelation: Trouble or new paradigm?', *Ecology*, **74(6)**, pp. 1659-1673.

Lehning, M, Bartelt, P, Brown, B, Russi, T, Stöckli, U and Zimmerli, M (1999) 'SNOWPACK model calculations for avalanche warning based upon a new network of weather and snow stations', *Cold Regions Science and Technology*, **30**, pp. 145-157.

Lehning, M, Doorschot, J, Raderschall, N and Bartelt, P (2000) 'Combining snow drift and SNOWPACK models to estimate snow loading in avalanche slopes', in Hjorth-Hansen, Holand, Løset and Norem (eds) *Snow Engineering*, Balkema, Rotterdam, pp.113-122.

Lehning, M, Fierz, C, Brown, B and Jamieson, B (2004) 'Modeling instability for the snow cover model SNOWPACK', *Annals of Glaciology*, **38**, pp. 331-338.

Lenton, SM, Fa, JE and Del Val, JP (2000): 'A simple non-parametric GIS model for predicting species distribution: endemic birds in Bioko Island, West Africa', *Biodiversity and Conservation*, **9**, pp. 869-885.

Levick, GM (1915) 'Natural history of the Adélie penguin', *British Antarctic ("Terra Nova") Expedition, 1910-1913, Natural History Reports (Zoology, Vol 1)*, pp. 55-84.

Linacre and Geerts (1999) 'Roughness length', University of Wyoming, http://www.da.uwyo.edu/~geerts/cwx/notes/chap14/roughness.html, accessed 25/07/2005.

Ling, HU and Seppelt, RD (1998) 'Non-marine algae and cyanobacteria of the Windmill Islands region, Antarctica, with descriptions of two new species', *Archiv für Hydrobiologie Supplement 124, Algological Studies,* **89**, pp. 49-62.

Liston, GE and Sturm, M (1998) 'A snow-transport model for complex terrain', *Journal of Glaciology*, **44(148)**, pp. 498-516.

Lucieer, A (2005) *Interpolated Digital Elevation Model of Whitney Point, Windmill Islands, East Antarctica*, University of Tasmania.

Lucieer, A (2005) *Interpolated Digital Elevation Model of the Windmill Islands, East Antarctica*, University of Tasmania.

Lucieer, A (2005) *Grid2Col tool for ArcGIS 9.1*, University of Tasmania.

Lucieer, A (2006) *Interpolated Digital Elevation Model of Shirley Island, Windmill Islands, East Antarctica*, University of Tasmania.

Manel, S, Dias, J-M and Ormerod, SJ (1999): Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: A case study with a Himalayan river bird", *Ecological Modelling*, **120(2-3)**, 337-347.

Martin, M, Johnstone, G and Woehler, E (1990) 'Increased numbers of Adélie penguins (*Pygoscelis adeliae*) breeding near Casey, Wilkes Land, East Antarctica', *Corella*, **14**, pp. 119-122.

Maurer, BA (1995) *Methods in Ecology: Geographical Population Analysis: Tools for the Analysis of Biodiversity*, Blackwell Science Ltd., Oxford.

McRae, CF, Hocking, AD, Seppelt, RD (1999) '*Penicillium* species from terrestrial habitats in the Windmill Islands, East Antarctica, including a new species *Penicillium antarcticum*', *Polar Biology*, **21**, pp. 97-111.

Melick, DR, Hovenden, MJ, Seppelt, RD (1994) Phytogeography of bryophyte and lichen vegetation in the Windmill Islands, Wilkes Land, Continental Antarctica. *Plant Ecology*, **111(1)**, pp. 71-87.

Michaux, J-L, Naaim-Bouvet, F, Naaim, M, Lehning, M and Guyomarc'h, G (2002) 'Effect of unsteady wind on drifting snow: first investigations', *Natural Hazards and Earth System Sciences*, **2**, pp. 129-136.

Micol, T and Jouventin, P (2001) 'Long-term population trends in seven Antarctic seabirds at Pointe Géologie (Terre Adélie): Human impact compared with environmental change', *Polar Biology*, **24**, pp. 175-185.

Moczydlowski, E (1986) 'Microclimate of the nest-sites of pygoscelid penguins (Admiralty Bay, South Shetland Islands)', *Polish Polar Research*, **7(4)**, pp. 377-394.

Moczydlowski, E (1989) 'Protection of eggs and chicks against flooding as a part of nesting strategy of pygoscelid penguins at King George Island, South Shetlands', *Polish Polar Research*, **10(2)**, pp. 163-181.

Morgan, PJ (2005) *Ground control data for aerial photographs of Whitney Point, Windmill Islands,* *East Antarctica*, University of Tasmania.

Murray, MD and Luders, DJ (1990) 'Faunistic studies at the Windmill Islands, Wilkes Land, East Antarctica, 1959-80', *ANARE Research Notes,* **73**.

Olivier, F., Lee, AV and Woehler, EJ (2004) 'Distribution and abundance of snow petrels *Pagodroma nivea* in the Windmill Islands, East Antarctica', *Polar Biology,* **27**, pp. 257-265.

Olmastroni, S, Pezzo, F, Volpi, V and Focardi, S (2004) 'Effects of weather and sea-ice on the reproductive performance of the Adélie penguin at Edmonson Point, Ross Sea', *CCAMLR Science*, **11**, pp. 99-109.

Orndorff, R. L. and Van Hoesen, J. G., 2001, 'Using GIS to predict the spatial distribution of perennial snow under modern and late Pleistocene conditions in the Snake Range, Nevada', *69th Annual Meeting of the Western Snow Conference, 2001*, Las Vegas, Nevada, pp. 119-122.

Orton, MN (1963) 'A brief survey of the fauna of the Windmill Islands, Wilkes Land, Antarctica', *The Emu,* **63(1)**, pp. 14-22.

Osborne, PE, Alonso, JC and Bryant, RG (2001): 'Modelling landscape-scale habitat use using GIS and remote sensing: a case study with great bustards', *Journal of Applied Ecology*, **38**, pp. 458-471.

Patterson, DL, Easter-Pilcher, AL and Fraser WR (2003): 'The effects of human activity and environmental variability on long-term changes in Adélie penguin populations at Palmer Station, Antarctica' in Huiskes, AHL, Gieskes, WWC, Rozema, J, Schorno, RML, van der Vies, SM and Wolf, WJ (eds.) *Antarctic Biology in a Global Context*, Backhuys Publishers, Leiden, Netherlands.

Pearsall, J (ed.) (1999) *Concise Oxford Dictionary, Tenth Edition*, Oxford University Press, Oxford.

Penney, RL (1968) 'Territorial and Social Behaviour in the Adélie Penguin', in Austin, OL Jnr (ed.) *Antarctic Bird Studies. Antarctic Research Series,* **12,** American Geophysical Union, Washington DC, pp. 83-131.

Pereira, JMC and Duckstein, L (1993) 'A multiple criteria decision-making approach to GIS-based land suitability evaluation', *International Journal of Geographical Information Systems*, **7(5)**, pp. 407-424.

Pereira, JMC and Itami, RM (1991) 'GIS-based habitat modeling using logistic multiple-regression – A study of the Mt Graham red squirrel', *Photogrammetric Engineering and Remote Sensing*, **57(11)**, pp. 1475-1486.

Petz, W (1997): Ecology of the active soil microfauna (Protozoa, Metazoa) of Wilkes Land, East Antarctica, *Polar Biology*, 18:1, pp. 33-44.

Pfeiffer, S and Peter, H-U (2004) 'Ecological studies toward the management of an Antarctic tourist landing site (Penguin Island, South Shetland Islands)', *Polar Record*, **40(215)**, pp. 345-353.

Pomeroy, JW, Gray, DM and Landine, PG (1993) 'The prairie blowing snow model: characteristics, validation, operation', *Journal of Hydrology*, **144**, pp.16-192.

Pomeroy, JW, Marsh, P and Gray, DM (1997) 'Application of a distributed blowing snow model to the Arctic', *Hydrological Processes*, **11**, pp. 1451-1464.

Purves, RS, Barton, JS, Mackaness, WA and Sugden, DE (1998) 'The development of a rule-based spatial model of wind transport and deposition of snow', *Annals of Glaciology*, **26**, pp. 197-202.

Quinlan, JR (1993) *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, Inc., San Francisco, CA.

Quinn, GP and Keough, MJ (2002) *Experimental Design and Data Analysis for Biologists*, Cambridge University Press, Cambridge, UK,

Raphael, MG, Young, JA and Galleher, BM (1995) 'A landscape-level analysis of marbled murrelet habitat in western Washington', *USDA Forest General Technical Report*, **PSW-152**, pp.177-191.

Robertson, G, Kent, S and Seddon, J (1994) Effects of the water-offloading technique on Adélie penguins, *Journal of Field Ornithology*, **65**, pp. 376–380.

Ropert-Coudert, Y, Wilson, RP, Daunt, F and Kato, A (2004) 'Patterns of energy acquisition by a central place forager: benefits of alternating short and long foraging trips', *Behavioural Ecology*, **15(5)**, pp. 824-830.

Ryan, B. C., 1977, 'A mathematical model for diagnosis and prediction of surface wind in mountainous terrain', *Journal of Applied Metrology*, **16**, pp. 571 -584.

Ryan, U. (1999, updated 2006) 'Windmill Islands 1:50000 GIS Dataset', *Australian Antarctic Data Centre - CAASM Metadata,* http://aadc-db.aad.gov.au/KeywordSearch/Home.do?Portal=amd_au&MetadataType=0, accessed 8/7/2006.

Ryan, U. and Woehler, E. (2000, updated 2006) 'Windmill Islands penguin colonies digitised from 1990 Linhof aerial photography', *Australian Antarctic Data Centre - CAASM Metadata,* http://aadc-db.aad.gov.au/KeywordSearch/Home.do?Portal=amd_au&MetadataType=0, accessed 10/11/2006.

SAS Institute (1989-2003): *JMP 5.1*, www.jmp.com, accessed 21/9/2006.

Seppelt, RD, and Connell, D, 2006, *Windmill Islands*, Australian Antarctic Division, http://www.aad.gov.au/default.asp?casid=1979, accessed 5/8/2006.

Sieg, CH and Becker, DM (1990) 'Nest-site habitat selected by merlins in southeastern Montana', *The Condor*, **92**, pp. 688-694.

Store, R and Kangas, J (2001): "Integrating spatial multi-criteria evaluation and expert knowledge for GIS-based habitat suitability modelling", *Landscape and Urban Planning*, **55**, pp. 79-93.

Tappeiner, U, Tappeiner, G, Aschenwald, J, Tasser, E and Ostendorf, B (2001) 'GIS-Based Modelling of Spatial Pattern of Snow Cover Duration in an Alpine Area', *Ecological Modelling*, **138**, pp. 265-275.

Tobler, WR (1970) 'A computer model simulation of urban growth in the Detroit region', *Economic Geography*, **46(2)**, pp. 234-240.

Trivelpiece, WZ and Fraser, WR (1996) 'The breeding biology and distribution of Adélie penguins:

adaptations to environmental variability', *Foundations for Ecological Research West of the Antarctic Peninsula, Antarctic Research Series*, **70**, American Geophysical Union, pp. 273-285.

Uematsu, T, Nakata, T, Takeuchi, K, Arisawa, Y and Kaneda, Y (1991) 'Three-dimensional numerical simulation of snowdrift', *Cold Regions Science and Technology*, **20**, pp. 65-73.

Van Dam, O (2001) 'Forest filled with gaps: effects of gap size on water and nutrient cycling in tropical rain forest: a study in Guyana', PhD thesis, Universiteit Utrecht, http://igitur-archive.library.uu.nl/dissertations/1954687/inhoud.htm, accessed 20/07/2006.

van Deursen, W, Wesseling, C, Burrough, P, Karssenberg, D, Pebesma, E and de Jong, K (2006), PCRaster Version 2, http://pcraster.geo.uu.nl, accessed 18/11/2006.

VirtuoZo v3.6 (2004), www.supresoft.com.

Vogiatzakis, IN (2003): "GIS-based modelling and ecology: a review of tools and methods", *Geographical Paper No.170*, University of Reading, UK.

Voous, KH (1965) 'Antarctic Birds', in van Oye, P and van Mieghem, J (eds.) *Biogeography and Ecology in Antarctica*, Junk, The Hague, Holland.

Walker, R and Craighead, L (1997): "Least-cost-path corridor analysis: analyzing wildlife movement corridors in Montana using GIS", *Proceedings ESRI User's Conference, 1997.*

Wallace, L (2005) 'GIS based snow accumulation model for the Windmill Islands', Geomatics Studio Project, University of Tasmania, Hobart.

Walter, MT, McCool, DK, King, LG, Molnau, M and Campbell, GS (2004) 'Simple snowdrift model for distributed hydrological modeling', *Journal of Hydrologic Engineering*, **July/August 2004**, pp. 280-287.

Watanuki, Y, Kato, A, Naito, Y, Robertson, G and Robinson, S (1997) 'Diving and foraging behaviour of Adélie penguins in areas with and without fast sea-ice', *Polar Biology*, **17**, pp. 296-304.

Weka Manual (2006) http://grb.mnsu.edu/grbts/doc/manual/J48_Decision_Trees.html, accessed 20/11/2006.

Whittow, JB (2000) *The Penguin Dictionary of Physical Geography*, Penguin Books, London.

Wienecke, BC, Lawless, R, Rodary, D, Bost, C-A, Thomson, R, Pauly, T, Robertson, G, Kerry KR and LeMaho Y (2000) 'Adélie penguin foraging behaviour and krill abundance along the Wilkes and Adélie land coasts, Antarctica', *Deep Sea Research Part II: Topical Studies in Oceanography*, **47(12)**, pp. 2573-2587.

Wilson, RP, Coria, NR, Spairani, HJ, Adelung, D and Culik, B (1989) 'Human-induced behaviour in Adélie penguins *Pygoscelis adeliae*', *Polar Biology*, **10**, pp.77-80.

Wilson, RP, Culik, B, Danfield, R and Adelung, D (1991) 'People in Antarctica – How much do Adélie penguins *Pygoscelis adeliae* care?', *Polar Biology*, **11**, pp. 363-370.

Wilson, K-J, Taylor, RH and Barton, KJ (1990) 'The impact of man on Adélie penguins at Cape Hallett, Antarctica', in Kerry, KR and Hempel, G (eds.) *Antarctic Ecosystems. Ecological Change and Conservation*, Springer-Verlag Berlin, Heidelberg, pp. 183-190.

Witten, IH and Frank, E (2005) *Data Mining: Practical machine learning tools and techniques*, 2nd Edition, Morgan Kaufmann, San Francisco.

Wittenberger, JF and Hunt, GL Jr. (1985) 'The adaptive significance of coloniality in birds', in Farner, DS, King, JR and Parkes, KC (eds.), *Avian Biology*, **VIII**, Academic Press Inc, Orlando.

Woehler, EJ (1993): "Antarctic Seabirds: Their status and conservation in the AAT", *RAOU Conservation Statement No. 9,* Supplement to *Wingspan,* **12**.

Woehler, EJ, Cooper, J, Croxall, JP, Fraser, WR, Kooyman, GL, Miller, GD, Nel, DC, Patterson, DL, Peter, H-U, Ribic, CA, Salwicka, K, Trivelpiece, WZ and Weimerskirch, H (2001) 'A statistical assessment of the status and trends of Antarctic and Subantarctic seabirds', *Report on SCAR BBS Workshop on Southern Ocean seabird populations*, SCAR 2001.

Woehler, EJ & Croxall, JP (1997) 'The status and trends of Antarctic and sub-Antarctic seabirds', *Marine Ornithology,* **25**, pp. 43–66.

Woehler, EJ, Slip, DJ, Robertson, LM, Fullagar, PJ and Burton, HR (1991) 'The distribution, abundance and status of Adélie penguins *Pygoscelis adeliae* at the Windmill Islands, Wilkes Land, Antarctica', *Marine Ornithology*, **19(1)**, pp. 1-18.

Woehler, EJ, Penney, RL, Creet, SM and Burton, HR (1994) 'Impacts of human visitors on breeding success and long-term population trends in Adélie penguins at Casey, Antarctica', *Polar Biology*, **14**, pp. 269-274.

Yeates, GW (1968) 'Studies on the Adélie penguin at Cape Royds 1964-65 and 1965-66', *New Zealand Journal of Marine and Freshwater Research*, **2(3)**, pp. 472-496.

Yeates, GW (1975) 'Microclimate, climate and breeding success in Antarctic penguins', in Stonehouse, B (ed.) *The Biology of Penguins*, Macmillan Press Ltd, London, pp. 397-409.

Young, EC (1990) 'Long-term stability and human impact in Antarctic skuas and Adélie penguins', in Kerry, KR and Hempel, G (eds.) *Antarctic Ecosystems. Ecological Change and Conservation*, Springer-Verlag, Berlin, Heidelberg, pp. 231-236.

Zar, JH (1999) *Biostatistical Analysis 4th Ed.*, Prentice-Hall, Englewood Cliffs, New Jersey.

Appendix 1: Summary Statistics

Summary statistics for Whitney Pt Adélie penguin colony distribution

			ELEVATION	WIND EXP	PLAN	PROFILE	ROUGH NORM	ROUGH STDEV	SLOPE	SNOW 2005	SNOW DIFF	SOLAR	WETNESS
Individual Cell Values	Colonies	MEAN	14.89	57.27	0	0	2.31	0.36	0.2	0.5	0.04	3159.58	4.32
		MEDIAN	14.62	57	0	0	1.73	0.32	0.18	0.3	0.08	3154.47	3.89
		MIN	3.99	0	-0.54	-0.83	1.54	0.03	0	0	-2.64	1719.34	1.82
		MAX	29.82	175	0.83	0.62	351.47	1.67	0.92	8.74	2.9	4022.13	12.9
		ST DEV	1567.9	1592.38	1716.45	1729.66	6.14	6.13	6.12	0.57	0.65	1503	1502.11
		KURTOSIS	0.49	-0.08	13.89	13.89	1093.5	4.41	2.88	39.06	5.99	0.13	2.89
	Control Plots	MEAN	16.74	60.54	0	0	2.06	0.54	0.3	0.64	0.01	3061.2	4.07
		MEDIAN	16.4	57	-0.01	0	1.7	0.43	0.24	0.38	0.07	3197.75	3.72
		MIN	5.85	0	-0.84	-1.06	1.57	0.05	0.02	0	-7.7	1324.21	1.19
		MAX	27.36	198	1.07	0.74	22.7	2.19	1.35	11.77	5.22	4044.05	10.18
		ST DEV	3.93	49.06	0.13	0.17	1.46	0.36	0.22	1.08	0.89	606.51	1.63
		KURTOSIS	0.02	-0.32	9.09	8.75	77.26	2.68	3.39	52.21	28.39	-0.15	1.49
Mean Values	Colonies	MEAN	16.49	61.79	0	0	2.8	0.4	0.22	0.46	0.06	3135.79	4.23
		MEDIAN	15.58	58.93	0	0	1.8	0.39	0.22	0.35	0.03	3099.24	4.14
		MIN	6.22	13.13	-0.1	-0.13	1.62	0.12	0.01	0	-0.68	2446.3	2.75
		MAX	29.24	126	0.1	0.11	21.04	0.85	0.48	1.63	0.79	3824.11	7.91
		ST DEV	5.75	30.93	0.04	0.05	3.76	0.15	0.09	0.38	0.34	327.82	0.97
		KURTOSIS	-0.1	-0.74	1.18	1.1	18.14	1.03	1.22	1.89	0.4	-0.35	4.44
	Control Plots	MEAN	16.75	60.52	0	0	2.06	0.54	0.3	0.64	0.01	3061.47	4.07
		MEDIAN	16.32	57.86	0	0	1.86	0.47	0.25	0.55	0	3183.4	4.15
		MIN	7.58	0	-0.13	-0.13	1.64	0.16	0.09	0	-1.41	1813.2	2.21
		MAX	26.6	174.84	0.07	0.09	4.13	1.53	0.91	3.91	0.76	3824	5.06
		ST DEV	3.86	39.68	0.03	0.04	0.51	0.31	0.18	0.65	0.41	517.5	0.65
		KURTOSIS	0.06	0.55	5.79	2.13	5.8	2.16	3.45	15.22	2.9	-0.37	1.14

Summary statistics for Shirley I Adélie penguin colony distribution

			ELEVATION	WIND EXP	PLAN	PROFILE	ROUGH NORM	ROUGH STDEV	SLOPE	SNOW 2005	SNOW DIFF	SOLAR	WETNESS
Individual Cell Values	Colonies	MEAN	20.7	50.37	0	0	1.67	0.15	0.09	0.92	0.2	3382.48	5.58
		MEDIAN	15.55	50	0	0	1.64	0.13	0.08	0.58	0.09	3381.23	5.22
		MIN	3.13	0	-0.07	-0.05	1.59	0.01	0	0	-4.72	2958.55	2.86
		MAX	35.46	101	0.04	0.07	6.96	0.41	0.25	14.05	8.04	3769.61	12.27
		ST DEV	10.18	15.18	0.01	0.01	0.17	0.07	0.05	1.62	0.84	120.37	1.52
		KURTOSIS	-1.77	0.85	3.2	3.92	411.32	-0.06	-0.06	24.12	21.66	0.1	2.38
	Control Plots	MEAN	14.43	48.77	0	0	1.69	0.27	0.17	1.03	0.02	3330	5.21
		MEDIAN	11.28	48	0	0	1.64	0.22	0.13	0.36	0	3309.12	4.94
		MIN	1.51	0	-0.09	-0.16	1.58	0.01	0	0	-9.22	2377.26	1.89
		MAX	34	153	0.06	0.09	9.67	1.34	0.82	20.6	7.22	4120.81	15.38
		ST DEV	9.04	32.48	0.01	0.02	0.29	0.22	0.13	2.41	1.14	310.99	1.57
		KURTOSIS	-1.28	-0.68	4.31	4.18	307.32	4.48	29.14	25.74	29.14	0.28	5.02
Mean Values	Colonies	MEAN	17.24	45.41	-0.01	0.01	1.7	0.15	0.09	1.43	0.26	3385.16	5.06
		MEDIAN	12.48	46.57	0	0.01	1.66	0.14	0.09	0.74	0.02	3384.63	4.93
		MIN	3.2	3	-0.03	-0.01	1.63	0.04	0.02	0	-1.99	3043.7	3.56
		MAX	33.14	85.17	0.01	0.04	2.12	0.34	0.21	8.86	4.77	3698.97	7.67
		ST DEV	10.9	15.98	0.01	0.01	0.09	0.07	0.04	2	0.89	139.78	0.88
		KURTOSIS	-1.72	0.48	0.91	3.56	8.91	0.2	0.1	5.23	12.47	-0.15	0.63
	Control Plots	MEAN	14.74	46.92	0	0	1.69	0.27	0.16	1.06	0.03	3337.62	5.21
		MEDIAN	11.72	48.22	0	0	1.65	0.21	0.13	0.59	0	3298.62	5.06
		MIN	2.62	0	-0.02	-0.02	1.63	0.03	0.02	0	-1.99	2619	3.17
		MAX	32.16	111.71	0.01	0.03	2.06	1.11	0.68	6.48	1.02	4047.78	9.17
		ST DEV	9.11	29.99	0.01	0.01	0.09	0.2	0.12	1.35	0.5	292.99	1.13
		KURTOSIS	-1.38	-0.76	0.57	0.94	6.3	4.9	4.87	5.56	4.91	0.48	1.43

Summary statistics for Whitney Pt Adélie penguin colony population trend classes

			SnowDiff	Snow2005	WindExp	Casey Dis
INDIVIDUAL CELL VALUES	TREND 1 (n=13)	MEAN	0.22	0.47	77.69	3464.04
		MEDIAN	0.47	0.6	81	3465.63
		MIN	-0.52	0.12	47	3450.77
		MAX	0.67	0.67	89	3471.59
		ST DEV	0.42	0.21	12.63	6.25
		KURTOSIS	-0.76	-1.18	2.01	0.87
	TREND 2 (n=8)	MEAN	0.29	0.3	85.75	3266.85
		MEDIAN	0.24	0.25	86.5	3267.1
		MIN	0.13	0.16	64	3264.1
		MAX	0.47	0.47	100	3268.12
		ST DEV	0.17	0.16	10.75	1.48
		KURTOSIS	-2.51	-2.47	2.03	-0.14
	TREND 3 (n=141)	MEAN	-0.01	1.21	20.38	3316.45
		MEDIAN	0.11	0.41	19	3300.81
		MIN	-2.48	0	0	3262.27
		MAX	1.51	8.74	110	3615.75
		ST DEV	0.49	2.14	21	60.81
		KURTOSIS	4.43	4.77	3.34	11.87
	TREND 4 (n=29)	MEAN	-0.31	1.21	38	3544.35
		MEDIAN	-0.05	0.54	30	3544.13
		MIN	-2.15	0	0	3540.56
		MAX	0.85	5.51	120	3549.1
		ST DEV	0.9	1.78	37.3	2.52
		KURTOSIS	0.12	2.58	0.86	-0.88
	TREND 5 (n=847)	MEAN	0.07	0.37	61.81	3420.52
		MEDIAN	0.1	0.28	61	3415.65
		MIN	-2.64	0	0	3270.35
		MAX	2.9	5.56	175	3665.54
		ST DEV	0.47	0.39	30.9	109.04
		KURTOSIS	7.04	46.7	0.51	-1.12

			SnowDiff	Snow2005	WindExp	Ice Dist	Casey Dist
MEAN VALUES	TREND 1	MEAN	0.41	1.59	46.69	610.94	1350.6
		MEDIAN	0.02	0.71	47.83	679.26	1446.53
		MIN	-1.99	0	20	159.82	857.39
		MAX	4.77	8.86	74.33	978.32	1766.43
		ST DEV	1.46	2.33	16.53	227.03	285.58
		KURTOSIS	6	6.84	-0.56	-0.2	-0.69
	TREND 2	MEAN	0.42	0.83	41.28	406.1	1080.85
		MEDIAN	0	0.04	40.16	333.47	997.84
		MIN	-0.01	0	27.2	220.03	818.71
		MAX	2.16	2.84	52.33	674.22	1440.85
		ST DEV	0.81	1.21	7.68	191.45	278.62
		KURTOSIS	2.79	-0.78	0.99	-1.63	-2.11
	TREND 3	MEAN	-0.07	1.26	48.23	546.52	1277.58
		MEDIAN	0.01	0.02	45.5	588.23	1341.73
		MIN	-1.72	0	15.18	221.62	849.4
		MAX	0.6	8.24	85.17	760.04	1546.64
		ST DEV	0.65	2.67	19.75	179.97	255.05
		KURTOSIS	6.83	8.05	1.03	0.03	0
	TREND 4	MEAN	0.02	1.21	30.58	812.86	1602.67
		MEDIAN	0.02	1.21	30.58	812.86	1602.67
		MIN	-0.03	1.19	3	773.68	1563.68
		MAX	0.07	1.22	58.15	852.03	1641.66
		ST DEV	0.07	0.02	39	55.4	55.14
		KURTOSIS	N/A	N/A	N/A	N/A	N/A
	TREND 5	MEAN	2.31	2.31	42.24	744.96	1523.66
		MEDIAN	1.77	1.77	45.25	861.67	1649.56
		MIN	0.71	0.71	14.33	196.68	949.26
		MAX	5.89	5.89	63.52	930.09	1718.65
		ST DEV	1.7	1.7	13.77	235.49	251.35
		KURTOSIS	0.69	0.69	1.18	2.58	2.23

Summary statistics for Shirley I Adélie penguin colony population trend classes

			SnowDiff	Snow2005	WindExp	Ice Dist	Casey Dist
INDIVIDUAL CELL VALUES	TREND 1 (n=188)	MEAN	0.18	1.02	44.56	607.05	1338.53
		MEDIAN	0.02	0.32	45	678.73	1443.92
		MIN	-2.68	0	9	150.33	840.92
		MAX	4.77	12.37	82	1108.34	1896.85
		ST DEV	0.92	1.83	15.96	266.26	329.4
		KURTOSIS	6.61	16.29	-0.37	-1.08	-1.33
	TREND 2 (n=117)	MEAN	0.12	0.38	40.49	312.81	942.43
		MEDIAN	0	0	40	242.95	823.86
		MIN	-1.46	0	16	149.31	791.25
		MAX	2.16	3.84	64	828.61	1610.33
		ST DEV	0.6	0.93	7.25	180.7	246.29
		KURTOSIS	4.67	6.92	1.27	1.48	1.39
	TREND 3 (n=789)	MEAN	0.19	1.02	52.88	690.02	1466.44
		MEDIAN	0	0.6	54	776.1	1564.12
		MIN	-4.72	0	2	148.5	844.33
		MAX	8.04	12.85	101	868.75	1658.03
		ST DEV	1.1	1.84	15.05	172.85	192.12
		KURTOSIS	17.47	15.64	0.15	2.72	2.97
	TREND 4 (n=275)	MEAN	0.05	1.2	49.86	841.05	1630.72
		MEDIAN	-0.02	0.86	54	852.21	1642.19
		MIN	-1.72	0.3	0	484.26	1274.19
		MAX	7.04	14.05	87	1032.01	1822
		ST DEV	0.87	1.58	23.9	126.31	126.27
		KURTOSIS	23.05	51.68	-0.15	-1.12	-1.12
	TREND 5 (n=755)	MEAN	0.25	1.46	51.38	867.36	1654.03
		MEDIAN	0.17	0.77	51	886.09	1674.48
		MIN	-4.46	0	2	158.42	907.89
		MAX	4.99	12.18	96	1108.17	1896.75
		ST DEV	1.05	2.14	18.19	171.9	179.24
		KURTOSIS	9.39	10.68	-0.74	7.69	8.35

			SnowDiff	Snow2005	WindExp	Ice Dist	Casey Dist
MEAN VALUES	TREND 1	MEAN	0.41	1.59	46.69	610.94	1350.6
		MEDIAN	0.02	0.71	47.83	679.26	1446.53
		MIN	-1.99	0	20	159.82	857.39
		MAX	4.77	8.86	74.33	978.32	1766.43
		ST DEV	1.46	2.33	16.53	227.03	285.58
		KURTOSIS	6	6.84	-0.56	-0.2	-0.69
	TREND 2	MEAN	0.42	0.83	41.28	406.1	1080.85
		MEDIAN	0	0.04	40.16	333.47	997.84
		MIN	-0.01	0	27.2	220.03	818.71
		MAX	2.16	2.84	52.33	674.22	1440.85
		ST DEV	0.81	1.21	7.68	191.45	278.62
		KURTOSIS	2.79	-0.78	0.99	-1.63	-2.11
	TREND 3	MEAN	-0.07	1.26	48.23	546.52	1277.58
		MEDIAN	0.01	0.02	45.5	588.23	1341.73
		MIN	-1.72	0	15.18	221.62	849.4
		MAX	0.6	8.24	85.17	760.04	1546.64
		ST DEV	0.65	2.67	19.75	179.97	255.05
		KURTOSIS	6.83	8.05	1.03	0.03	0
	TREND 4	MEAN	0.02	1.21	30.58	812.86	1602.67
		MEDIAN	0.02	1.21	30.58	812.86	1602.67
		MIN	-0.03	1.19	3	773.68	1563.68
		MAX	0.07	1.22	58.15	852.03	1641.66
		ST DEV	0.07	0.02	39	55.4	55.14
		KURTOSIS	N/A	N/A	N/A	N/A	N/A
	TREND 5	MEAN	2.31	2.31	42.24	744.96	1523.66
		MEDIAN	1.77	1.77	45.25	861.67	1649.56
		MIN	0.71	0.71	14.33	196.68	949.26
		MAX	5.89	5.89	63.52	930.09	1718.65
		ST DEV	1.7	1.7	13.77	235.49	251.35
		KURTOSIS	0.69	0.69	1.18	2.58	2.23

Appendix 2: Discriminant Analysis Formulae

Adélie penguin colony distribution models:

Whitney Pt

Individual Cell Values:

SqDist[0] = 0.0534627677653561 * [Elevation] * [Elevation] - 0.0925164711997984 * [Roughness (St Dev)] * [Elevation] + 20.7029979809649 * [Roughness (St Dev)] * [Roughness (St Dev)] - 0.0000688508027802848 * [Solar Radiation] * [Elevation] + 0.0152541417947432 * [Solar Radiation] * [Roughness (St Dev)] + 0.00000708051487295363 * [Solar Radiation] * [Solar Radiation]

SqDist[Absent] = [SqDist_0] - 1.53576999254331 * [Elevation] - 67.259504138339 * [Roughness (St Dev)] - 0.0504878121571665 * [Solar Radiation] + 108.363225246094

SqDist[Present] = [SqDist_0] - 1.33252540810463 * [Elevation] - 61.4041467733707 * [Roughness (St Dev)] - 0.0490763238411542 * [Solar Radiation] + 98.1724887963393

Prob[0] = Exp(-0.5 * [SqDist_Absent]) + Exp(-0.5 * [SqDist_Pres])

Prob[Absent] = Exp(-0.5 * [SqDist_Absent]) / [Prob_0]

Prob[Present] = Exp(-0.5 * [SqDist_Pres]) / [Prob_0]

Covariance Matrices

Within Cov	Elevation	Roughness (St Dev)	Solar Radiation
Elevation	18.765722	0.0137889	76.385484
Roughness (St Dev)	0.0137889	0.0800923	-86.20766
Solar Radiation	76.385484	-86.20766	234466.22
Within Corr	Elevation	Roughness (St	Solar Radiation

		Dev)	
Elevation	1	0.0112474	0.0364156
Roughness (St Dev)	0.0112474	1	-0.629087
Solar Radiation	0.0364156	-0.629087	1

Colony/Control Plot Mean Values

Sqdist[0] = 0.00123585320191969 * [Wind Exposure] * [Wind Exposure] + 0.00502152653063165 * [Roughness (Norm)] * [Wind Exposure] + 0.134419706830413 * [Roughness (Norm)] * [Roughness (Norm)] - 0.16918291609256 * [Roughness (St Dev)] * [Wind Exposure] - 0.668261971951428 * [Roughness (St Dev)] * [Roughness (Norm)] + 30.8273619561492 * [Roughness (St Dev)] * [Roughness (St Dev)] + 0.00241480625010695 * [Snow Difference] * [Wind Exposure] - 0.327590123794121 * [Snow Difference] * [Roughness (Norm)] - 4.32336043329957 * [Snow Difference] * [Roughness (St Dev)] + 7.67254372659123 * [Snow Difference] * [Snow Difference] + 0.0153650948338209 * [Wetness Index] * [Wind Exposure] - 0.182614364124296 * [Wetness Index] * [Roughness (Norm)] + 6.45260097269701 * [Wetness Index] * [Roughness (St Dev)] - 1.70088427916617 * [Wetness Index] * [Snow Difference] + 2.08267825583303 * [Wetness Index] * [Wetness Index]

SqDist[Absent] = [SqDist_0] - 0.13587102626854 * [Wind Exposure] + 0.250412667774322 * [Roughness (Norm)] - 48.2253944544141 * [Roughness (St Dev)] + 10.0453557821323 * [Snow Difference] - 21.4663798042278 * [Wetness Index] + 61.7725804826357

SqDist[Present] = [SqDist_0] - 0.163644699622735 * [Wind Exposure] - 0.0199630328365665 * [Roughness (Norm)] - 40.0792313894895 * [Roughness (St Dev)] + 8.75502877854204 * [Snow Difference] - 20.6689377579887 * [Wetness Index] + 56.9842678626558

Prob[0] = Exp(-0.5 * [SqDist_Absent]) + Exp(-0.5 * [SqDist_Presnt])
Prob[Absent] = Exp(-0.5 * [SqDistAbsent]) / [Prob_0]

Prob[Present] = Exp(-0.5 * [SqDistPresent]) / [Prob_0]

Covariance Matrices

Within Cov	Wind Exposure	Roughness (Norm)	Roughness (St	Wetness Index

			Dev)	
Wind Exposure	1236.2676	-21.71073	4.4296762	-12.70849
Roughness (Norm)	-21.71073	8.5056587	-0.009638	0.5680598
Roughness (St Dev)	4.4296762	-0.009638	0.0553056	-0.101015
Wetness Index	-12.70849	0.5680598	-0.101015	0.7358718

Within Corr	Wind Exposure	Roughness (Norm)	Roughness (St Dev)	Wetness Index
Wind Exposure	1	-0.211721	0.5357118	-0.421344
Roughness (Norm)	-0.211721	1	-0.014053	0.2270591
Roughness (St Dev)	0.5357118	-0.014053	1	-0.500727
Wetness Index	-0.421344	0.2270591	-0.500727	1

Shirley I

Individual Cell Values

SqDist_0 = 0.0113377884643721 * [Elevation] * [Elevation] - 4.82851112342201 * [Roughness (St Dev)] * [Elevation] + 74832.7309135573 * [Roughness (St Dev)] * [Roughness (St Dev)] + 8.02386428357845 * [Slope] * [Elevation] - 242600.998925054 * [Slope] * [Roughness (St Dev)] + 196767.480547947 * [Slope] * [Slope] - 0.0205322880802495 * [Snow 2005] * [Elevation] - 51.5057614003503 * [Snow 2005] * [Roughness (St Dev)] + 79.3789462336679 * [Snow 2005] * [Slope] + 0.309547430071808 * [Snow 2005] * [Snow 2005] - 0.00010016311324339 * [Solar Radiation] * [Elevation] - 0.204754187148813 * [Solar Radiation] * [Roughness (St Dev)] + 0.240553758076989 * [Solar Radiation] * [Slope] + 0.000438272417241657 * [Solar Radiation] * [Snow 2005] + 0.0000748771944452865 * [Solar Radiation] * [Solar Radiation] - 0.000889025855991192 * [Wind Exposure] * [Elevation] - 1.4307857133681 * [Wind Exposure] * [Roughness (St Dev)] + 1.59214065290944 * [Wind Exposure] * [Slope] + 0.0146210373872546 * [Wind Exposure] * [Snow 2005] + 0.00114857449769329 * [Wind Exposure] * [Solar Radiation] + 0.00605288069706808 * [Wind Exposure] * [Wind Exposure]

SqDist[Absent] = [SqDist_0] + 0.0573220773726119 * [Elevation] + 280.824839256881 * [Roughness (St Dev)] - 162.184083778947 * [Slope] - 1.62884466997249 * [Snow 2005] - 0.537634531946183 * [Solar Radiation] - 4.29046318269748 * [Wind Exposure] + 975.155620783523

SqDist[Present] = [SqDist_0] - 0.0754399651622438 * [Elevation] + 548.631913561636 * [Roughness (St Dev)] - 569.48612895412 * [Slope] - 1.87548518224877 * [Snow 2005] - 0.554150222435931 * [Solar Radiation] - 4.42195651075929 * [Wind Exposure] + 1035.425248205

Prob[0] = Exp(-0.5 * [SqDistAbsent]) + Exp(-0.5 * [SqDistPresent])

Prob[Absent] = Exp(-0.5 * [SqDistAbsent]) / [Prob_0]

Prob[Present] = Exp(-0.5 * [SqDistPresent]) / [Prob_0]

Covariance Matrices

Within Cov	Elevation	Wind Exposure	Roughness (St Dev)	Slope	Solar Radiation	Snow 2005
Elevation	92.509461	-9.358537	0.1031595	0.0493347	134.05139	3.4770533
Wind Exposure	-9.358537	741.99684	-0.358132	-0.182939	-5651.381	-16.03334
Roughness (St Dev)	0.1031595	-0.358132	0.0328354	0.0202113	11.99948	0.1370447
Slope	0.0493347	-0.182939	0.0202113	0.0125644	7.2538379	0.0832261
Solar Radiation	134.05139	-5651.381	11.99948	7.2538379	70022.307	160.75684
Snow 2005	3.4770533	-16.03334	0.1370447	0.0832261	160.75684	4.2967976

Within Corr	Elevation	Wind Exposure	Roughness (St Dev)	Slope	Solar Radiation	Snow 2005
Elevation	1	-0.03572	0.0591895	0.0457603	0.0526696	0.1743997
Wind Exposure	-0.03572	1	-0.072556	-0.059915	-0.784035	-0.283956
Roughness (St Dev)	0.0591895	-0.072556	1	0.9950669	0.2502493	0.3648534
Slope	0.0457603	-0.059915	0.9950669	1	0.2445565	0.3581929

Solar Radiation	0.0526696	-0.784035	0.2502493	0.2445565	1	0.293075
Snow 2005	0.1743997	-0.283956	0.3648534	0.3581929	0.293075	1

Colony/ Control Plot Means

SqDist[0] = 210.773103515622 * [Slope] * [Slope] - 3.20902949541339 * [Snow 2005] * [Slope] + 0.414053198067366 * [Snow 2005] * [Snow 2005] - 0.124529483606092 * [Solar Radiation] * [Slope] - 0.000752518161675032 * [Solar Radiation] * [Snow 2005] + 0.0000882694767886766 * [Solar Radiation] * [Solar Radiation] + 16.4743238786638 * [Wetness Index] * [Slope] + 0.295851270439193 * [Wetness Index] * [Snow 2005] - 0.00294815654533994 * [Wetness Index] * [Solar Radiation] + 1.40792200808666 * [Wetness Index] * [Wetness Index] - 0.963648602576033 * [Wind Exposure] * [Slope] + 0.0120618995124814 * [Wind Exposure] * [Snow 2005] + 0.00135066786036723 * [Wind Exposure] * [Solar Radiation] - 0.0251471618410039 * [Wind Exposure] * [Wetness Index] + 0.00707443125508775 * [Wind Exposure] * [Wind Exposure]

Sqdist[Absent] = [SqDist_0] + 309.315479717998 * [Slope] + 0.0796705282217051 * [Snow 2005] - 0.616022491955657 * [Solar Radiation] - 6.52142537217429 * [Wetness Index] - 4.89146890176223 * [Wind Exposure] + 1133.89312820219

SqDist[Present] = [SqDist_0] + 347.598422398211 * [Slope] - 0.483032411526596 * [Snow 2005] - 0.630506914397322 * [Solar Radiation] - 5.22651320557493 * [Wetness Index] - 5.01468305062957 * [Wind Exposure] + 1177.41573294364

Prob[0] = Exp(-0.5 * [SqDist_Absent]) + Exp(-0.5 * [SqDist_Presnt])

Prob[Absent] = Exp(-0.5 * [SqDist_Absent]) / [Prob_0]

Prob[Present] = Exp(-0.5 * [SqDist_Present]) / [Prob_0]

Covariance Matrices

Within Cov	Slope	Snow 2005	Solar Radiation	Wetness Index	Wind Exposure
Slope	0.0090767	0.0667995	8.2506729	-0.054368	-0.323004

Snow 2005	0.0667995	3.3080644	169.90928	-0.701024	-15.73625
Solar Radiation	8.2506729	169.90928	53843.172	-52.7652	-4816.629
Wetness Index	-0.054368	-0.701024	-52.7652	1.0811758	3.8533405
Wind Exposure	-0.323004	-15.73625	-4816.629	3.8533405	599.4201

Within Corr	Slope	Snow 2005	Solar Radiation	Wetness Index	Wind Exposure
Slope	1	0.3854988	0.3732165	-0.548827	-0.138478
Snow 2005	0.3854988	1	0.4025917	-0.370679	-0.353385
Solar Radiation	0.3732165	0.4025917	1	-0.218693	-0.847837
Wetness Index	-0.548827	-0.370679	-0.218693	1	0.1513643
Wind Exposure	-0.138478	-0.353385	-0.847837	0.1513643	1

Snow accumulation patterns and Adélie penguin colony population trends

Whitney Pt

Individual Cell Values

SqDist[0] = 0.00119480599248143 * [Wind Exposure] * [Wind Exposure] - 0.00495075037305612 * [Snow 2005] * [Wind Exposure] + 1.28750499965548 * [Snow 2005] * [Snow 2005]

SqDist [1] = [SqDist0] - 0.181849610614292 * [Wind Exposure] + -0.737327811452706 * [Snow 2005] + 7.16135062926116

SqDist[2] = [SqDist0] + -0.210749485702305 * [Wind Exposure] + -0.384284916302041 * [Snow 2005] + 9.42479843636275

SqDist[3] = [SqDist0] + -0.042176345676994 * [Wind Exposure] + -2.732731581583 * [Snow 2005] + 1.9226510099326

SqDist[4] = [SqDist0] + -0.0928165607794011 * [Wind Exposure] + -3.17283096438601 * [Snow 2005] + 4.01021061450554

SqDist[5] = [SqDist0] + -0.145561937355603 * [Wind Exposure] + -0.634467011940916 * [Snow 2005] + 4.60422355960992

Prob[0] = Exp(-0.5 * [SqDist_1]) + Exp(-0.5 * [SqDist_2]) + Exp(-0.5 * [SqDist_3]) + Exp(-0.5 * [SqDist_4]) + Exp(-0.5 * [SqDist_5])

Prob[1] = Exp(-0.5 * [SqDist_1]) / [Prob_0]

Prob[2] = Exp(-0.5 * [SqDist_2]) / [Prob_0]

Prob[3] = Exp(-0.5 * [SqDist_3]) / [Prob_0]

Prob[4]= Exp(-0.5 * [SqDist_4]) / [Prob_0]

Prob[5] = Exp(-0.5 * [SqDist_5]) / [Prob_0]

Covariance Matrices

Within Cov	Wind Exposure	Snow 2005
Wind Exposure	840.30309	1.6155785
Snow 2005	1.6155785	0.7798021
Within Corr	Wind Exposure	Snow 2005
Wind Exposure	1	0.0631129
Snow 2005	0.0631129	1

Shirley I

Individual Cell Values

SqDist[0] = 0.00355133437660017 * [Wind Exposure] * [Wind Exposure] - 0.0122666431229158 * [si_Snow Difference] * [Wind Exposure] + 1.44400609918618 * [si_Snow Difference] * [si_Snow Difference] + 0.0191491013415595 * [Snow 2005] * [Wind Exposure] - 0.833530039669805 * [Snow 2005] * [si_Snow Difference] + 0.448077767126803 * [Snow 2005] * [Snow 2005]

SqDist[1] = [SqDist_0] - 0.326509096720394 * [Wind Exposure] + 0.931717261940955 * [si_Snow Difference] - 1.60042448751615 * [Snow 2005] + 7.83533108016429

SqDist[2] = [SqDist_0] - 0.29111562134154 * [Wind Exposure] + 0.457877029907374 * [si_Snow Difference] - 0.999052890366228 * [Snow 2005] + 6.00768672435545

SqDist[3] = [SqDist_0] - 0.391907053591727 * [Wind Exposure] + 0.867650398892529 * [si_Snow Difference] - 1.73080959359006 * [Snow 2005] + 11.128983294363

SqDist[4] = [SqDist_0] - 0.374071002536028 * [Wind Exposure] + 1.49665824484529 * [si_Snow Difference] - 1.99495909001073 * [Snow 2005] + 10.428255516494

SqDist[5] = [SqDist_0] - 0.38557381647713 * [Wind Exposure] + 1.11727824354657 * [si_Snow Difference] - 2.00858044222978 * [Snow 2005] + 11.0733093087087

Prob[0] = Exp(-0.5 * [SqDist_1]) + Exp(-0.5 * [SqDist_2]) + Exp(-0.5 * [SqDist_3]) + Exp(-0.5 * [SqDist_4]) + Exp(-0.5 * [SqDist_5])

Covariance Matrices

Within Cov	Wind Exposure	Snow Difference	Snow Cover 2005
Wind Exposure	299.4487	-0.785819	-7.129541
Snow Difference	-0.785819	0.9487041	0.8991978
Snow Cover 2005	-7.129541	0.8991978	3.2204596
Within Corr	Wind Exposure	Snow Difference	Snow Cover 2005
Wind Exposure	1	-0.046623	-0.229584
Snow Difference	-0.046623	1	0.5144355
Snow Cover 2005	-0.229584	0.5144355	1

Colony Mean Values

SqDist[0] = 0.00418545304162922 * [Wind Exposure] * [Wind Exposure] - 0.0409177714949934 * [si_Snow Difference] * [Wind Exposure] + 1.25394763665964 * [si_Snow Difference] * [si_Snow Difference] + 0.0182112772863148 * [Snow 2005] * [Wind Exposure] - 0.556128757926968 * [Snow 2005] * [si_Snow Difference] + 0.296192161457333 * [Snow 2005] * [Snow 2005]

SqDist[1] = [SqDist_0] - 0.403009228852826 * [Wind Exposure] + 1.77255512481391 * [si_Snow Difference] - 1.56385985452615 * [Snow 2005] + 10.2864780106369

SqDist[2] = [SqDist_0] - 0.339422180734924 * [Wind Exposure] + 1.07463220940373 * [si_Snow Difference] - 1.00045262612069 * [Snow 2005] + 7.11326805658983

Sqdist[3] = [SqDist_0] - 0.429360758965208 * [Wind Exposure] + 2.83961653826747 * [si_Snow Difference] - 1.65987525601527 * [Snow 2005] + 11.4921206070945

SqDist[4] = [SqDist_0] - 0.276606287361157 * [Wind Exposure] + 1.86878763386281 * [si_Snow Difference] - 1.25850871673196 * [Snow 2005] + 4.96057557062363

SqDist[5] = [SqDist_0] - 0.381960186680773 * [Wind Exposure] + 1.93531116931298 * [si_Snow Difference] - 1.91015401314318 * [Snow 2005] + 9.95382437560013

Prob[0] = Exp(-0.5 * [Sqdist_1]) + Exp(-0.5 * [Sqdist_2]) + Exp(-0.5 * [SqDist_3]) + Exp(-0.5 * [SqDist_4]) + Exp(-0.5 * [SqDist_5])

Prob[1] = Exp(-0.5 * [Sqdist_1]) / [Prob_0]

Prob[2] = Exp(-0.5 * [Sqdist_2]) / [Prob_0]

Prob[3] = Exp(-0.5 * [Sqdist_3]) / [Prob_0]

Prob[4] = Exp(-0.5 * [Sqdist_4) / [Prob_0]

Prob[5] = Exp(-0.5 * [Sqdist_5]) / [Prob_0]

Covariance Matrices

Within Cov	Wind Exposure	Snow Difference	Snow 2005
Wind Exposure	265.77018	3.1880962	-5.17742
Snow Difference	3.1880962	1.0453926	0.8834021
Snow 2005	-5.17742	0.8834021	4.364688

Within Corr	Wind Exposure	Snow Difference	Snow 2005
Wind Exposure	1	0.1912664	-0.152014

Snow Difference	0.1912664	1	0.4135634
Snow 2005	-0.152014	0.4135634	1

Proximity to human activities and population trends of Adélie penguin colonies

Whitney Pt

Individual Cell Values

SqDist[0] = 0.000098067734866 1567 * [Casey Distance] * [Casey Distance]

SqDist[1] = [SqDist0] + -0.679268214076451 * [Casey Distance] + 1176.2413684897

SqDist[2] = [SqDist0] + -0.640709630755633 * [Casey Distance] + 1046.49309863046

SqDist[3] = [SqDist0] + -0.650437435608941 * [Casey Distance] + 1078.51185259592

SqDist[4] = [SqDist0] + -0.695103006372688 * [Casey Distance] + 1231.72058100399

SqDist[5] = [SqDist0] + -0.670647532859328 * [Casey Distance] + 1146.57515528463

Prob[0] = Exp(-0.5 * [SqDist_1]) + Exp(-0.5 * [SqDist_2]) + Exp(-0.5 * [SqDist_3]) + Exp(-0.5 * [SqDist_4]) + Exp(-0.5 * [SqDist_5])

Prob[1] = Exp(-0.5 * [SqDist_1]) / [Prob_0]

Prob[2] = Exp(-0.5 * [SqDist_2]) / [Prob_0]

Prob[3] = Exp(-0.5 * [SqDist_3]) / [Prob_0]

Prob[4] = Exp(-0.5 * [SqDist_4]) / [Prob_0]

Prob[5] = Exp(-0.5 * [SqDist_5]) / [Prob_0]

Covariance Matrices

Within Cov	Casey Distance
Casey Distance	10197.034

Within Corr	Casey Distance

Casey Distance	1

Colony Mean Values

SqDist[0] = 0.0000641817780638834 * [Casey Distance] * [Casey Distance]

SqDist[1] = [SqDist0] + -0.444012107917068 * [Casey Distance] + 767.923380763649

SqDist[2] = [SqDist0] + -0.419344483335995 * [Casey Distance] + 684.967762693098

SqDist[3] = [SqDist0] + -0.436944410516673 * [Casey Distance] + 743.670647811168

SqDist[4] = [SqDist0] + -0.454879366578845 * [Casey Distance] + 805.973456879061

SqDist[5] = [SqDist0] + -0.446269576387421 * [Casey Distance] + 775.751859861134

Prob[0] = Exp(-0.5 * [SqDist_1]) + Exp(-0.5 * [SqDist_2]) + Exp(-0.5 * [SqDist_3]) + Exp(-0.5 * [SqDist_4]) + Exp(-0.5 * [SqDist_5])

Prob[1] = Exp(-0.5 * [SqDist_1]) / [Prob_0]

Prob[2] = Exp(-0.5 * [SqDist_2]) / [Prob_0]

Prob[3] = Exp(-0.5 * [SqDist_3]) / [Prob_0]

Prob[4] = Exp(-0.5 * [SqDist_4]) / [Prob_0]

Prob[5] = Exp(-0.5 * [SqDist_5]) / [Prob_0]

Covariance Matrices

Within Cov	Casey Distance
Casey Distance	15580.746

Within Corr	Casey Distance
Casey Distance	1

Shirley I

Individual Cell Values:

SqDist[0] = 0.00144737517073122 * [Casey Distance] * [Casey Distance] + 0.00010193234367731 8 * [Wind Exposure] * [Casey Distance] + 0.00334570328953665 * [Wind Exposure] * [Wind Exposure] - 0.00328731675902068 * [Ice Distance] * [Casey Distance] - 0.00014162814513091 9 * [Ice Distance] * [Wind Exposure] + 0.00190426355162413 * [Ice Distance] * [Ice Distance]

SqDist[1] = [SqDist_0] - 1.8817955886733 * [Casey Distance] - 0.343253339763072 * [Wind Exposure] + 2.09659673289585 * [Ice Distance] + 628.803805364984

SqDist[2] = [SqDist_0] - 1.70098629110724 * [Casey Distance] - 0.320036974585171 * [Wind Exposure] + 1.90720816126373 * [Ice Distance] + 509.122367903401

SqDist[3] = [SqDist_0] - 1.98865696402544 * [Casey Distance] - 0.405846574045663 * [Wind Exposure] + 2.20834597338445 * [Ice Distance] + 708.994535458669

SqDist[4] = [SqDist_0] - 1.94552360935418 * [Casey Distance] - 0.377062838897193 * [Wind Exposure] + 2.14540130069008 * [Ice Distance] + 689.083977089765

SqDist[5] = [SqDist_0] - 1.91746719390296 * [Casey Distance] - 0.384474528871473 * [Wind Exposure] + 2.10834255344528 * [Ice Distance] + 675.462971469201

Prob[0] = Exp(-0.5 * [SqDist_1]) + Exp(-0.5 * [SqDist_2]) + Exp(-0.5 * [SqDist_3]) + Exp(-0.5 * [SqDist_4]) + Exp(-0.5 * [SqDist_5])

Prob[1] = Exp(-0.5 * [SqDist_1]) / [Prob_0]

Prob[2] = Exp(-0.5 * [SqDist_2]) / [Prob_0]

Prob[3] = Exp(-0.5 * [SqDist_3]) / [Prob_0]

Prob[4] = Exp(-0.5 * [SqDist_4]) / [Prob_0]

Prob[5] = Exp(-0.5 * [SqDist_5]) / [Prob_0]

Covariance Matrices

Within Cov	Casey Distance	Wind Exposure	Sea-Ice Distance

Casey Distance	34932.826	106.13114	30156.092
Wind Exposure	106.13114	299.4487	102.74235
Sea-Ice Distance	30156.092	102.74235	26558.083
Within Corr	Casey Distance	Wind Exposure	Sea-Ice Distance
Casey Distance	1	0.0328144	0.9900562
Wind Exposure	0.0328144	1	0.0364326
Sea-Ice Distance	0.9900562	0.0364326	1

Colony Means

SqDist[0] = 0.00377009434175355 * [Wind Exposure] * [Wind Exposure] + 0.0000213494170722331 * [Ice Distance] * [Wind Exposure] + 0.0000153055246896057 * [Ice Distance] * [Ice Distance]

SqDist[1] = [Sqdist_0] - 0.364588332443925 * [Wind Exposure] - 0.0190802926907297 * [Ice Distance] + 14.1454030938387

Sqdist[2] = [Sqdist_0] - 0.316363528583488 * [Wind Exposure] - 0.013365901038925 * [Ice Distance] + 9.18186045334272

SqDist[3] = [Sqdist_0] - 0.374055303559716 * [Wind Exposure] - 0.0159058835321895 * [Ice Distance] + 12.8856768822215

SqDist[4] = [Sqdist_0] - 0.24822328525483 * [Wind Exposure] - 0.0265988871600862 * [Ice Distance] + 15.0610941863943

SqDist[5] = [Sqdist_0] - 0.339762543620913 * [Wind Exposure] - 0.0261371093399869 * [Ice Distance] + 18.0276073022348

Prob[0] = Exp(-0.5 * [Sqdist_1]) + Exp(-0.5 * [sqDist_2]) + Exp(-0.5 * [sqdist_3]) + Exp(-0.5 * [sqdist_4]) + Exp(-0.5 * [sqdist_5])

Prob[1] = Exp(-0.5 * [SqDist_1]) / [Prob_0]

Prob[2] = Exp(-0.5 * [SqDist_2]) / [Prob_0]

Prob[3] = Exp(-0.5 * [SqDist_3]) / [Prob_0]

Prob[4] = Exp(-0.5 * [SqDist_4]) / [Prob_0]

Prob[5] = Exp(-0.5 * [SqDist_5]) / [Prob_0]

Covariance Matrices

Within Cov	Wind Exposure	Ice Distance
Wind Exposure	265.77018	-185.3592
Ice Distance	-185.3592	65465.162

Within Corr	Wind Exposure	Ice Distance
Wind Exposure	1	-0.044438
Ice Distance	-0.044438	1

Appendix 3 Decision Trees

Adélie penguin colony distributions

Whitney Pt

Individual cell values

Solar Radiation <= 2434.05: Absent (154.0/14.0)
Solar Radiation > 2434.05
| Elevation <= 13.43053: Present (449.0/97.0)
| Elevation > 13.43053
| | Aspect <= 139.285
| | | Snow 2005 <= 0.488844
| | | | Elevation <= 15.65099: Present (101.0/47.0)
| | | | Elevation > 15.65099: Absent (171.0/56.0)
| | | Snow 2005 > 0.488844: Absent (155.0/38.0)
| | Aspect > 139.285
| | | Wind Exposure <= 29: Absent (287.0/62.0)
| | | Wind Exposure > 29: Present (403.0/103.0)

Colony/control plot mean values

Roughness (St Dev) <= 0.85
| Wetness Index <= 3.46: Present (7.0)
| Wetness Index > 3.46
| | Wetness Index <= 5.06: Absent (45.0/21.0)
| | Wetness Index > 5.06: Present (6.0)
Roughness (St Dev) > 0.85: Absent (6.0)

Shirley I

Individual cell values

Slope <= 0.217657
| Elevation <= 29.38974
| | Aspect <= 108.408
| | | Snow Difference <= 0.00002
| | | | Wind Exposure <= 48: Present (201.0/80.0)
| | | | Wind Exposure > 48: Absent (285.0/111.0)
| | | Snow Difference > 0.00002: Present (1096.0/218.0)
| | Aspect > 108.408
| | | Roughness (St Dev) <= 0.135416
| | | | Elevation <= 15.0789
| | | | | Elevation <= 9.950747: Absent (333.0/83.0)
| | | | | Elevation > 9.950747: Present (200.0/52.0)
| | | | Elevation > 15.0789: Absent (200.0/38.0)
| | | Roughness (St Dev) > 0.135416: Absent (1139.0/147.0)
| Elevation > 29.38974: Present (1104.0/54.0)
Slope > 0.217657: Absent (663.0/11.0)

Colony/control plot mean values

Slope <= 0.13
| Planar Curvature <= -0.02: Present (6.0)
| Planar Curvature > -0.02
| | Elevation <= 29
| | | Planar Curvature <= -0.01
| | | | Snow 2005 <= 0.97: Absent (8.0/3.0)
| | | | Snow 2005 > 0.97: Present (5.0)
| | | Planar Curvature > -0.01
| | | | Elevation <= 12.65
| | | | | Planar Curvature <= 0
| | | | | | Slope <= 0.11
| | | | | | | Solar Radiation <= 3254.44: Absent (4.0)
| | | | | | | Solar Radiation > 3254.44
| | | | | | | | Solar Radiation <= 3392.32

| | | | | | | | | Roughness (St Dev) <= 0.04: Absent (3.0/1.0)
| | | | | | | | | Roughness (St Dev) > 0.04: Present (10.0)
| | | | | | | | Solar Radiation > 3392.32: Absent (6.0/1.0)
| | | | | | Slope > 0.11: Present (6.0)
| | | | | Planar Curvature > 0: Present (2.0)
| | | | Elevation > 12.65: Absent (7.0)
| | Elevation > 29: Present (10.0)
Slope > 0.13: Absent (31.0/4.0)

Snow accumulation patterns and Adélie penguin colony population trends

Whitney Pt

Individual cell values

Wind Exposure <= 39
| Snow 2005 <= 0.072101: Strong Increase (66.0/3.0)
| Snow 2005 > 0.072101
| | Wind Exposure <= 5: Stable (51.0/13.0)
| | Wind Exposure > 5: Strong Increase (167.0/75.0)
Wind Exposure > 39: Strong Increase (548.0/32.0)

Colony mean values

Snow 2005 <= 1.14: Strong Increase (31.0/8.0)
Snow 2005 > 1.14: Stable (2.0/1.0)

Shirley I

Individual cell values

Snow 2005 <= 0.201634
| Snow 2005 <= 0.000087
| | Wind Exposure <= 56: Moderate Decrease (139.0/58.0)
| | Wind Exposure > 56: Stable (64.0)
| Snow 2005 > 0.000087
| | Snow 2005 <= 0.003957: Stable (64.0/14.0)
| | Snow 2005 > 0.003957: Strong Decrease (89.0/39.0)

```
Snow 2005 > 0.201634
| Wind Exposure <= 36
| | Wind Exposure <= 16: Moderate Increase (54.0/19.0)
| | Wind Exposure > 16
| | | Snow Difference <= 0.755802: Strong Increase (181.0/69.0)
| | | Snow Difference > 0.755802: Stable (54.0/32.0)
| Wind Exposure > 36
| | Wind Exposure <= 67
| | | Snow 2005 <= 1.697746
| | | | Snow 2005 <= 1.139009
| | | | | Snow Difference <= 0.013476
| | | | | | Snow 2005 <= 0.693254
| | | | | | | Snow 2005 <= 0.390971: Strong Increase (54.0/29.0)
| | | | | | | Snow 2005 > 0.390971: Stable (152.0/70.0)
| | | | | | Snow 2005 > 0.693254: Moderate Increase (77.0/41.0)
| | | | | Snow Difference > 0.013476
| | | | | | Snow 2005 <= 0.571027: Stable (62.0/11.0)
| | | | | | Snow 2005 > 0.571027
| | | | | | | Wind Exposure <= 52: Strong Increase (111.0/46.0)
| | | | | | | Wind Exposure > 52
| | | | | | | | Snow 2005 <= 0.842806: Strong Increase (70.0/33.0)
| | | | | | | | Snow 2005 > 0.842806: Stable (76.0/15.0)
| | | | Snow 2005 > 1.139009
| | | | | Snow Difference <= 0.414414: Moderate Increase (71.0/33.0)
| | | | | Snow Difference > 0.414414: Stable (56.0/29.0)
| | | Snow 2005 > 1.697746: Strong Increase (88.0/53.0)
| | Wind Exposure > 67: Strong Increase (239.0/108.0)
```

Colony mean values

| Snow 2005 <= 0.02
| | Snow Difference <= 0
| | | Wind Exposure <= 36.98: Stable (2.0)
| | | Wind Exposure > 36.98
| | | | Wind Exposure <= 48.51: Moderate Decrease (4.0)
| | | | Wind Exposure > 48.51: Strong Decrease (2.0/1.0)
| | Snow Difference > 0: Stable (2.0)
| Snow 2005 > 0.02: Strong Decrease (2.0/1.0)
Snow 2005 > 0.07
| Snow 2005 <= 0.38: Strong Decrease (5.0)
| Snow 2005 > 0.38
| | Snow 2005 <= 1.22
| | | Wind Exposure <= 38.24: Strong Decrease (4.0/1.0)
| | | Wind Exposure > 38.24: Strong Increase (7.0/3.0)
| | Snow 2005 > 1.22
| | | Wind Exposure <= 53.3
| | | | Snow 2005 <= 3.06
| | | | | Snow Difference <= 0.85
| | | | | | Wind Exposure <= 47.67: Strong Decrease (3.0/1.0)
| | | | | | Wind Exposure > 47.67: Strong Increase (4.0/2.0)
| | | | | Snow Difference > 0.85: Moderate Decrease (2.0)
| | | | Snow 2005 > 3.06: Strong Increase (4.0/1.0)
| | | Wind Exposure > 53.3: Strong Decrease (3.0)

Proximity to human activities and population trends of Adélie penguin colonies

Whitney Pt

Individual cell values

Casey Distance <= 3310.06

| Casey Distance <= 3292.787: Strong Increase (75.0/8.0)

| Casey Distance > 3292.787: Stable (139.0/38.0)

Casey Distance > 3310.06

| Casey Distance <= 3536.243: Strong Increase (435.0/11.0)

| Casey Distance > 3536.243

| | Casey Distance <= 3549.416: Moderate Increase (50.0/25.0)

| | Casey Distance > 3549.416: Strong Increase (133.0/1.0)

Colony mean values

Casey Distance <= 3303.17: Stable (5.0/2.0)

Casey Distance > 3303.17: Strong Increase (28.0/6.0)

Shirley I

Individual cell values

Ice Distance <= 814.8865

| Casey Distance <= 860.6161: Moderate Decrease (103.0/22.0)

| Casey Distance > 860.6161

| | Casey Distance <= 1559.853

| | | Ice Distance <= 711.3762

| | | | Ice Distance <= 567.3976

| | | | | Ice Distance <= 176.9633: Stable (60.0/7.0)

| | | | | Ice Distance > 176.9633: Strong Decrease (61.0/9.0)

| | | | Ice Distance > 567.3976: Stable (236.0/12.0)

| | | Ice Distance > 711.3762: Strong Decrease (67.0/12.0)

| | Casey Distance > 1559.853: Stable (326.0/31.0)

Ice Distance > 814.8865

| Casey Distance <= 1682.517

| | Ice Distance <= 858.5243

| | | Wind Exposure <= 30: Moderate Increase (51.0/19.0)

| | | Wind Exposure > 30: Strong Increase (78.0/35.0)

| | Ice Distance > 858.5243

| | | Wind Exposure <= 39: Strong Increase (61.0/24.0)

| | | Wind Exposure > 39: Moderate Increase (198.0/30.0)

| Casey Distance > 1682.517: Strong Increase (460.0/19.0)

Colony mean values

Casey Distance <= 1588.42

| Casey Distance <= 855.58: Moderate Decrease (5.0/1.0)

| Casey Distance > 855.58

| | Wind Exposure <= 46.13

| | | Casey Distance <= 1470.6

| | | | Casey Distance <= 935.88: Stable (4.0/1.0)

| | | | Casey Distance > 935.88

| | | | | Casey Distance <= 1438.98: Moderate Decrease (3.0)

| | | | | Casey Distance > 1438.98: Stable (3.0/1.0)

| | | Casey Distance > 1470.6: Strong Decrease (3.0)

| | Wind Exposure > 46.13

| | | Casey Distance <= 1547.58: Strong Decrease (11.0/1.0)

| | | Casey Distance > 1547.58: Stable (3.0/1.0)

Casey Distance > 1588.42

| Casey Distance <= 1778.12: Strong Increase (10.0/2.0)

| Casey Distance > 1778.12: Strong Decrease (2.0)