Methods for mapping the tundra vegetation of sub-Antarctic Macquarie Island

by

Phillippa Kate Bricher

M.Env.Mgt, Grad Cert GIS, B.A(Journ)

School of Geography and Environmental Studies

Submitted in fulfilment of the requirements for the Doctor of Philosophy

University of Tasmania

July, 2012

Statement of Co-Authorship

The following people and institutions contributed to the publication of the work undertaken as part of this thesis:

Paper 1 (Ch. 2)   Phillippa Bricher (90%), Arko Lucieer (5%), Dana Bergstrom (5%)

Paper 2 (Ch. 3)   Phillippa Bricher (85%), Arko Lucieer (10%), Dana Bergstrom (5%)

Paper 3 (Ch. 4)   Phillippa Bricher (90%), Arko Lucieer (5%), Dana Bergstrom (5%)

Paper 4 (Ch. 5)   Phillippa Bricher (85%), Arko Lucieer (10%), Dana Bergstrom (5%)

Details of the Authors' roles:

Paper 1: All authors contributed to the idea and its development; Bricher and Bergstrom conducted the fieldwork; Bricher performed all data processing, analysis, and writing; Lucieer and Bergstrom contributed editorial advice.

Paper 2: All authors contributed to the idea and its development; Bricher performed most of the data processing, all analysis and writing; Lucieer performed the image classification provided a written description of this analysis; and Lucieer and Bergstrom provided editorial advice.
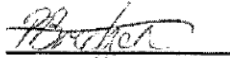
Paper 3: All authors contributed to the idea and its development; Bricher performed all data analysis and writing; Bergstrom and Lucieer provided editorial advice.

Paper 4: Bricher and Lucieer contributed to the idea and its development; Bricher performed 90% of the data processing and analysis and all the writing; Lucieer applied the random forest models to the satellite image; Lucieer and Bergstrom provided editorial advice.
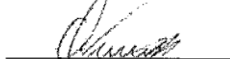

We the undersigned agree with the above stated "proportion of work undertaken" for each of the above published (or submitted) peer-reviewed manuscripts contributing to this thesis:
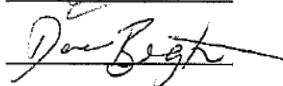

Signed:

Candidate: Phillippa Bricher

Author 1: Arko Lucieer

Author 2: Dana Bergstrom

Date: 16/5/12

Abstract

Vegetation mapping is increasingly used for detecting changes in plant distributions at landscape scales. This is aided by the development of sophisticated machine learning tools and very high resolution satellite imagery, which together allow for the detection of changes at finer spatial and attribute resolutions. Even with these new tools, the sub-Antarctic poses particular challenges for mapping. The remoteness demands particular efficiency in field sampling and regular cloud-cover hampers satellite image acquisition and the tundra vegetation is small-statured and dominated by ecologically generalist species. Additionally, High accuracy mapping is of particular interest on sub-Antarctic Macquarie Island, where the vegetation is changing rapidly. Reliable maps of current species distributions would provide valuable baselines for monitoring future change.

This thesis develops and tests methods for all steps of the mapping process, including field methods, sampling design, examining whether plant communities are an appropriate mapping unit for change detection, and testing the best methods of image classification. A new photographic field sampling method is demonstrated to collect equivalent data to an existing field quadrat method but with less field time needed. Next, a geographically-stratified random sampling design is developed, applied and tested for its capacity to capture the variation in terrain variables that are thought to drive plant species distributions.

Previous descriptions of the vegetation of Macquarie Island are based on the idea that the vegetation can be more or less grouped into discrete communities. I test this hypothesis and find that floristic classification is unable to identify any stable communities on this highly disturbed island.

I therefore finally focus on single-species mapping, determining the best combination of very high resolution satellite imagery, terrain modelling and random forest classification to map the extent of the endemic, critically-endangered cushion plant, *Azorella macquariensis*, with very high accuracy. The successful mapping of *A. macquariensis*, a small-statured and patchily distributed species,

demonstrates the utility of this mapping technique. I describe initial results for applying this technique to other species, highlighting the ecological and spectral characteristics that distinguish those tundra species that are amenable to mapping from VHR satellite imagery and terrain modelling.

In summation, in this thesis, I develop and test tools for all stages of mapping the distribution of plant species on sub-Antarctic Macquarie Island at a resolution suitable for monitoring rapid distributional changes.

Acknowledgements

Contents

List of Figures

14

16

List of Tables

1. Introduction

Mapping of vegetation is a common scientific or conservation action to record ecosystems within both spatial and temporal frameworks. In particular, serial or sequential mapping is one of the first tools used to detect environmental change. In any environment that is experiencing rapid change, it is crucial however that vegetation mapping and spatial monitoring tools are able to accurately describe that change. The ability to perceive environmental problems and to assess effectiveness of management responses or actions to them is contingent on the sensitivity and accuracy of the monitoring tools (Harris 2003) . As landscape-scale management intervention is generally costly and resources are limited, there are pragmatic as well as scientific reasons for wanting to accurately monitor environmental changes (Field et al. 2007). To be effective, monitoring programs must choose suitable variables to monitor and specify the degree of change that would trigger a management response (Field et al. 2007). In recent years, it has become apparent that the data from many ecological monitoring datasets are inadequate to detect changes induced through management programs (Harris and Heathwaite 2012).

This thesis sets out to examine methods employed at each stage of the mapping process to determine if a suite of tools can be developed to detect change in a rapidly changing ecosystem, sub- Antarctic Macquarie Island.

1.1. The sub-Antarctic islands

High latitude ecosystems are of considerable scientific interest, largely due to a combination of their relatively simple trophic structures, the rapid rate of observed climatic changes, and the opportunities they provide to understand the interactions between native and alien species (Bergstrom and Chown 1999). The low species richness and relatively simple trophic structures of terrestrial ecosystems in this region simplify analysis of ecological processes (Smith and Steenkamp 1990). Sub-Antarctic islands are also of significant conservation interest due to their rarity, small

spatial extent, importance as habitat for endangered species, high levels of endemism, and the concern that their paucity of taxa leaves them vulnerable to invasions of alien species (Chown et al. 2001; Frenot et al. 2005; Convey 2007). Primary production on many of the islands is high, and native macroherbivores are generally absent (Smith and Steenkamp 1990). The vegetation of sub-Antarctic islands is also significantly affected by the region's geographical isolation and harsh climate. Herbs, grasses, bryophytes and lichens dominate the flora, while woody plants are generally absent (Selkirk et al. 1990). This is in stark contrast to the flora at similar latitudes in the northern hemisphere, where shrubs and trees dominate the vegetation (Selkirk 2007). The islands provide valuable breeding habitat for pelagic birds and marine mammals, which in turn affect the vegetation through burrowing, trampling, grazing and nitrification (Selkirk et al. 1990).

One example of the impact of introduced macroherbivores comes from the tall coastal vegetation common to many sub-Antarctic islands, where it is thought that the absence of native vertebrate herbivores from sub-Antarctic islands encouraged the evolution of megaherbs and large tussock grass species, such as *Stillbocarpa polaris* and *Poa foliosa* (Mitchell et al. 1999). The introduction of herbivorous vertebrates in the past two hundred years has resulted in significant negative impacts on the tussock grasses and megaherbs (Frenot et al. 2005; Convey 2007).

Researchers have warned that specific management is needed to mitigate the effects of invasive species, especially in light of additional ecological effects of climate change and increased visits to the region by tourists and scientists (Kriwoken and Holmes 2007). Better understanding of the spatial patterns of the distributions of plant communities and of the changes in these distributions is likely to lead to better understanding of the relative ecological effects of management interventions, invasive species and climate.

1.1.1. **Monitoring change**

On sub-Antarctic islands, vegetation patterns and dynamics have traditionally been quantified through monitoring of quadrats or through mapping homogeneous patches of vegetation.

Generally, monitoring plots have had very small spatial extents (e.g. Chown & Smith 1993; Bergstrom et al. 2002; Scott & Kirkpatrick 2008). While such studies provided highly detailed information about the changes occurring within a small area, they are of limited value in quantifying changes at the landscape scale, and have tended to not explicitly account for the spatial distribution of plants, plant communities and their environments.

1.1.2. **Macquarie Island**

Macquarie Island is one of Australia's two sub-Antarctic territories and is located approximately halfway between Tasmania and Antarctica (158°55'E, 54°30'S). Covering an area of 12,390 ha, it is protected as a Nature Reserve and World Heritage Area. The Macquarie Island flora includes up to 46 vascular plant species, including five alien plant species (ABRS 1993) of which three are still reported (*Poa annua, Cerastium fontanum* and *Stellaria media*). Traditionally, five major plant communities have been recognised on the island – wet tussock grassland, herbfield, fen, bog and feldmark (Taylor 1955a). More recently, the grasslands have been divided into tall and short grasslands, fernbrake communities have been identified and the bog and fen communities have been combined into a single mire community (Selkirk et al. 1990). Recently, a new category of "grazed vegetation" has been recognised by Bergstrom et al. (2009a).

Historically, efforts to monitor the effects of herbivory on Macquarie Island have been conducted at very fine spatial scales. These include the monitoring of species composition in 1 m² quadrats within 25 m^2 plots (Bergstrom et al. 2009a); of changes in dominant species at fixed photo points (Shaw et al. 2005) and along transects, which could be analysed at multiple scales (Kirkpatrick & Scott 2002; Scott & Kirkpatrick 2008). These studies have produced detailed evidence of regime shifts within sites that have been subjected to landslips or intensive grazing. Recently, very high resolution QuickBird satellite images have been used to investigate changes in the Normalised Difference Vegetation Index (NDVI) and these have indicated the replacement of vegetation by bare ground at many sites (Bergstrom et al. 2009a). Data from the satellite images and fine-scale plots

23

have demonstrated a trophic cascade on Macquarie Island. To date, however, it has been difficult to monitor successional changes in vegetation communities across the entire island.

Vegetation patterns on Macquarie Island are changing rapidly, making accurate monitoring at all spatial scales crucial. Firstly, the coastal vegetation of Macquarie Island is undergoing rapid change due to a rabbit and rodent eradication program being funded by the Tasmanian and Australian governments. One of the goals of the eradication project is the recovery of vegetation, and the eradication plan calls for ongoing monitoring of changes in the distributions of individual vascular plant species (Parks and Wildlife Service 2007). Indeed, early signs of recovery have already been observed (Shaw et al. 2011). To assess whether this goal has been met, baseline maps must be produced of current vegetation patterns. Secondly, the discovery of widespread die-off in the endemic cushion plant, *Azorella macquariensis*, during fieldwork for this thesis provided added incentive for producing baseline maps of vegetation patterns.

Macquarie Island is also a useful environment for testing the applicability of field and remote sensing methods to sub-Antarctic tundra environments. It is relatively accessible, compared to Australia's other sub-Antarctic territory. This simplifies the testing of sampling regimes and field techniques compared to many other sub-Antarctic islands (e.g. Heard Island). High resolution terrain data and satellite imagery is available for the island. The small stature of sub-Antarctic vegetation has limited the usefulness of satellite imagery for mapping anything but very broad vegetation classes (e.g. Selkirk et al. 2000), but the advent of very high resolution satellite sensors (pixel size of 5 m or less) provides the potential for finer categories to be identified. The highly disturbed nature of the vegetation and the dominance of a few ecological generalist species pose particular challenges for robustly identifying communities.

1.2.   Theoretical Framework

The technical and ecological challenges of producing repeatable, accurate vegetation maps have led to the development of a wide variety of both simple and sophisticated tools by researchers in the disparate fields of ecology, remote sensing, and machine learning. There are three major approaches to mapping vegetation, each of which has limitations: field mapping, interpretation of remotely sensed imagery, and species distribution modelling (SDM). Field-based mapping requires extensive field time and expert knowledge to identify vegetation types and to locate boundaries between them (Hearn et al. 2011). Secondly, mapping from remotely sensed data has typically either involved digitising of large relatively homogeneous stands of vegetation or classification of pixels according to their spectral signatures (Nagendra 2001). Classification of satellite images has historically relied on sensors with medium or coarse spatial resolutions, with pixel resolutions ranging from 30 m to 100 m (e.g. Aspinall & Veitch 1993; Debinski et al.1999; Townsend & Walsh 2001). Such coarse resolutions have limited value for detecting fine-scale changes in vegetation patterns, such as those caused by the spread of an invasive species or grazing damage. Now however, the ready availability of satellite sensors with very high spatial resolutions, such as DigitalGlobe's QuickBird and WorldView-2 satellites, is providing scientists with the capacity to map spatial patterns in vegetation at resolutions less than 3 m. The third approach, SDM, uses spatial data of environmental variables, such as terrain and geology, to predict the distribution of species or communities, on the basis of known distributions (Guisan and Zimmermann 2000).

Each of these approaches has strengths and limitations. Field mapping is time-intensive, and hence impractical for large areas. Additionally, this approach involves an assumption that hard classes with relatively clear boundaries can be reliably identified, and requires surveyors to make subjective decisions about the location of boundaries between vegetation types. This can result in low repeatability (Hearn et al. 2011), which in turn limits the utility of such maps for change detection. Access to remote islands is also often limited, further reducing the practicability of this approach in

the sub-Antarctic, though it has been successfully implemented on some very small islands (Chapuis et al. 2004). Satellite image interpretation has been successfully used to classify vegetation based on spectral information (Lu and Weng 2007) . However, in alpine regions, it has been found that many plant communities exhibit similar spectral responses (Dirnböck et al. 2003). This can considerably reduce the accuracy of a classification based purely on satellite imagery. Terrain analysis has been widely used to predict the distributions of individual species of both plant and animal, and it is known that topography has a significant impact on the distribution of plant communities. For example, on Macquarie Island, feldmark only occurs at high altitudes and tall tussock grasslands tend to occur on steep coastal slopes (Selkirk et al. 1990). However, purely topographic models of the distribution of plant and animal species tend to have significant residual errors, suggesting that terrain is not the only determinant of distribution (Dirnböck et al. 2003). Hybrid approaches that combine the benefits of species distribution modelling and image classification methods are beginning to be used for vegetation mapping (e.g. Dirnböck et al. 2003; Dobrowski et al. 2008).

The availability of very high resolution satellite imagery, and the rapid development over the past decade of sophisticated machine learning tools for classifying such imagery, has resulted in a plethora of mapping approaches, many combinations of which have not been widely tested. In particular, there has been comparatively little testing of these approaches in high latitude tundra environments (Stow et al. 2004; Murray et al. 2010). Given that tundra vegetation tends to be small in stature and not structurally diverse, such environments pose particular problems for mapping based on image classification. The rapid changes observed in Macquarie Island vegetation require accurate and repeatable maps if we are to better understand the patterns in those changes. There is thus a need to test whether these newly emerging tools can be used to produce accurate, very high resolution maps of vegetation patterns on Macquarie Island.

Very high resolution mapping for change detection requires the definition of mapping units that can be robustly and repeatably defined. To date, landscape-scale descriptions and maps of vegetation on

Macquarie Island have focussed plant communities rather than individual species (Taylor 1955; Selkirk et al. 1990, 2000). Recent research, however, has raised renewed questions about the repeatability of community mapping (e.g. Cushman et al. 2010; Hearn et al. 2011) and multiple attempts to define communities on Macquarie Island have produced little consensus on those definitions (e.g. Taylor 1955; Selkirk et al. 1990; Bergstrom et al. 2009). Given this lack of consensus and the aim of this study to produce maps for change detection, it is uncertain whether the mapping units should describe communities or individual species. Several numerical classification tools have been developed to identify groups in floristic data and there is little guidance on which are most appropriate for any given application. There is thus a need to investigate both whether stable groups can be found in floristic data and, if so, which combination of tools would best generate such groups.

Training the definitions of communities and the image classifications in turn requires a large, representative sample of vegetation abundance across the island. There are two key aspects to this: sampling design and field sampling methods. Sampling design for environmentally-explicit mapping, such as species distribution modelling, further requires that the sample is representative of the true distribution of both vegetation and environmental variables. Environmentally-explicit sampling designs have been developed (e.g. Goedickemeier et al. 1997; Franklin et al. 2001) but new terrain morphometry and classification methods (Burrough et al. 2000; 2001) may improve their efficiency.

Collecting field data on sub-Antarctic islands places particular demands on the efficiency of field sampling methods. The limited field access means field work must be rapid, and photo-sampling, long used in marine ecology, has potential to replace quadrat-based assessments of vegetation cover, but must be tested for reliability.

1.3.   Research Aims:

The aim of this thesis is to develop techniques for field sampling and spatial analysis that will enable plant communities to be mapped in great detail and high accuracy for the whole of Macquarie Island. These techniques must be practicable in the harsh climate and remote locations of the sub-Antarctic.

The primary research question of this thesis is:

> *How can field data, satellite imagery, and terrain analysis be used to improve the mapping of sub-Antarctic plant communities?*

The research objectives of this thesis are:

1. to establish a reliable and efficient field sampling method to capture species abundance data for vascular plants at the required scale of image analysis;
2. to establish a geographically stratified random sampling method to capture the full variation of both the terrain and vegetation patterns;
3. to establish a robust statistical method for identifying plant communities that can then be translated into mapping units;
4. to produce accurate high resolution maps of the plant communities, and to use these maps to describe the spatial ecology of the vegetation of Macquarie Island. These maps should be sufficiently accurate to be used as baselines for future change detection.

1.4.   Thesis Structure

Chapter 2: Close-range photo-sampling as a rapid and effective replacement for field-based vegetation sampling in sub-Antarctic tundra landscapes for up-scaling to satellite imagery

Training of satellite image classifications generally requires large sample sizes. Sub-Antarctic islands are difficult to access and it is therefore important to develop field sampling methods that are

efficient *and* provide robust, repeatable data. This chapter describes PoleCam, a photo-sampling method for gathering species cover data in tundra vegetation. Photographic sampling has potential as a replacement for field quadrat sampling methods in areas that are difficult to access, because it is quicker than manual sampling. This chapter demonstrates that point-intercept interpretation of photographs collected with PoleCam enables the rapid collection of vegetation cover classes that is comparable to traditional field methods in the tundra vegetation of Macquarie Island. The point intercept analysis does tend to overlook very rare species, and so should be coupled with the collection of a species list in the field to ensure a comprehensive dataset for each site. The main advantages of PoleCam are its field efficiency and digital data record.

Chapter 3: Geographically-stratified random sampling design for vegetation survey: its implementation on sub-Antarctic Macquarie Island

This chapter describes GeoStrat, a geographically-stratified proportional random sampling design that aims to capture the variation in both terrain parameters and vegetation patterns. The PoleCam method of field sampling is carried out at the sites chosen through GeoStrat, as well as 54 purposively-sampled sites. The effectiveness of GeoStrat in sampling the terrain is assessed by comparing it to other simulated randomised sampling designs. One of the challenges of applying randomised sampling in remote environments is that access may be limited by safety or environmental concerns, both of which are applicable on Macquarie Island. Since access limitations to GeoStrat sites are not uniform across the strata, all the simulated randomised samples are down-weighted by the same proportions as the GeoStrat strata for the purposes of comparison. This comparison shows that GeoStrat functions as a compromise between true random sampling and purposive sampling, which is biased towards areas that are perceived to contain interesting or complex vegetation patterns.

Chapter 4: Are plant communities discrete mappable entities or fluid groupings? A test case using sub-Antarctic tundra vegetation

The traditional approach to understanding vegetation patterns on Macquarie Island has been to divide the vegetation into plant communities, and examine the distributions of those communities. An examination of previous attempts to define communities shows little agreement between authors on where boundaries should be drawn. This chapter explores whether stable and reliably identifiable groups exist in the floristic dataset collected using the methods described in the previous two chapters. Given the lack of agreement on the number and identity of communities for the island, unsupervised classification provides the opportunity for a data-driven test of whether stable groupings can be found. The effects of sampling design, included taxa, data transformations, number of clusters, and clustering algorithms on the stability of groupings is tested. No combination of these variables produces stable groupings in the data, and the least unstable groups of sites are those dominated by a single species. This indicates that there is little clustering structure in the floristic data, making it difficult to identify clear communities that could be used as mapping units.

Chapter 5: Mapping sub-Antarctic cushion plants: using the random forests to classify very high resolution satellite imagery and terrain modelling

Given the failure to define stable plant communities that can be used as mapping units, this chapter explores methods to map a single species, the endemic cushion plant *Azorella macquariensis*. Random forest classification is used to explore the utility of three image training methods and the relative importance of terrain variables and spectral reflectance for predicting the presence or absence of *Azorella*.

Object-based image analysis has gained popularity in recent years because it incorporates contextual information in classifications (Blaschke 2010), but there has been limited investigation of whether object- or pixel-based classification is most accurate for small-statured vegetation (Addink et al.

2007). This chapter tests the impact of object-based and two pixel-based training methods on the accuracy of the image classification.

There are two key data-driven approaches to mapping vegetation. Terrain data is used for species distribution modelling, to determine potential habitat rather than actual distribution (Guisan and Zimmermann 2000). Satellite imagery, in contrast, is used to identify structure and disturbance patterns but has limited capacity to distinguish spectrally similar species. Recent hybrid approaches have exploited the strengths of both approaches (e.g. Dirnböck et al. 2003; Dobrowski et al. 2008) and this chapter examines the relative effects of spectral and terrain data in improving the accuracy of the image classification. Finally, this chapter examines whether the classification can differentiate sparse and moderate cover of *Azorella*.

Chapter 6: Preliminary results of application of random forest classification for other dominant plant species on Macquarie Island.

Here, preliminary results for applying the techniques developed in chapter 5 to other species on Macquarie Island are presented. Classifying an image into multiple classes is computationally more difficult than the binary classification presented in the previous chapter. Mapping for individual species also introduces complexity into standard image classifications because species ranges may overlap, but random forest classification calculates the probability of membership for each class that may be used to produce overlapping classes that optimise the accuracy of the predicted distribution of each species. This chapter explores methods to optimise a multi-class classification of species with intergraded distributions and ends by suggesting directions for investigations to refine species distribution maps.

Chapter 7: Conclusions

This thesis describes and tests the effectiveness of a range of approaches to mapping vegetation at very high resolution and in remote tundra environments. It demonstrates that the plant community

paradigm is inappropriate for change detection mapping on Macquarie Island, but that a

combination of comparatively new techniques can be used to map individual species with high

accuracy.

2. Close-range photo-sampling as a rapid and effective replacement for field-based vegetation sampling in sub-Antarctic tundra landscapes for up-scaling to satellite imagery

2.1. Abstract

There is a growing need to monitor changes in the vegetation of many tundra areas, driven by factors including climate change and herbivory by non-indigenous vertebrate species. Evidence-based management policies are dependent on accurate quantitative data on community structure at multiple spatial scales. At landscape scales, there is a need for large samples, and in remote areas, access is often limited. These two factors drive demand for rapid field sampling techniques. We use a case study on sub-Antarctic Macquarie Island to test whether close-range photo-sampling can acquire species composition data that is equivalent to that collected by field quadrat sampling. We compare the efficiency of two photographic cover estimation techniques (photographic visual estimation (PVE) and photographic point intercept (PPI)) with field-based visual estimation (FVE).

While FVE had the greatest capacity to detect small and rare species, photo-sampling took less field-time (sometimes less than half the field time) and plots could still be reliably differentiated using multivariate clustering. Furthermore, photo-sampling allowed for greater surface areas to be sampled. Close–range photo-sampling coupled with an *in situ* species list with either of the cover estimation methods tested here proved to be robust tools for rapidly collecting vegetation data, comparable to FVE in remote tundra environments. These methods aim to provide substantive and efficiently acquired training and validation data sets for satellite image analyses.

2.2.   Introduction

Monitoring vegetation change in high latitude tundra requires robust and efficient field sampling methods. Recent rapid changes in these environments, due to factors including climate change and grazing by non-native herbivores (Serreze et al. 2000; Smith 2002; Cairns and Moen 2004; Chapuis et al. 2004; Walther et al. 2005), increases the importance of accurate and efficient monitoring methods. Increasingly, evidence-based conservation policies in these regions require high quality data at multiple spatial scales (Chown et al. 2009) that can be used to monitor significant changes (Archaux et al.2007). New remote sensing and geographic information systems (GIS) tools allow researchers to scale-up studies of vegetation dynamics from plots to landscapes, but are dependent on large ground referencing samples to train and validate the image analysis (Wilson et al. 2011), which in turn require robust and rapid field sampling techniques.

There are two key aspects to creating a ground-referencing system to maximise biological information extraction from satellite imagery. Firstly, the validation sampling plots must be representative of the vegetation and terrain and secondly, floristic survey methods must adequately describe the vegetation at those plots. This paper focuses on the latter of these aspects as part of a broader research program aimed at optimising satellite image interpretation for mapping vegetation distribution.

Scientists and land managers need field sampling methods that balance repeatability, exhaustiveness and efficiency (Archaux et al. 2007), especially when working in remote environments. The focus of this study is remote sub-Antarctic Macquarie Island, 1200 km SE of Australia, which has recently undergone significant change due to grazing pressure from a population explosion of non-indigenous rabbits (Scott & Kirkpatrick 2008; Bergstrom 2009a). Traditional plant ecology methods used in previous vegetation studies on the island (e.g. Copson and Whinam 1998; Kirkpatrick and Scott 2002; Bergstrom et al. 2009a) were considered too time consuming to be suitable for landscape-scale analyses requiring large data sets. Close-range photo-

sampling potentially offers a number of advantages over other field methods including speed, simplicity and the capacity to archive the photographs for future research (Laliberte et al. 2007).

Perhaps because of the specific demands on sampling efficiency in marine environments, the use of photo-sampling to collect quantitative data on community structure is more common there than in terrestrial environments, but there are some conflicting conclusions in the marine ecology literature about which techniques are most accurate, repeatable and time-efficient (see Foster et al. 1991; Tkachenko 2005; Drummond and Connell 2005; Alquezar and Boyd 2007; Leujak and Ormond 2007) In general, photographic methods have been considered most useful in communities with simple vertical structures and where field access is difficult.

In terrestrial ecology, close-range photo-sampling has been used to estimate biomass (Paruelo et al. 2000); quantify total canopy coverage (Northup et al. 1999; Laliberte et al. 2007); identify broad taxonomic classes (Zhou and Robson 1998; Luscier et al. 2006); and for phenological research, in which repeated sampling is required (Sparks et al. 2006; Crimmins & Crimmins 2008; Liang et al. 2011). A few studies have used close-range photo-sampling to quantify the coverage of individual species in low-statured vegetation. These studies have thus far focussed on environments with extremely sparse vegetation (Gilbert and Butt 2009) or focussed on a few, easily identifiable species (Bennett et al. 2000; Vanha-Majamaa et al. 2000; Booth et al. 2008). We are unaware, however, of any studies testing the use of close-range photo-sampling to quantify the coverage of all species in dense communities in tundra or other cool temperate environments characterised by low-statured plants. There is therefore a need to test whether close-range photo-sampling can acquire species composition and cover data that is sufficiently accurate and comparable to field quadrat sampling data.

This study presents a technique for rapid sampling of vegetation plots using photo-sampling and tests whether two photographic cover estimation methods (photographic point intercept and photographic visual estimation) can capture sufficient data on species composition and cover to

identify field plots established using field-based visual estimations of vegetation cover in an ongoing

vegetation study (see Bergstrom et al. 2009a). There is a body of literature exploring the precision

and repeatability of individual field data collection techniques (Sykes et al. 1983; Nilsson and Nilsson

1985; Klimeš et al. 2001; Gray and Azuma 2005; Archaux et al. 2006; Vittoz and Guisan 2007; Vittoz

et al. 2010) and comparing different field-based methods for assessing vegetation cover (e.g.

Archaux et al. 2007; Bråkenhielm and Qinghong 1995). We use similar analyses to assess the

capacity of rapid photo-sampling methods to quantify the coverage of plant species in a plot and to

distinguish plots from each other.

2.3. Methods

2.3.1. Study Site

Macquarie Island (54°30'S, 158°57'E) is a World Heritage-listed oceanic island located approximately

halfway between Australia and Antarctica (Fig. 1). The 12,390 ha island is low-lying and dominated

by a plateau bounded by steep coastal slopes. Its vegetation is dominated by grasses, herbs and

bryophytes (Selkirk et al. 1990). In the coastal regions where this study was conducted, the

vegetation is dominated by the grasses *Agrostis magellanica*, *Poa foliosa* and *P. annua*; the sedge

Luzula crinita; the mat-forming herb *Acaena magellanica;* and the megaherbs *Stilbocarpa polaris* and

Pleurophyllum hookeri. The vegetation has a relatively simple vertical structure with rarely more

than two strata in any vegetation type (e.g. a vascular plant canopy layer and underlying bryophyte

layer). With the exception of three species (*P. foliosa*, *S. polaris* and *Polystichum vestitum*) the

vegetation is generally less than 50 cm in height. The coastal vegetation has been extensively

modified by rabbit grazing since rabbits were introduced in the 19th century, and the grazing severity

has increased in the past decade (Bergstrom 2009a).

The island's vegetation has been studied since its discovery in 1810 (Copson 1984), with most

research examining small plots with high attribute resolution, relatively small sample sizes and

purposive sampling strategies (e.g. Copson and Whinam 1998; Kirkpatrick and Scott 2002; Scott and Kirkpatrick 2008; Bergstrom et al. 2009). Recently, satellite image interpretation has been used to investigate landscape-scale processes, but currently with very low attribute resolution (Lucieer 2008; Bergstrom et al. 2009). More detailed analyses have not been possible at the landscape scale because of the relatively small sizes of the field samples. Such study designs reflect the challenges of working in a remote, physically challenging environment, in which field time and access can be limited.
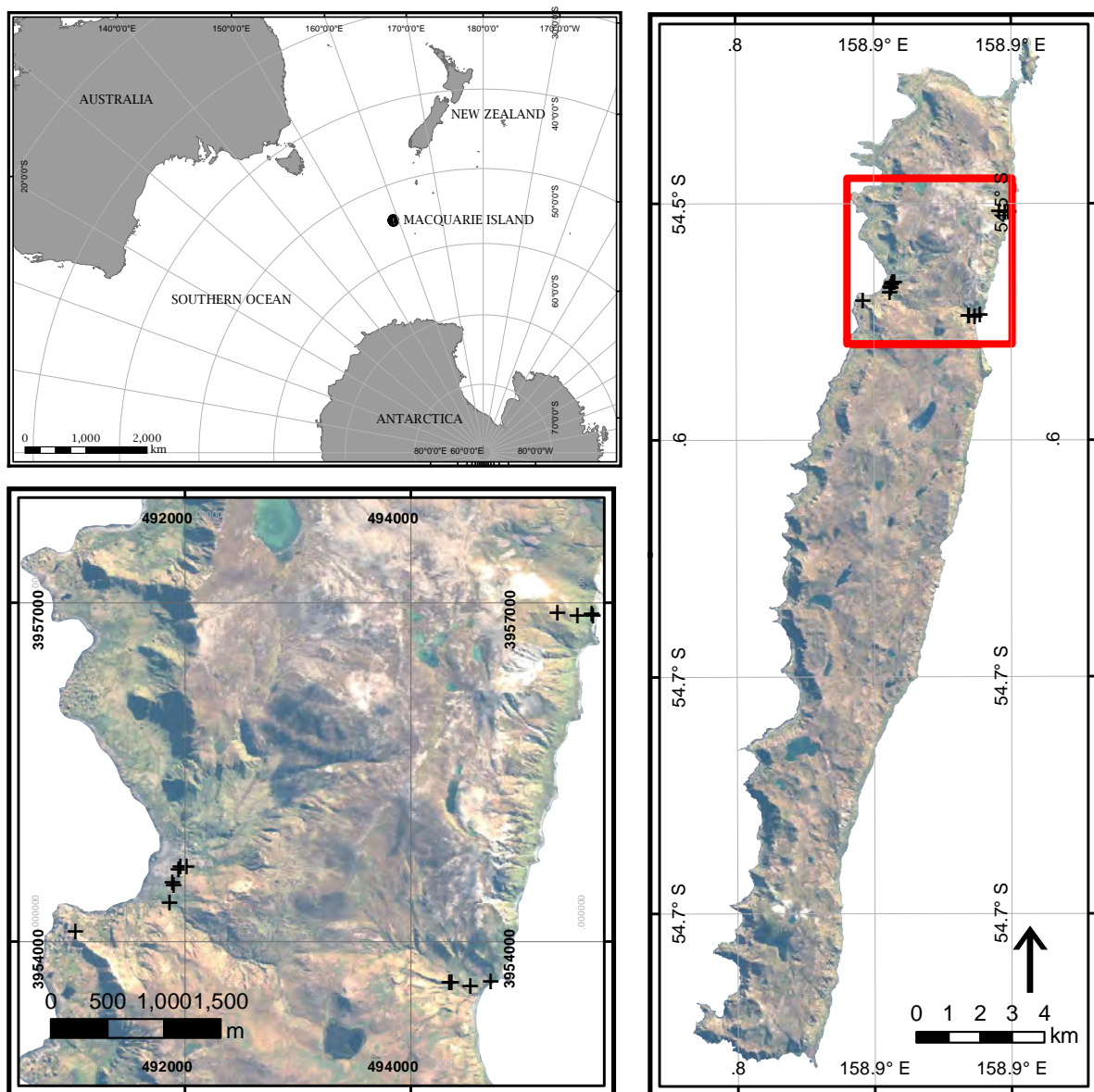
Figure 1: Macquarie Island is located in the Southern Ocean, equidistant between Australia, New Zealand and Antarctica. (Data source: Center for Applied Geoinformatics TerraSpace, www.terraspace.ru)

2.3.2. **Site Descriptions**

In March 2008, 15 sites were visited over a 7 day period to assess plant species composition and

cover (Fig. 1). Study plots had been established at 11 of these sites either in 2001 (10 sites) or 2007

(one site) (Bergstrom *et al.* 2009). The previously established sites were chosen in advance of the

field season as being feasible to visit in a short trip, as well as having the potential to have exhibited

significant changes in vegetation cover from rabbit grazing since the previous visit. These sites

encompass a range of low-altitude vegetation types, ranging from almost bare, severely grazed sites

through relatively species-rich short grasslands to *S. polaris* dominated herbfields.

2.3.3. **Plot photography**

At each site, a 5 x 5 m plot was defined with a tape measure, with intersecting measuring poles

dividing the plot into quarters. The location of the south-east corner of each plot was recorded with

a Trimble ProXH single-frequency differential GPS. A Canon EOS 400D SLR camera with an 18 mm

lens mounted on a 2 m carbon fibre pole was used to take a single photograph of each quarter of the

plot from approximately 2.5 m above the ground (i.e. four photographs per plot). Each photograph

covered an area of approximately 6.5 m^2. Guy lines were used to stabilise the camera. Five randomly

located 1 m^2 quadrats were placed within the 5 m x 5 m area.

2.3.4. **Data Collection and Vegetation Cover Estimates**

2.3.4.1.   **In situ fractional cover**
For each plot, field-based visual estimates (FVE) of percentage cover were made for all ground cover

classes within the five quadrats, following the methods of Bergstrom et al*.* (2009a). The cover

estimates were made by two botanists in consultation, using decimetre markings on the edges of

the quadrat to assist the cover estimation. For each cover class, the canopy coverage was estimated

without regard for any overlap with other cover classes, so total percentages for a quadrat exceeded

100% in plots with overlapping strata. For each plot, mean class coverage for all quadrats was

calculated and normalised to 100% for comparison with the results of the photographic

interpretation. A species list was recorded for each plot based on the species identified in the five quadrats, and supplemented with any species observed in the plot and outside the quadrats.

The cover classes included vascular plant species identified to species level, with other taxa identified to broader taxonomic levels (e.g. algae or fungi). Bryophytes were separated into *Marchantia berteroana* and other bryophytes, as *M. berteroana* is the only bryophyte species that we were able to reliably identify in the photographs and it is structurally very different to the other, mostly leafy, bryophytes on the island. Vascular plant species that could not be reliably identified to species level in photographs (*Deschampsia spp.*, *Acaena spp*. and two species of *Colobanthus*, *C. apetalus* and *C. affinis*) were identified to genus level. The final two ground cover classes were water and bare ground/dead vegetation.

2.3.4.2.   Fractional cover and point-intercept counts on digital photographs

Two photographic cover estimation methods were assessed. The canopy coverage of each ground cover class was first visually estimated (Photographic Visual Estimates - PVE) for the entire plot. To aid this, each photograph was georeferenced to a 10 cm grid using an affine transformation in ArcGIS 9.3. This approach has been found to be time consuming but accurate (Zhou et al. 1998; Roush et al. 2007; Michel et al. 2010). Next, a point intercept method (Photographic Point Intercept - PPI) was used to estimate percentage area of each cover class. One hundred randomly located points were positioned over each of the four photographs, and the cover class under each point recorded using Coral Point Count software (Kohler and Gill 2006). The PVE and PPI assessments were conducted by the same researcher on different days to minimise the potential for biasing the results. The species list acquired in the field was used to aid identification. For each photographic cover estimation method, the results for the four photographs in a plot were converted into a mean percentage for that plot.

In the photographs, only those cover classes forming the upper canopy could be observed. Vegetation that could not be confidently identified due to shadowing (a small proportion of total

cover: mean = 0.83%, SD=0.98%) was excluded from the analysis. This mean was inflated by a single

plot that was photographed in misty, dim conditions (unidentified vegetation at this plot covered

3.5%). Canopy cover for each ground cover class was calculated as a percentage of the canopy in the

plot that was identifiable and not obscured by field equipment (mean=1.55%, SD = 0.71%), with

covers summing to 100%. Thus, FVE measured the percentage of the total vegetation cover, whereas

the two photographic methods represented percentages of the visible canopy cover. Before analysis,

the photographs were cropped to exclude areas outside the plot boundaries and overlaps among

photographs.

2.3.5. **Statistical analysis**

Statistical tools for analysing dissimilarity among ecological data are generally limited in their

capacity to test hypotheses or to capture multiple aspects of that dissimilarity (Økland 2007). We

therefore used multiple statistical analyses to explore both the differences in total numbers of cover

classes detected and their abundance by each observation method.

For each plot, the total number of cover classes was defined as the number of classes found using at

least one cover estimation method (FVE, PVE or PPI), and the proportion of classes found using each

observation method was compared with this total. We assumed no species were misidentified. Since

the photographic interpretation was conducted after the end of the field season, there was no

opportunity to revisit plots to reconcile the species lists from the various cover estimation methods.

Previous studies have shown that misidentifications affected the reliability of some vegetation

surveys (Klimeš et al. 2001) though other studies have shown it to be less of a problem than

overlooking taxa (Scott and Hallam 2002; Archaux et al. 2006). Since the flora on Macquarie Island is

small (44 vascular plant species), we considered that the risk of misidentifications was small.

Errors of overlooking of classes were assessed by comparing pseudoturnover rates between paired

lists of the three sampling methods, following the formula of Nilsson and Nilsson (1985).

$$Pseudoturnover = 100 \times \frac{2n(A \cap B)}{n(A) + n(B)} \qquad [1]$$

where *A* and *B* are the species found with each sampling method. Pseudoturnover is the inverse of

the Sørensen Index, or Bray-Curtis similarity coefficient based on presence-absence data (Archaux *et*

al. 2009) and hence does not take account of the relative abundance of different species (Clarke

1993).

Linear regression was used to explore the relationship between pseudoturnover and taxon-richness

at a given plot (R Core Development Team 2010). The taxa were then grouped according to their

typical morphology (megaherbs, medium graminoids, mat-forming herbs, interstitial species and

non-vegetated ground) and the rate of overlooking was plotted for each morphological group.

To compare the reliability of the cover estimates, we used two approaches. First, one-way ANOVA

was used to test whether the three cover estimation methods produced significantly different cover

estimates for each morphological group. We then examined the effect of the differing cover

estimates on the ability of multivariate clustering to group all observations from each plot. For the

multivariate analyses, we followed a common approach in community ecology of using Bray-Curtis

distance matrices, non-metric multidimensional scaling (nMDS) and clustering to explore whether

the three different observation methods for each plot were more similar to each other than

observations from other plots (Bray & Curtis 1957; Field et al. 1982; Clarke 1993).

Bray-Curtis similarity matrices were calculated with four increasing levels of transformation:

untransformed data, square-root, fourth-root and presence/absence transformations (Clarke 1993;

Quinn & Keough 2002), to explore the relative importance of abundant, mid-range and rare species.

The nMDS ordination and all classification methods were repeated for all four transformations.

To ensure that the clusters were not artefacts of a single clustering algorithm, we applied three

different clustering methods: Partitioning Around Medoids (PAM), which is a form of fuzzy

clustering; and two hierarchical clustering tools (one divisive (*diana*) and one agglomerative (*hclust*))

were used with an average weighted linkage method (Maechler et al. 2005; R Core Development

Team 2010).

Three tools were used to test the validity of the clusters. Firstly, silhouette plots were used to test

how well each observation fitted its assigned cluster for all clustering techniques. These plots show

the average dissimilarity of each observation to all other observations in the cluster, compared to its

average dissimilarity to members of the closest neighbouring cluster (Rousseeuw 1987; Kaufman

and Rousseeuw 2005). Secondly, permutational similarity profile (SIMPROF) tests were applied to

the agglomerative hierarchical clustering dendrograms to determine which of the apparent clusters

were significant (Clarke et al. 2008). SIMPROF could not be applied to the diana and PAM clusters

using available tools. ANOSIM was used to test the null hypothesis that observations within each

plot were no more similar than observations from different plots (Clarke et al. 2006).

2.4.   Results

2.4.1. Field results

2.4.1.1.   Timing of measurements

Photo-sampling took approximately 20 to 25 minutes per plot. This compared with 25 to 50 minutes

per plot for field-based visual estimation (FVE) (incorporating 5 quadrats per plot) (Table 1). FVE

sampling times were influenced by the complexity of the vegetation, with species-rich short

grasslands taking the longest.

In the laboratory, Photographic Point Intercept (PPI) estimates took approximately 5 minutes per

photograph (i.e. 20 minutes per plot). Photographic Visual Estimation (PVE) of taxa cover required 5

– 30 minutes per photograph (i.e. 20 – 120 minutes per plot), with the variation determined by the

complexity of the vegetation (Table 2). As with FVE, species-rich short grassland plots took the

greatest amount of time.

Table 1: The three observation methods tested here, field visual estimates (FVE), photographic visual estimates (PVE), and photographic point intercept (PPI), all compromised between time investment and sampling comprehensiveness.

Method	Equipment Costs	Field Time (per plot)	Lab Time (per plot)	Mean Percentage of Total Species
FVE	Nil	25-50 mins	2 mins	83%
PVE	Low	20-25 mins	20-120 mins	76%
PPI	Low	20-25 mins	20 mins	55%

2.4.1.2.   Species presence and absence

The three cover estimation techniques varied significantly in their capacity to detect cover classes within the study plots ($p < 0.001$). FVE identified 83% ± 9.1% of the total number of cover classes per plot, compared with 76% ± 8.6% for PVE and 55% ± 12.6% for PPI (Fig. 2).
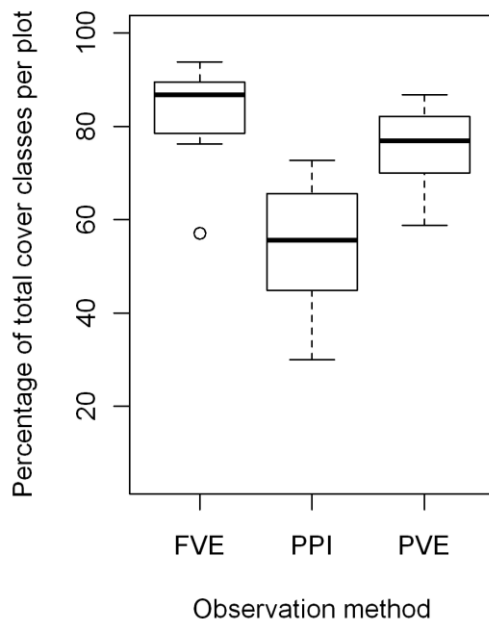
Figure 2: The total number of species per plot was calculated as the number of species identified in the plot using any one of the three observation methods. The percentage of total species found using each method demonstrates the comparatively poor performance of PPI. The bars represent the 5th and 95th percentiles.

Taxa that were detected at a given plot using only one of the three cover estimation techniques had very low coverage. The mean maximum coverage for a taxon found at a plot by only one observation method was 0.14% ± 0.26 for PPI, 0.19% ± 0.23 for PVE, and 0.8% ± 0.7 for FVE. Of the taxa that were only detected by a single observation method, 46% were interstitial species (most commonly *Cardamine corymbosa*, a rosette herb that is typically <2 cm in diameter); 25% were medium

graminoids (most commonly *Agrostis magellanica*, a morphologically variable species); and 18%

were mat-forming herbs. Of these, the fine-leaved grass *Deschampsia spp.* had the greatest extent,

covering 3.1% of the field quadrats (FVE) at one plot. This grass can be difficult to distinguish from

Agrostis magellanica, especially in photographs, and is likely to have been misidentified in the two

photographic observation methods. This is the only obvious example of a likely misidentification in

our dataset.

2.4.1.3.   Pseudoturnover

The mean pseudoturnover in this study was 24.6% ± 10.5 and was smallest between FVE and PVE,

and largest between FVE and PPI (Fig. 3). Plotting mean pseudoturnover against taxon richness

revealed two groups in the data. For the majority of plots, mean pseudoturnover increased with the

total number of species in a plot. Three plots with very low species richness, however, exhibited high

levels of pseudoturnover. These plots (MI0801, MI0802 and MI0804) had been recently denuded by

rabbits, and at least 96% of the cover at these three plots was provided by bare ground and one

taxon. Plants in these plots were only just becoming established, and consisted of very small,

sparsely located individuals. There was a significant linear relationship (β = 1.36, t = 2.36, p = 0.04)

between taxon richness and pseudoturnover for plots where the rabbit browsing damage appeared

older or where there were no signs of rabbit damage (all other plots) (Fig. 4). One outlier (MI009)

had strong leverage on the regression but there was no obvious biological difference that would

account for this outlier.
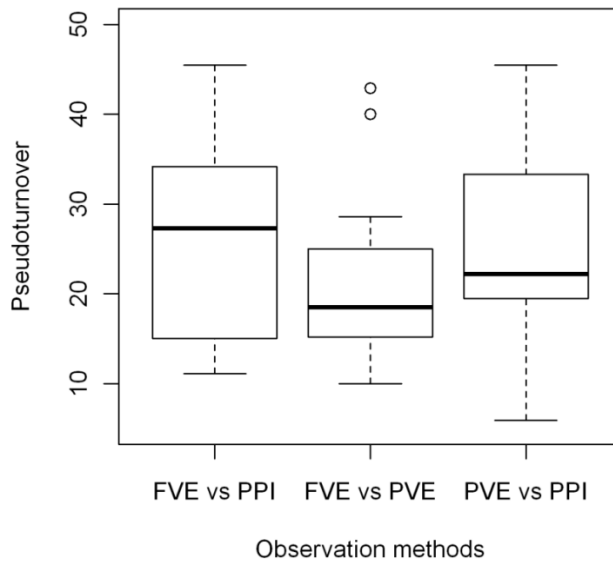
Figure 3: Relationship between pseudoturnover (%) and observation method (FVE=Field visual estimation, PVE = photographic visual estimation, and PPI = photographic point intercept estimation). The bars represent the 5th and 95th percentiles.
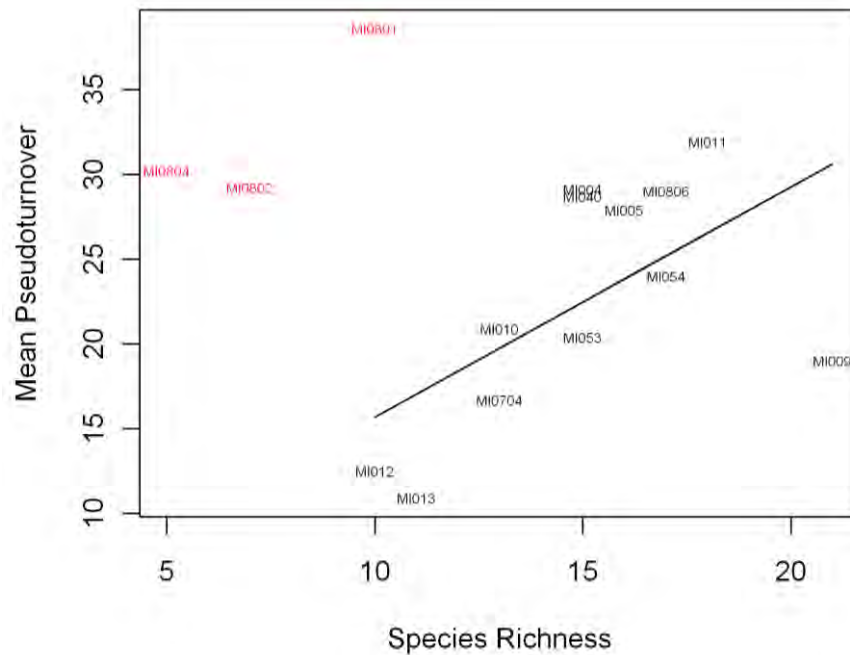
Figure 4: Relationship between pseudoturnover rates (%) and species richness. The three plots in the top left of the corner have been recently heavily grazed. The line shows the regression through all plots other than the three recently heavily grazed plots (slope = 1.36, p = 0.04).

2.4.1.4.   ANOVA comparisons

One-way ANOVA was used to test for differences between the three cover estimates for the total

cover estimates for each morphological group. Box-Cox transformation analysis indicated that

fourth-root transformation was appropriate for all groups. The cover estimation methods did not

differ significantly for any morphological group (Table 2).

Table 2: Results of one-way ANOVA tests for differences among the three cover estimation methods for each morphological group. For all analyses, there were two degrees of freedom between groups and 42 degrees of freedom within groups.

Morphological Group	F	P
Bare	0.003	0.997
Graminoids	0.002	0.998
Interstitial Herbs	0.623	0.541
Megaherbs	0.052	0.949
Mat Herbs	0.295	0.746

2.4.1.5.   nMDS plots

Non-metric multi-dimensional scaling (nMDS) showed that the variation between methods (FVE, PPI, PVE) was sufficiently small that intra-plot observations were exclusively grouped together and that there was good separation among plots when based on raw (linear R^2 = 0.766; non-metric R^2 = 0.956) or square-root transformed (linear R^2 = 0.842; non-metric R^2 = 0.968) data (Fig. 5). Separation among plots decreased as the strength of the transformation increased and the impact of abundance measures were lost in the analysis. The ordination exhibited the least separation for plots with very similar vegetation (complex short grassland communities dominated by a mix of *Agrostis magellanica* and *Luzula crinita*, e.g. MI0704, MI040, MI053 and MI009).
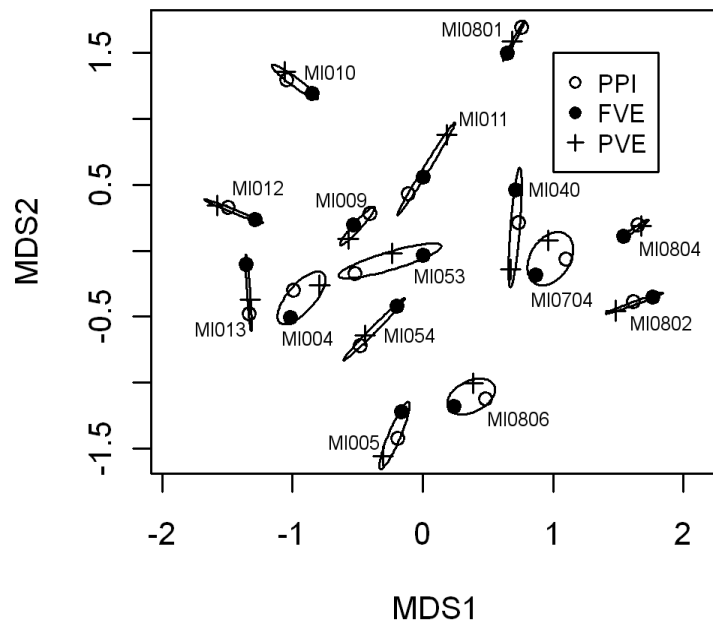
(a)
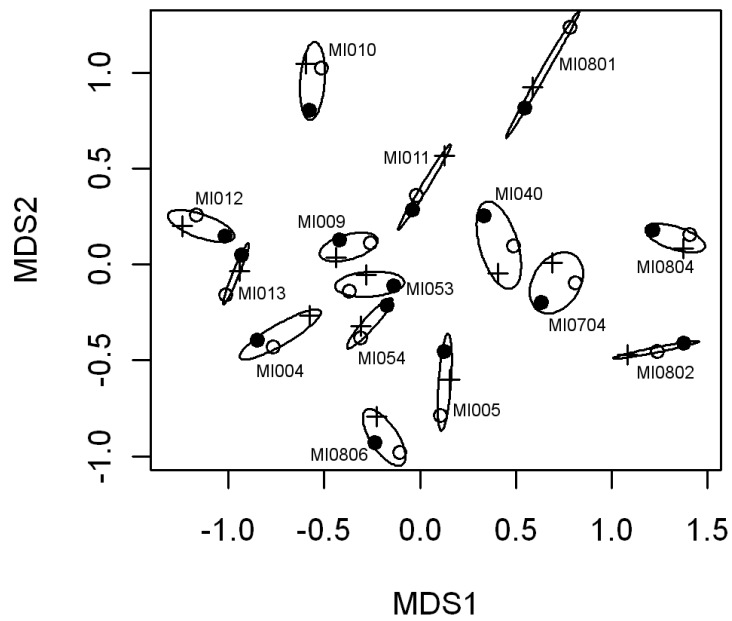
(b)

Figure 5: 2-dimensional non-metric multi-dimensional ordination of observations, based on (a) raw and (b) square-root transformed data, with 68% confidence limit ellipses to demonstrate the study plots. These show that the differences among the observation methods within each plot tended to be much smaller than the differences between the plots.

2.4.2. **Cluster analysis**

The hierarchical clustering algorithms based on untransformed, square-root transformed and fourth-root transformed data produced similar clustering results (Table 3). For the untransformed data, the two hierarchical clustering algorithms produced similar dendrograms (not presented here) and an MSW of 0.66, indicating a strong cluster structure. Neither hierarchical clustering algorithm, however, could separate plots MI004, MI009 and MI053, which contain very similar species-rich short grassland vegetation. Rates of misidentification and cover estimate errors were likely to be highest at these plots due to the species richness and the small size of the graminoid leaves, as found in other grasslands (Klimeš et al. 2001). Square-root transformed data could be used to correctly assign each cover estimate to its correct plot, regardless of which clustering algorithm was applied (Table 3 and Fig. 6), with a mean silhouette width (MSW) of 0.59, which indicates a reasonable cluster structure. When the data were subjected to a fourth-root transformation, the observations from most plots could be exclusively clustered, but the cluster structure was weak (MSW = 0.36 – 0.4) and there were no consistent patterns in the errors of the two hierarchical clustering techniques. Presence-absence transformed data could not be reliably clustered using hierarchical clustering algorithms (MSW = 0.27-0.28).

Table 3: The number of correctly clustered plots (all three observation techniques for a plot assigned to an exclusive cluster) organised by data transformation and clustering method. Maximum = 15.

	Validation*			
	Raw Data	**Square-Root Data**	**4th Root Data**	**Presence-Absence**
Divisive Clustering (diana)	12	15	13	6
Agglomerative Clustering (hclust)**	12 (9)	15 (12)	13 (10)	6 (2)
PAM	15	15	15	2

* Validation is a count of the number of clusters that contain all the observations for an individual plot and no other observations.

**Numbers in parentheses indicate SIMPROF statistically significant clusters.

(a)

(b)
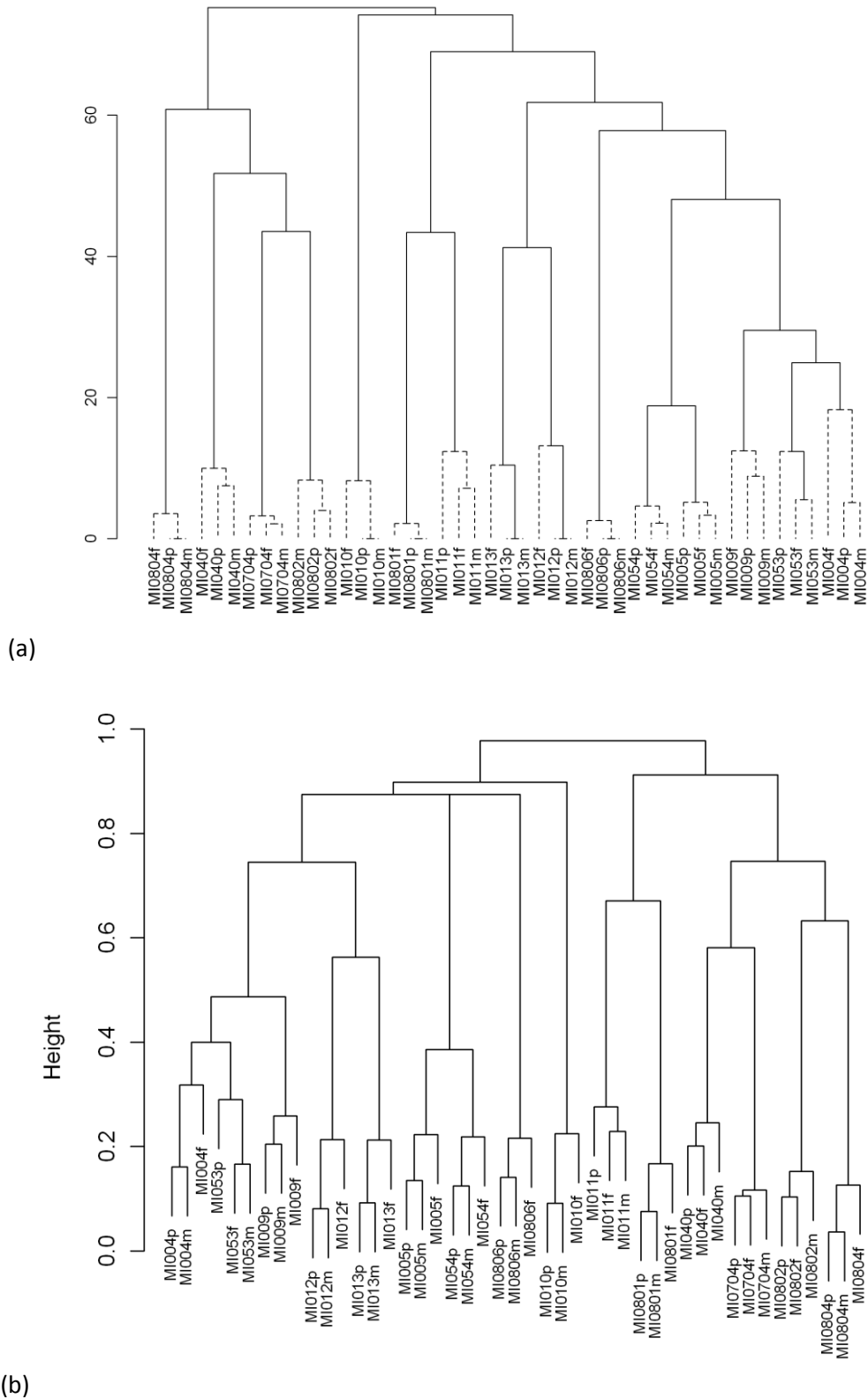
Figure 6: Dendrograms of (a) agglomerative hierarchical clustering and (b) divisive hierarchical clustering based on square-root transformed data. In (a) statistically significant differences are shown with solid lines. The statistical significance could not be calculated for (b). In both cluster analyses, the three observations within each plot were more similar to each other than to any observation in any other plot.

Partitioning around medoids (PAM) successfully assigned all observations to their individual plots when based on any transformation other than presence-absence. The silhouette plots (Fig. 7) show how well each observation fits its assigned cluster. While the mean silhouette width was largest for PAM based on raw data (0.64), the minimum silhouette width of -0.05 showed that the observations for plot MI004 could equally be assigned to other clusters. The slightly lower mean silhouette width (0.59) for PAM based on square-root transformed data was accompanied by a minimum silhouette width of 0.18. These plots showed that a reasonable partition structure was found (Rousseeuw 1987). As all clustering algorithms based on square-root transformed data produced identical groupings and were based on the same distance matrix, their silhouette plots were identical.

The mean silhouette width for PAM based on fourth-root transformed data was 0.46. PAM was unable to uniquely assign the presence-absence transformed cover estimates to their individual plots.

2.4.1. **ANOSIM**

Global analysis of similarities (ANOSIM) showed that grouping the cover estimates into their plots produced statistically significant clusters regardless of which data transformation was applied (R = 0.848 – 0.9993, P=0.0001).
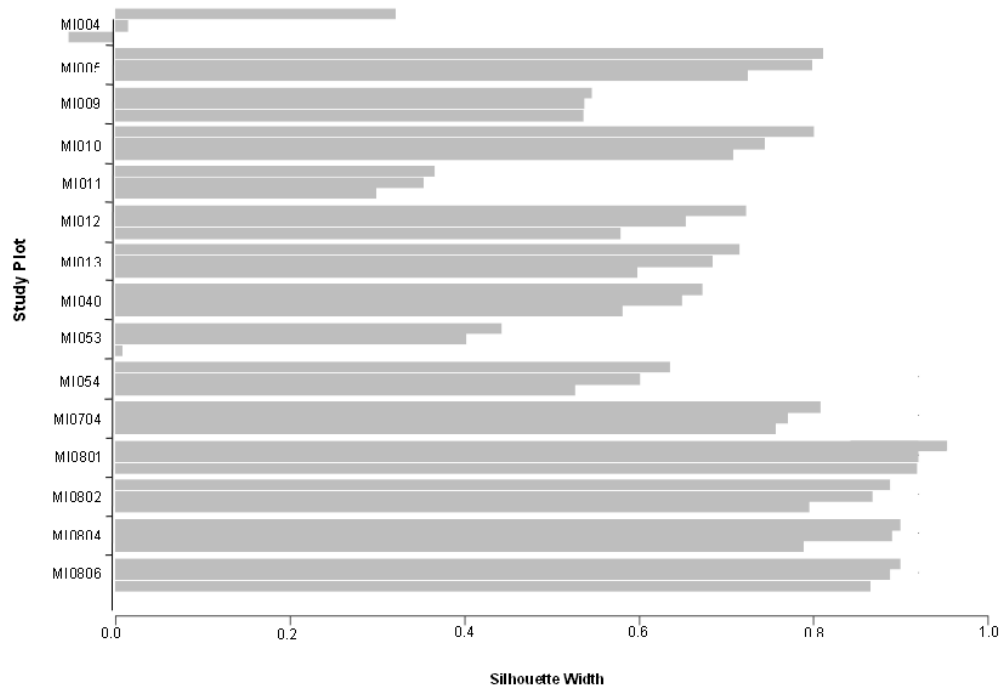
a)
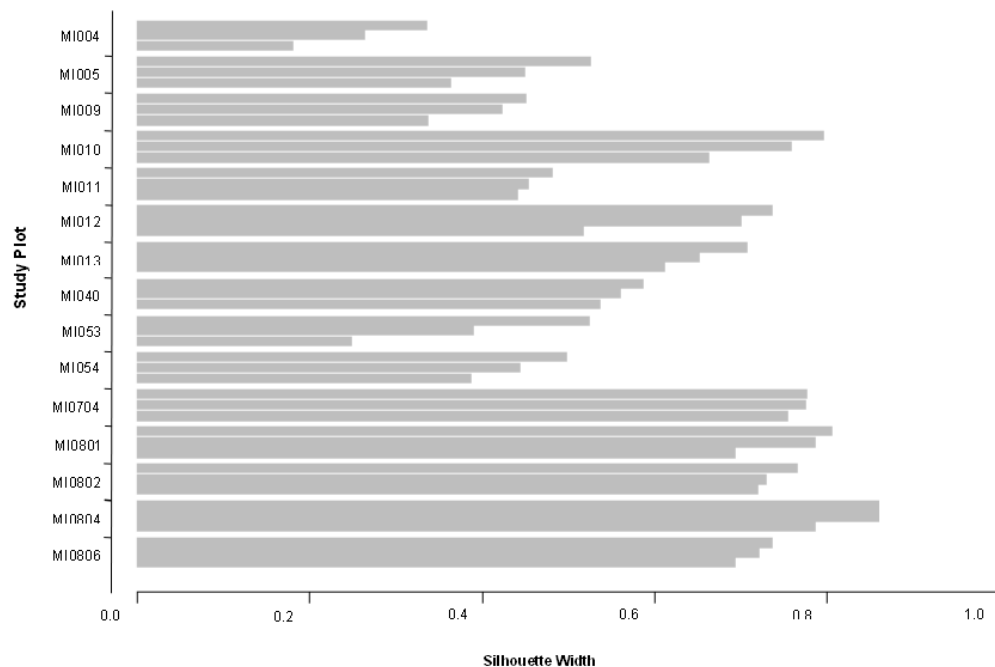
b)

Figure 7: Silhouette plots for Partitioning Around Medoids (PAM) based on (a) raw data and (b) square-root transformed data. A mean silhouette width greater than 0.5 indicates that a reasonable clustering structure has been found. A negative silhouette width indicates that an observation would have fitted better in the nearest neighbouring cluster. This occurred for a single observation (PPI) in

plot MI004, when the clustering was based on raw data (a). Overall, the mean silhouette plots show that the clusters were well separated.

2.5.   Discussion

Pseudoturnover rates among the three cover estimation methods tested here were similar to those found in inter-observer calibration studies elsewhere (Archaux et al. 2009). Multivariate clustering was generally able to group observations into their correct plots when species abundance was accounted for in the distance metric.

This study has shown that although close-range photo-sampling is less successful at identifying small and interstitial species than field visual estimates, it can be used successfully to collect quantitative data on individual species cover in low-growing dense vegetation with a relatively simple vertical structure. This is the first study that systematically tested the use of close-range photo-sampling for cover estimates of all species in a terrestrial environment with dense vegetation. Other researchers have shown that it can be used to quantify the cover of individual species in very sparse low-growing vegetation (Gilbert and Butt 2009) or to quantify the cover of selected, visually distinct species (Bennett et al. 2000; Vanha-Majamaa et al. 2000; Booth et al. 2008). Photo-sampling has been more extensively used in marine environments, where it has been most successfully used in benthic communities with simple vertical structure, and in deep water where few cover estimation methods are practicable (Foster et al. 1991; Drummond &Connell 2005; Tkachenko 2005; Alquezar & Boyd 2007). In benthic environments, PPI has been shown to underestimate species richness (Foster et al. 1991). This study extends the range of environments in which photo-sampling has been successfully used to quantify the coverage of individual species.

In general, pseudoturnover increased with species richness for plots that had not recently been denuded by rabbits. This, however, did not predict the ease with which ordination and clustering tools could exclusively group the cover estimates for each plot. Where the vegetation had recently been severely grazed (a combination of bare ground and a single taxon making >96% of the cover)

the rest of the vegetation consisted of small, sparse, newly established individuals that were easily missed due to the locations of quadrats or intercept points. The small size of the plants and the very low species richness at these plots inflated the pseudoturnover rate. These plots were easily separable using the clustering analyses.

Pseudoturnover and the proportion of total taxa found using each technique were affected by the difficulty in identifying interstitial species, especially in photographs. The differing proportions of the plots covered by each observation method (400 points, 20% and 100% of the plot area respectively) may also have contributed to pseudoturnover, but given that each of these cover estimation methods is supposed to characterise the vegetation of the entire 25 m^2 plot, we consider it valid to compare them directly. The mean pseudoturnover rates of 21.5 to 26.9% were higher than the best-case scenarios for studies examining inter-observer variability using a single cover estimation method (Nilsson & Nilsson 1985; Lepš & Hadincová 1992), but well within the range reviewed by Archaux et al. (2009). Given that the studies cited by Archaux et al. (2009) were comparing inter-observer differences using identical field protocols in a range of different vegetation types, including grasslands, it is encouraging to obtain similar levels of accuracy using different observation methods. In those studies, the mean Sørensen Index ranged from 66.6% to 89.1%, equivalent to a pseudoturnover rate of 10.9% - 33.3%. Of most immediate relevance, it is within the range of the Sørensen Index reported by Klimeš et al. (2001) in European grasslands (SI = 60-88%; PT = 12–40%).

Pseudoturnover is limited in its capacity to compare different observation methods because it does not account for the relative abundance of each species, as it is based on a presence-absence Bray-Curtis distance metric. ANOVA comparisons of the total cover for each morphological group showed no significant differences among each cover estimate, and this lack of difference was also reflected in the multivariate analyses. Wherever abundance measures were taken into account (i.e. all data transformations other than presence-absence) multivariate clustering tools were generally able to exclusively group observations to their individual plots, despite several plots containing apparently

similar vegetation. Macquarie Island's flora is small and distinguished by a small number of species that occur in all vegetation types at varying abundance levels. For example, the grass *Agrostis magellanica* occurs sparingly in almost every habitat from beaches to dry fell field, but dominates the canopy in some short dense grasslands. A small number of interstitial species occur sparsely in all vegetation types on Macquarie Island, and are responsible for much of the pseudoturnover. The presence-absence clustering attempts use pseudoturnover as a distance metric, and their failure to group observations to their correct plots highlights the limitations of pseudoturnover as a measure of the repeatability of vegetation surveys.

There is some evidence that multivariate analyses of vegetation data are more robust to sampling errors than univariate analyses like pseudoturnover (Carlsson et al. 2005). Lepš and Hadincová (1992) showed that in multivariate analyses, pseudoturnover had a negligible influence on the result compared with the choice of ordination method. Here, despite relatively high pseudoturnover rates, all clustering techniques and all but the strongest data transformation were generally able to exclusively group the various cover estimates from individual plots.

Choosing an appropriate cover estimation method involves balancing time limitations in the field and for data processing, as well as the accuracy of the results (Table 1). Photo-sampling is suitable for applications focussed on the dominant canopy species, and where field access is limited, compared with applications in which detailed information on the distribution of small, rare or cryptic species is required or field access is uncomplicated. Photo-sampling has a further advantage in that photographs are data rich and can be easily archived for use in future research (Benedetti-Cecchi et al. 1996; Laliberte et al. 2007).

A major incentive for this study was to generate a reliable method for gathering vegetation cover information from field plots to train a semi-automated vegetation classification of the whole island using very high resolution satellite imagery. The scale of the field samples has to match the scale of the satellite imagery for representative classification results. The field sampling technique proposed

in this study provides very detailed vegetation information for sufficiently large areas (i.e. corresponding to 4 multispectral pixels in a QuickBird satellite image) to rigorously ground-reference a range of image classification techniques, as recommended by Wilson et al (2011). The main disadvantage of the photographic sampling technique is that small and low growing species might be overlooked, especially if they make up a very small proportion of the plot. For remote sensing applications, such as mapping of plant communities or structurally dominant species, these types of errors are likely to have less impact than errors in species that dominate the canopy, and sampling efficiency is likely to be a more important consideration.

2.6.   Conclusions

This study was designed to test whether photo-sampling could be used as a replacement for field-based visual estimates of species cover on sub-Antarctic Macquarie Island. The results suggest that photo-sampling is a reasonable alternative to field quadrats in situations where field access is limited and where the vertical structure of the vegetation is relatively simple. Photo-sampling is faster in the field than field-based visual estimates (FVE), especially in complex vegetation, where FVE may take twice as long. This advantage is counter-balanced by increased data processing time, and the relative importance of these two time costs will depend on the difficulty of field access.

Two photo-interpretation methods (PVE and PPI) were compared to FVE. Both photo-interpretation methods underestimated species richness compared with FVE, with PPI performing worst. On average, PPI recorded just over half the taxa found in a plot using any method (55%), compared with 83% for FVE. The lower species detection rate did not affect the ability of the clustering algorithms to successfully allocate PPI estimates to their correct plot, and ANOVA showed no significant differences in cover scores.

Of the three clustering methods tested, PAM produced the most stable results – correctly classifying all observations with any transformation except presence-absence. Of all the data transformations

tested, square-root transformation allowed the correct clustering of all observations, using any of the clustering algorithms. This is likely due to the nature of the vegetation on Macquarie Island. Several species have an island-wide distribution, but their abundance is extremely variable. This variability is ignored by the strongest transformations. In contrast, the analyses based on untransformed data allowed the clustering to be largely determined by one or a few dominant species. Such classifications would be limited in their capacity to distinguish sites that share those few dominant species. It is therefore likely that future analyses on similar samples will be most robust if square-root data transformation is combined with PAM clustering.

Coupled with a species presence list collected *in situ*, photo-sampling may be an effective rapid form of data capture in tundra or other structurally simple vegetation types where large samples are required for tasks such as satellite image interpretation and where field access is expensive or difficult.

3. Geographically stratified random sampling design for vegetation survey: its implementation on sub-Antarctic Macquarie Island

3.1. Abstract

Training vegetation classifications and maps requires a field sampling design that captures the full range of both the vegetation types and the environmental variables likely to affect vegetation distributions. The sampling design must also be sufficiently flexible to allow for access limitations in remote environments. GeoStrat is a geographically stratified proportional random sample design that uses fuzzy *c*-means classification of very high resolution terrain variables and a vegetation index to define eight landscape strata. The proportions of field sites in each stratum are set according to three criteria: area, variance of the NDVI, and a combined subjective assessment of vegetation complexity and access limitations. GeoStrat was implemented on sub-Antarctic Macquarie Island over two summers. To test its effectiveness, it was compared to an existing purposive sample and three simulated random samples: unrestricted, equal stratified, and area-proportional stratified random samples. All samples were down-weighted to simulate the effect of differential access limitations across strata, and compared against a very large unrestricted random sample to examine their capacity to capture variation in the terrain and vegetation. GeoStrat performed as a compromise between the other randomised designs and the purposive sample, which was strongly biased towards the vegetationally complex coastal areas.

GeoStrat allows researchers to use their knowledge of the study area to account for likely field difficulties and hotspots of biological diversity. The flexibility in the design stage allows for flexibility in the field, and the use of fuzzy c-means clustering of environmental variables overcomes the limitations of existing landscape stratification methods for vegetation survey.

3.1.   Introduction

Mapping the distributions of plant species and communities requires a robust field sampling design that captures information on the full range of environmental variables likely to affect vegetation patterns. The sampling design must balance statistical representativeness (e.g. independence, randomness) against ecological representativeness, and needs to capture both diversity of the environment (i.e. a full range of landscape variables) as well as the distributions of plant species. Additionally, the field design must be flexible enough to cope with the practical realities of ecological fieldwork, where randomly located sites are often inaccessible.

Vegetation sampling designs generally fit within the following categories: unrestricted random sampling), stratified random sampling, regular or grid sampling, or purposive sampling. There are few clear guidelines available to help researchers choose the most appropriate sampling method. Unrestricted random sampling is sometimes highlighted as an ideal for a statistical sampling design, but is rarely applied in vegetation surveys, probably because it is inefficient in time usage and requires large samples to adequately capture rare vegetation and environmental types (Goedickemeier et al. 1997; Pillar 1998; Hirzel and Guisan 2002). Although it may adequately capture geographic space, this does not guarantee that the sample will not be biased in environmental space (Aspinall & Lees, 1994; Roleček et al. 2007).

A number of studies have explored whether various forms of stratified random sampling can improve the efficiency of surveys, increasing their capacity to adequately sample rare vegetation types and combinations of environmental variables, without sacrificing statistical validity (Guisan and Zimmermann 2000; Austin 2002; Chiarucci 2007). The aim is to locate samples so that they are randomly distributed in environmental space rather than geographic space (Goedickemeier et al. 1997; Pillar 1998). Sampling to identify communities or vegetation types is most efficient if each community is represented by an equal number of samples (Franklin et al. 2001). If community definitions do not already exist, then efficiency is tested by classifying the collected floristic data,

which risks circular reasoning. It has been argued that efficiency is highest when both the number of strata and number of vegetation types are high (Goedickemeier et al. 1997). To maximise efficiency, Fortin et al. (1989) recommend stratifying the study area into geographic zones in which the variances are approximately equal, and then concentrating the sampling in those zones with the highest variance.

The best-known stratified random sampling system is probably Gradsect, in which the study area is stratified according to environmental variables that are believed to be major drivers of vegetation patterns (Austin and Heyligers 1989). In this approach, the chosen environmental variables are converted from continuous values to discrete classes, and these classes are then combined to generate large numbers of relatively small strata, representing the various combinations of all classes of each variable. A decision is then made as to how much effort to direct to sampling relatively rare strata compared with common strata. This approach has been implemented in a range of Australian, European and North American environments (Franklin et al. 2001; Hirzel and Guisan 2002; Grabherr et al. 2003). For example, such an approach was shown to improve the efficiency of sampling in a Swiss alpine forest (Goedickemeier et al. 1997) and to represent the range of climatic, topographic and geologic variation and efficiently captured the range of plant species and vegetation alliances in the Mojave desert (Franklin et al. 2001).

Due to the way GRADSECT sampling designs categorise each variable before intersecting them, the landscape is divided into large numbers of small strata, with the number of strata increasing exponentially as additional environmental variables are included or the number of classes increases. A very simple stratification with two environmental variables and four classes in each would potentially produce 16 strata, but a more complex design with 5 environmental variables, each divided into four classes, could produce 1024 strata. If those five variables were divided into five classes, then 3125 strata could result. Given that poor stratification (typically one that creates large numbers of very small strata with unequal areas and a small sample in each) can make this approach

inefficient (de Gruijter 1999), the number of environmental variables and the number of classes in each must be limited for the stratification to remain practical. A further disadvantage of this system is that strata with rare attribute combinations may be artefacts rather than representing true environmental extremes and do not necessarily contain rare vegetation types (Goedickemeier et al. 1997). Thus, a stratification system that allows the incorporation of multiple variables, avoids arbitrary classification of continuous variables and does not produce large numbers of very small strata would be an improvement on this system.

Much vegetation classification and mapping is still conducted on the basis of purposive or other subjective sampling (Bergstrom et al. 2002; Faber-Langendoen et al. 2007). Purposive sampling is common because it is a convenient way to capture samples of apparent vegetation communities, and ensure that rare vegetation types are captured (Roleček et al. 2007). Such designs, however, have long been criticised because of the risk that the resulting datasets will be biased towards sites that are easily accessible, species-rich, spatially clustered, or contain species of particular interest to a given researcher and will hence violate the assumptions of many statistical tests making them unsuitable for predicting distributions (Ashby 1935; Chytrý 2001; Chiarucci 2007). Despite these warnings, purposive sampling is still popular in the community ecology literature (Chytrý 2001). It may be that the lack of consensus in the literature about the best sampling design, and a belief that truly random designs are impossible to implement in many environments due to accessibility and practicality concerns, are discouraging researchers from adopting randomised sampling designs.

A key problem in designing vegetation sampling regimes is that geographic space, environmental (or resource) space and taxonomic space (i.e. how species are distributed) are often different (Aspinall and Lees 1994). Thus, a sample that is representative of the full range of geographic coordinates may not be randomly distributed in environmental space. Testing how well a sample represents environmental space is comparatively straightforward, where spatial datasets of the key environmental variables are available. The distribution of the observations for each environmental

variable within the sampling sites can be directly compared to the distribution of that variable across the entire study area. In contrast, researchers rarely know the 'true' distribution of the species being studied, so it is not obvious how to test the ability of various sampling designs to represent species distributions (Hirzel and Guisan 2002).

This study is part of a broader research program to map vegetation patterns on sub-Antarctic Macquarie Island, using topographic variables and very high resolution (VHR) satellite imagery. Here, we describe GeoStrat, a topographically and spectrally-explicit proportional stratified random sampling design to collect field calibration data. The stratification is designed to avoid the creation of large numbers of small strata, and the proportions of points per stratum take account of likely access difficulties and variation in the vegetation, as well as the area covered by each stratum. We demonstrate its implementation in a remote, physically challenging environment, and test its capacity to capture the range of environmental variables on the island. To assess the ability of GeoStrat to represent environmental variation and to capture vegetation patterns, we compare it to simulated unrestricted, equal-stratified and area-proportional stratified random sampling designs and the purposive sampling design used in a long-term vegetation monitoring program on the island.

3.2. Methods

3.2.1. Study site

Macquarie Island (54°30'S, 158°57'E) is located in the Southern Ocean, approximately 1500 km south of Australia's mainland (Fig. 1). The island is dominated by a plateau between 200 and 400 m MSL, with steep escarpments leading down to beaches and, on the north-west coast, a raised coastal terrace. The vegetation consists of 44 vascular plant species and approximately 135 bryophyte species, along with lichens, algae and fungi (Selkirk et al. 1990). The most exposed parts of the plateau are characterised by open feldmark, which consists of an extremely variable mixture of gravel, cushion plants and bryophytes interspersed with lesser amounts of other vascular plant

species. Less exposed sites are characterised by short grasslands that are dominated by several grass species individually or in combination. Areas where the water table is close to the surface contain bryophyte- and sedge-dominated mires. The coastal slopes are host to a complex array of short grasslands, heavily disturbed or grazed sites and the tall tussock and megaherb-fields for which Macquarie Island is known (Selkirk et al. 1990).
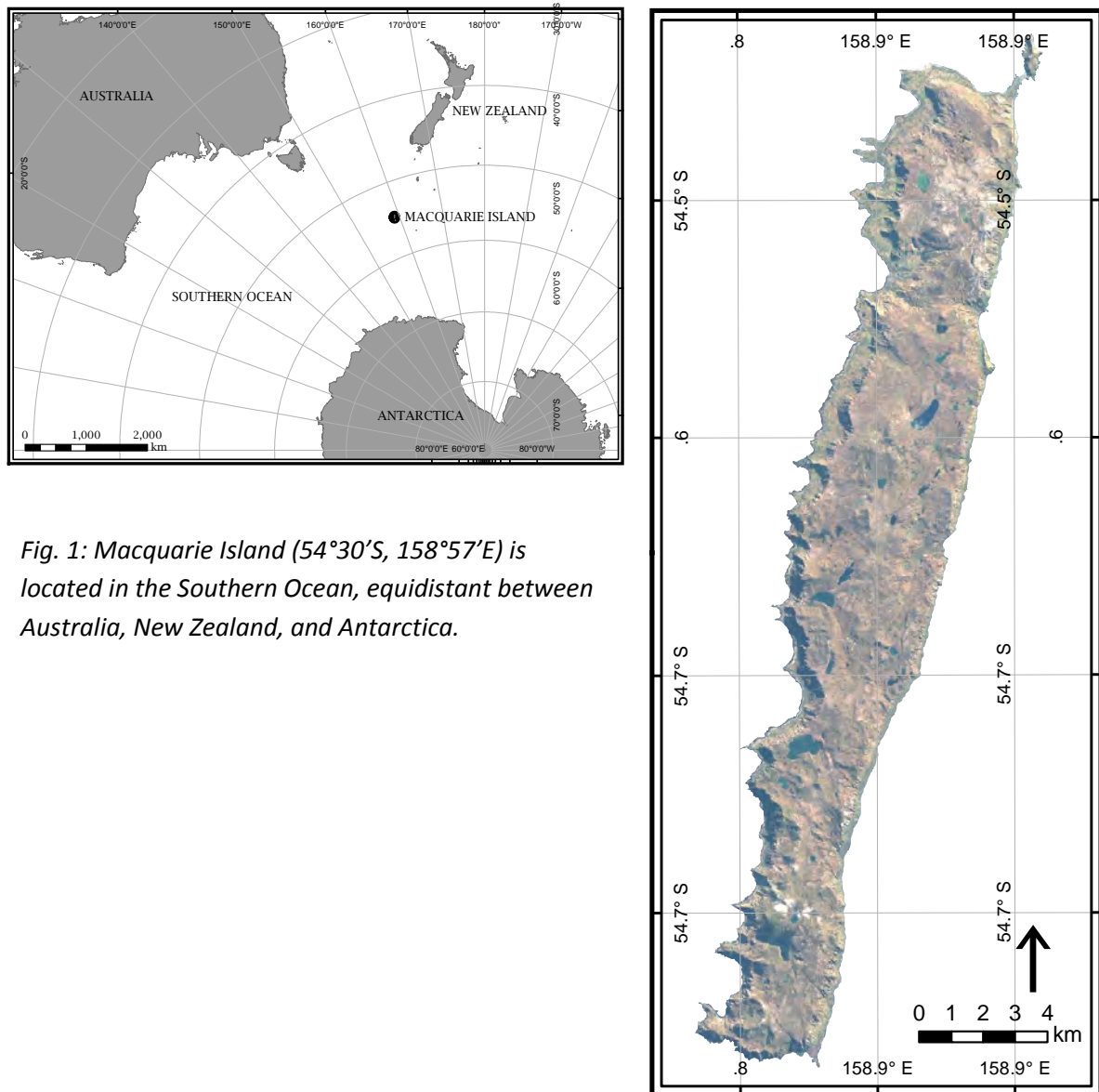
Fig. 1: Macquarie Island (54°30'S, 158°57'E) is located in the Southern Ocean, equidistant between Australia, New Zealand, and Antarctica.

3.2.2. **Digital datasets**

Elevation values were taken from the Australian Antarctic Division Digital Elevation Model (DEM) of Macquarie Island, a 5 m resolution raster, with heights accurate to 5 m, derived from Airborne Synthetic Aperture RADAR images acquired in 2000 by the NASA PACRIM Mission 2 (Tapley et al. 2004; Brolsma 2008).

A cloud-free QuickBird image of Macquarie Island was acquired on 15 March 2005. The QuickBird satellite sensor captures four multispectral bands (blue, green, red and near-infrared (NIR)) at 2.4 m resolution.

Vegetation distribution data were taken from a vegetation map that was digitised from 20 m resolution SPOT satellite imagery (Selkirk et al. 2000). This map shows the structure of the vegetation, rather than species composition, though the small number of vascular plant species on Macquarie Island makes it possible to link dominant canopy species to the structural groups identified in this map.

3.2.3. **Fuzzy *c*-means landscape stratification**

Vegetation distribution is known to be affected by the shape of terrain and its resulting impact on microclimate and soil moisture (Franklin 1995; Goedickemeier et al. 1997). Therefore, an efficient sampling strategy should be stratified according to those gradients believed to be major drivers of species distributions (Hirzel and Guisan 2002). Gradsect-style stratified random sampling designs are limited by the simple way in which the variables are combined to generate the strata. Converting continuous variables, such as elevation, slope, and solar radiation into discrete classes creates arbitrary divisions in the data. Furthermore it produces exponentially larger numbers of strata as the number of environmental variables or their classes increase. To avoid these problems, we looked to the field of morphometric classification to improve the stratification of the landscape into

homogeneous units (Pike 2000; Fisher et al. 2004). Terrain data and a satellite image-derived

vegetation index were used to stratify the island in an attempt to capture the full range of habitats.

Traditionally, clustering algorithms group samples into distinct classes with hard or discrete

boundaries. Terrain and vegetation characteristics often vary gradually over space, however, and a

fuzzy approach towards identification of strata might be more appropriate (Burrough and Frank

1996). The fuzzy *c*-means (FCM) clustering algorithm identifies groups of samples sharing similar

characteristics in a multivariate feature space (Bezdek et al. 1984). FCM produces overlapping fuzzy

clusters with transition zones. Mapping fuzzy membership values for each grid cell results in a

membership layer for each class; the higher the membership of a grid cell, the greater the likelihood

that the grid cell belongs to a certain class. Finally, the FCM output is 'defuzzified' by assigning hard

class labels to each of the grid cells, based on the maximum membership value.

Following Burrough et al. (2000, 2001) we used an unsupervised fuzzy *c*-means clustering algorithm

to divide Macquarie Island into relatively homogeneous landscape strata. The classification was

based on six terrain parameters (elevation, slope, surface curvature, wetness index, solar radiation

and topographically deflected wind speed) derived from the DEM and a normalised difference

vegetation index (NDVI) layer derived from the QuickBird imagery. Table 1 provides a summary of

the derivation of each layer.

The wetness index is sensitive to errors in the DEM as a small error in height can significantly change

a cell's flow direction, slope and flow accumulation (van Niel et al. 2004). To account for this, we

used the mean value from a Monte Carlo simulation with 100 realisations, where the error term was

based on a Gaussian error model with a mean of 0.0 and a standard deviation of 5 m, set from the

estimated root mean square error of the DEM (Tapley et al. 2004; Oksanen and Sarjakoski 2005;

Brolsma 2008; Januchowski et al. 2010). The topographically deflected wind speed layer was primed

using long-term averages for the mean wind speed (35.1 km/h) and direction (WNW) (Bureau of

Meterorology 2011).

All GIS analyses were conducted in PCRaster and ArcGIS 9.3 (ESRI 2008; PCRaster 2011). The remote

sensing software ENVI 4.7 was used to calculate the normalised difference vegetation index (NDVI)

and the fuzzy c-means algorithm was implemented in IDL (ENVI 4.7).

Table 1: The derivation of the spatial attribute layers used for fuzzy c-means classification of the landscape into homogeneous strata.

Input Layer	Description	Calculation	Reference
Elevation	Values extracted from the DEM	N/A	(Tapley et al. 2004; Brolsma 2008)
Slope	Average maximum rate of change in a 3 x 3 neighbourhood of the DEM	slope° = ATAN($\sqrt{([\Delta z/\Delta x]^2 + [\Delta z/\Delta y]^2)}$)*57.29578	(Horn 1981; Gallant and Wilson 2000)
Surface Curvature	A second-order derivative of the DEM, surface curvature of the DEM is a composite of planar and profile curvature.	A fourth-order polynomial is fitted to the surface of a 3 x 3 cell neighbourhood	(Zevenbergen and Thorne 1987; Moore et al. 1991)
Wetness Index	A measure of the relative position of a cell within a catchment, wetness index is a function of the upstream area and the slope of the target cell. It is used as proxy for variability in soil wetness and assumes a uniform substrate and ground cover.	WI = ln(A/tanß), where A = upstream area and ß = Slope in degrees	(Beven & Kirkby 1979; Wilson & Gallant 2000; Burrough et al. 2001; van Niel et al. 2004; Sørensen et al. 2006)
Solar Radiation	Modelled solar radiation, accounting for latitude, topography and atmospheric conditions over a year. Sky conditions in the model were set assuming a generally cloudy sky, as Macquarie Island experiences an average 289 cloudy days each year.	For each cell, an upward-looking hemispherical viewshed is calculated based on topography; a direct sun map and diffuse sky map are overlaid on this for each time interval (here, for every 30 minutes in a year) to calculate a composite insolation value in Wh/m^2.	(Fu & Rich 2000; 2002; Huang et al. 2008; Bureau of Meteorology 2010)
Wind Speed	Topography has a significant effect on wind direction and speed. An input wind speed is weighted by a function of surface slope and curvature. The sign of the weight for any given cell is determined by whether that cell is on a windward or lee slope, as determined by a topographically deflected wind direction.	The mean prevailing wind is weighted by the factor R, where R = 1.0 + 0.6*μs + 0.4*μc, where μs = slope scaled to -0.5<μ<0.5; and μc = curvature scaled to -0.5<μ<0.5. The weighting is positive for windward and convex slopes and negative for lee and concave slopes, as determined by the topographically deflected wind direction D, where D = -0.255Ssin(2(A-θ)), where S = slope (%); A = Aspect (°); and θ = initial wind direct (°)	(Liston and Sturm 1998; Wallace 2006)
Normalised Difference Vegetation Index (NDVI)	NDVI calculates the relative proportions of red and near-infra red reflections as a proxy for the amount of live vegetation in a pixel of a satellite image.	NDVI = (NIR - RED) / (NIR + RED)	(Sellers 1985)

3.2.4. **Setting the proportions**

Generally, a proportional stratified random sampling scheme assigns the number of sites per stratum to match the proportion of the study area covered by each stratum (Hirzel and Guisan 2002). On Macquarie Island, however, much of the variation in vegetation types occurs in the coastal areas, which make up a small proportion of the island's surface area. The coastal slopes are topographically complex and subject to frequent landslips, and at the time of study, exposed to extensive grazing by introduced rabbits. This has produced a complex mosaic of plant assemblages at different successional stages. Fortin et al. (1989) recommend that sampling effort be concentrated within those strata that exhibit the highest levels of variation. In addition, we knew in advance that parts of the coastal slopes would be unsafe or inaccessible due to the presence of nesting sea birds, so there was a need to generate extra sites to compensate for access limitations. This corresponds with warnings by other researchers that truly random samples are often not possible to collect in real-world field studies (Lepš and Šmilauer 2007; Roleček et al. 2007).

To balance these sampling concerns, we considered three criteria in determining the proportion of sites to be located in each stratum. We ranked the strata according to area; standard deviation of the NDVI as a surrogate for variation in the density of the vegetation; and a subjective assessment of the likelihood of accessibility issues and variation of vegetation types within each stratum. The ranking for each criterion was calculated using ranked-pairs voting, a technique used in multi-criteria decision analysis (e.g. Feick and Hall 2002). Every possible pair of strata was subjected to pairwise voting for that criterion, with the higher ranked stratum receiving a score of 1 for each pairwise vote it won. Thus, for each criterion, a total of 28 votes were available. The three criteria resulted in contradictory rankings. To create a composite ranking, the scores from the pairwise voting were weighted and summed, with the weights set according to the perceived importance of each criterion. Given the observed variability in coastal vegetation and the likely difficulty of accessing coastal areas, the access/variability criterion was assigned a weight of 0.6. Surface area and the

standard deviation of NDVI within each stratum were each given weights of 0.2. The weighted

rankings were then summed to calculate the proportion of all planned sites to be located in each

stratum.

3.2.5. **Assigning random point locations**

A random sampling tool was used to generate 380 sites within the strata, using the proportions

derived from the composite rankings (Beyer 2004). Since the ecological processes that drive

vegetation patterns are subject to spatial autocorrelation (Ricotta 2007), we attempted to improve

the statistical independence of the sites by enforcing a minimum distance between sites within a

stratum, following a suggestion from Guisan & Zimmermann (2000). To explore the spatial

autocorrelation of the environmental parameters used in the landform classification, we constructed

semivariograms of each parameter. The distance at which sites were spatially autocorrelated ranged

from 100 m for the wetness index to 2 km for elevation. To balance the need for efficient sampling

against minimising the effects of spatial autocorrelation in the environmental gradients, a minimum

distance of 100 m was enforced between sites within each stratum.
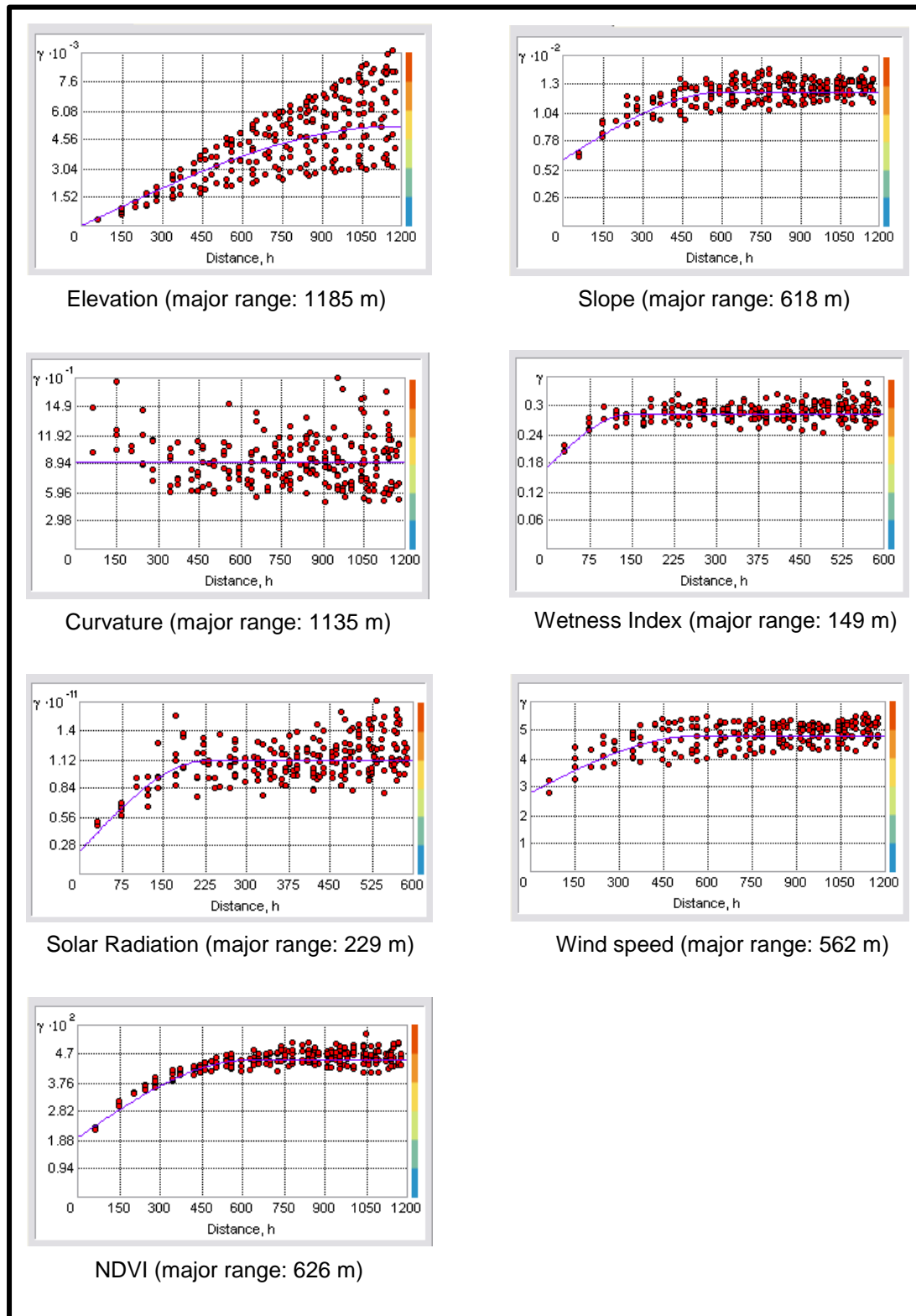
Figure 2: Semivariograms of the input variables for the fuzzy c-means classification of landscape strata. Spatial autocorrelation for each variable occurs at distances less than the major range. Thus, spatial autocorrelation occurs within 149 m for the wetness index, and 1185 m for elevation.

3.2.6. **Field application**

Macquarie Island was visited twice during the summer months of 2008/09 and 2009/10 and 349

sites were surveyed. Because inter-annual vegetation dynamics are relatively low for most parts of

Macquarie Island, we combined data from the two seasons (Sluiter and Pebesma 2010). Those sites

that were determined *a priori* to be in inaccessible locations were not visited (e.g. halfway down a

200 m cliff or in an albatross breeding site). If a site was approached and found to be inaccessible, it

was relocated by up to 30 m if a safer site with similar vegetation could be found. If not, that site

was discarded from the sample.

3.2.7. **Comparison samples**

We compared the ability of GeoStrat to capture environmental variation against a large control

sample and four other sampling methods – 1) an existing purposive sampling design; 2) an area-

proportional stratified random sampling; 3) an equal-stratified random sampling; and 4)

unrestricted random sampling designs. All the randomised samples included 380 sites and were

generated using the same sampling tool as used for GeoStrat (Beyer 2001-2011). For area-

proportional stratified random sampling, the number of points per stratum was set according to the

proportion of the land area covered by each stratum. The total number of points in each sample was

the same as the planned GeoStrat points. To represent the true distribution of environmental

variation, a control sample was generated with 5000 unrestricted random points.

To simulate the effect of testing these sampling regimes in the field, we reduced the size of each

sample by the same proportion of sites in each stratum that were excluded from the planned

GeoStrat due to access problems. For example, 35% of planned sites in stratum 2 (coastal flats) could

not be visited. Thus, for all sampling regimes, 35% of the points in stratum 2 were randomly

removed, to produce a simulated field sample.

Purposive site locations were taken from a long-term vegetation monitoring project on Macquarie

Island (see Bergstrom et al. 2009a). By definition, all of these sites were accessible in the field, so this

sample was not reduced as the others were. This sample was designed to capture representative samples of perceived major communities on the island and is dominated by sites in coastal areas, where the vegetation is extremely variable, and away from the exposed parts of the plateau that are dominated by the floristically-poor feldmark.

3.2.8. **Testing the performance of the sampling designs**

To test how well each of these sampling regimes captured variation in the terrain, for each point we extracted the values for the terrain and image variables used in the stratification: elevation, slope, surface curvature, solar radiation, wind speed, wetness index, and NDVI.

Empirical cumulative distribution function plots showed the distribution of each sample for each continuous variable. Kolmogorov-Smirnov tests were used to test whether each sample's distribution matched the distribution of the control sample. All analyses were conducted using the *stats* package in R (R Core Development Team 2010).

As a partial test of the ability of the sampling designs to capture vegetation patterns, we extracted the vegetation type for each sampling point from an existing 1:25 000 scale map of vegetation structures (Selkirk et al. 2000). We plotted the proportion of each field sample in each vegetation class and in each stratum of the stratification scheme.

3.3.   Results

3.3.1. **Stratification**

Seven strata were generated by fuzzy *c*-means classification of the terrain variables and NDVI (Fig. 3). Visual examination of the stratification shows that it clearly separated coastal zones and mountain tops from other parts of the landscape. Stratum 1 (s1) covered 2261.7 ha of mountain tops and high plateau areas with generally very sparse vegetation, especially those areas exposed to the prevailing westerly winds. Stratum 2 (s2) (1809.6 ha) covered the vegetated coastal flats but

excludes coastal rocky outcrops and some beaches. Strata 3 (s3) (2192.2 ha) and 4 (s4) (1137.9 ha) had patchy distributions throughout the plateau area, appearing to serve as boundary zones between the high exposed sites of s1 and the mid-altitude grasslands of stratum 5 (s5). S3 occurred around the fringes of s1, while s4 appeared to be transitional between s3 and s5. Stratum 5 (2024.6 ha) covered the mid-altitude short grasslands and dense herbfields, and includes the most gentle east coast slopes down to near sea level. The steep coastal slopes were divided between strata 6 (s6) (1277.4 ha) and 7 (s7) (1686.4 ha). The division between these two strata appeared to be largely driven by shadow effects in the QuickBird image. Stratum 6 covered most of the steep east coast slopes and the base of slopes on the west coast that are not facing south-west (and hence obscured by shadow). Stratum 7 was largely defined by the low NDVI, and it hence included the steepest and least vegetated slopes, shadowed areas, lakes, beaches and coastal rocks.

Combining votes from three criteria for ranking the number of observations in each class increased the number of sites in coastal regions (Table 2 and Fig. 4). Stratum 7, one of the coastal slope strata was ranked lowest in the subjective assessment of the likelihood of environmental change and access problems. This ranking was because a large part of the stratum contains very steep bare rocky ground, and was hence considered unlikely to change significantly. During fieldwork, however, it became apparent that both of the coastal slope strata were extremely difficult to access and contained complex vegetation. In hindsight, s7 should have been given higher ranking.
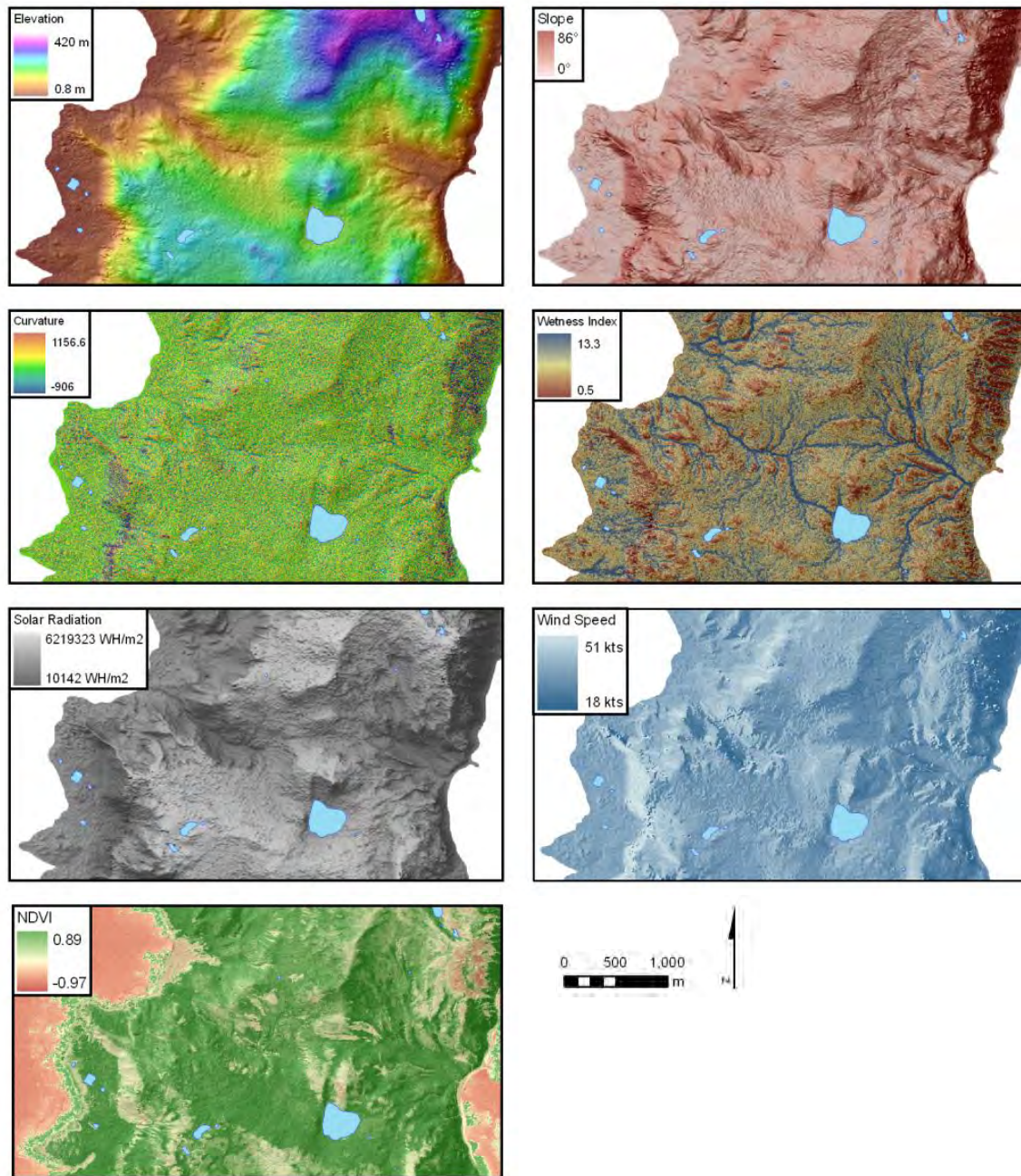
Figure 3: The environmental parameters used to drive the fuzzy c-means classification of the landscape. The first six parameters were derived from a 5 m resolution DEM, and the NDVI was calculated from a 2.5 m resolution QuickBird satellite image.
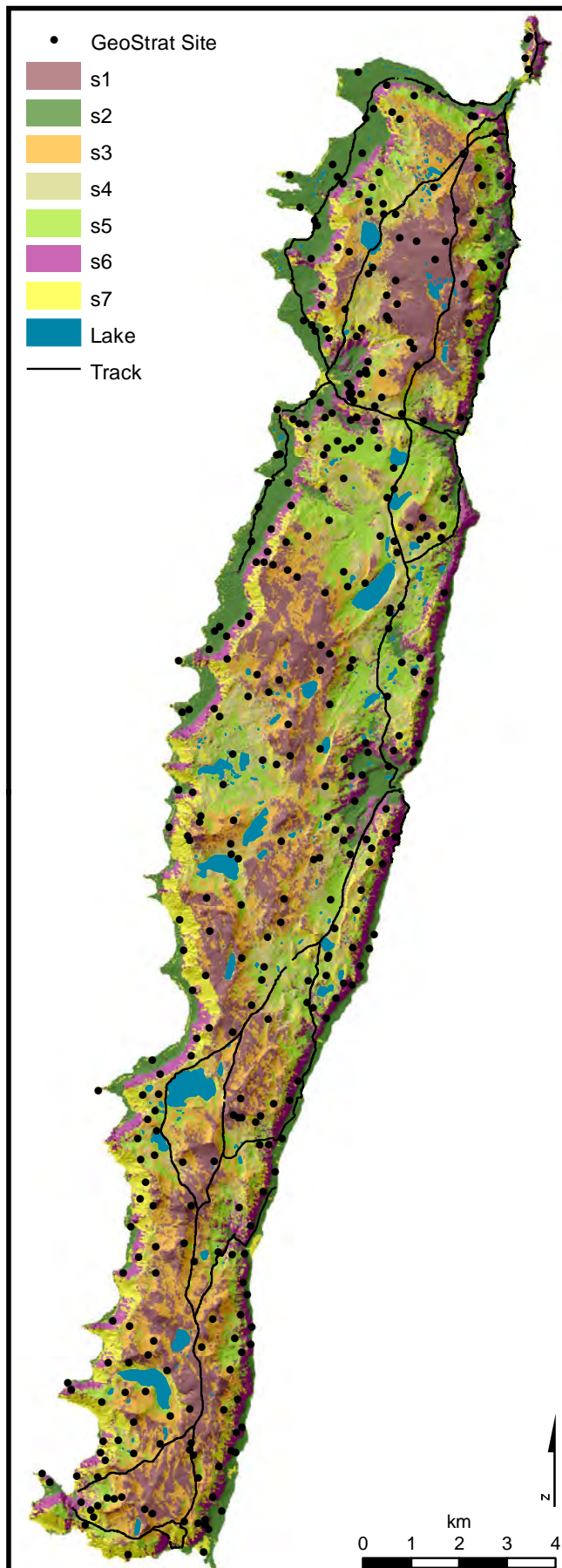
Figure 4: Fuzzy c-means classification of the input terrain and NDVI variables produced seven strata. 380 sites were located across the island, using proportional stratified random sampling.

Table 2: Composite rankings of strata as determined by pairwise voting show that no single criterion determined the proportion of sites to be located in each stratum.

Stratum	Primary Rankings			Composite Ranking	Sites per Stratum
	Area	NDVI	Access/Variation		
1: Mountain top	1	3	6	5	44
2: Coastal flat	4	2	3	2	65
3: Plateau I	2	6	4	4	51
4: Plateau II	7	5	5	7	33
5: Mid-altitude grassland	3	7	2	3	59
6: Coastal slope I	6	4	1	1	87
7: Coastal slope II	5	1	7	6	41

The strata produced for GeoStrat have extensive overlaps with the broad structural vegetation classes of Selkirk et al. (2000) (Table 3). The mountain tops (s1) were almost entirely covered in open-canopied short vegetation. The two plateau strata (s3 and s4) which fringe the mountain tops contained varying mixtures of open- and closed-canopy short vegetation. Stratum 5 was almost entirely covered by closed vegetation. The three coastal strata (s2, s6 and s7) all contained complex mixtures of vegetation structures.

Table S1: Comparison between the strata defined through fuzzy c-means classification of spatial data and the dominant vegetation structures, as mapped by Selkirk et al. (2000).

Stratum	Dominant vegetation types	Floristic description
1: Mountain top	open short (91%)	Open feldmark
2: Coastal flat	closed (52%); closed tall (22%); closed short (14%)	Complex mosaic of dense herbfields, mires and megaherbs
3: Plateau I	closed short (52%); open short (47%)	Complex mix of dense grasslands and herbfields, opening out into feldmark
4: Plateau II	open short (75%); closed short (23%)	Closed feldmark, tending towards short grasslands
5: Mid-altitude grassland	closed short (88%)	Short grasslands and herbfields
6: Coastal slope I	closed short (41%); vegetation complex (34%); closed tall (14%)	Complex mosaic of dense herbfields and megaherbs
7: Coastal slope II	vegetation complex (44%); open short (38%); closed short (10%)	Mosaic of megaherbs and rocky outcrops with small areas of short grassland and herbfield

3.3.2. **Field application**

A total of 288 out of 384 (75%) planned GeoStrat sites were visited, with a total of 26 sites being relocated by up to 30 m to avoid inaccessible areas. The 96 excluded sites were either physically inaccessible (43 sites) or could not be visited due to time constraints (53 sites). During the two field seasons, all 58 purposive sites from the long-term vegetation monitoring program were also visited. The three coastal strata were least accessible, with visitation rates between 51% and 65%, compared with 98% for the mountain tops (Table 4).

Table 4: The number of GeoStrat sites that were planned and visited for each stratum.

Stratum	Planned	Visited	Proportion Visited
1	44	43	0.98
2	66	43	0.65
3	52	37	0.71
4	33	30	0.91
5	59	56	0.95
6	89	58	0.65
7	41	21	0.51
Total	384	288	0.75

3.3.3. **Comparison with other sampling designs**

Using the proportions in Table 4, the down-weighting of the planned random samples to account for access problems in the field reduced the total sample size by 75-78.5%. The purposive sites were not down-weighted because the sites were all, by definition, accessible.

The coastal vegetation on Macquarie Island is restricted in area but extremely variable, ranging from rocky outcrops, through mires and short grasslands to dense stands of the tussock grass *Poa foliosa* and the megaherb *Stilbocarpa polaris*. In contrast, the feldmark is extensive but floristically simple, being dominated by a mixture of the cushion plant *Azorella macquariensis*, bryophytes and gravel. The purposive sample was strongly biased towards the coastal areas in order to capture this variation.

Plotting the proportions of the samples within each stratum showed that all the randomised samples had similar distributions to the control sample and none performed consistently better or worse than the others (Fig. 5). All designs under-sampled coastal slopes II (s7), which covered the steepest coastal slopes with low NDVI. This stratum contained very steep and unstable scarps, which were difficult to access safely. To fully compensate for the large proportion of inaccessible sites in s7, the number of sampling sites in this stratum would need to be significantly increased. GeoStrat is the only randomised sampling system tried here that could be adjusted in this way. GeoStrat over-sampled coastal slopes I (s6), but was otherwise very close to the control sample.

The purposive sample showed the largest deviations from the control, heavily over-sampling the vegetationally complex coastal flats (s2) at the expense of other strata, especially the mountain tops (s1) and plateau II (s4). These two strata have relatively little variation in vegetation. The biggest differences in sample size among the designs were in the mountain tops (s1), which were over-sampled by area-proportional and equal- stratified random sampling, and under-sampled by GeoStrat and the purposive samples.
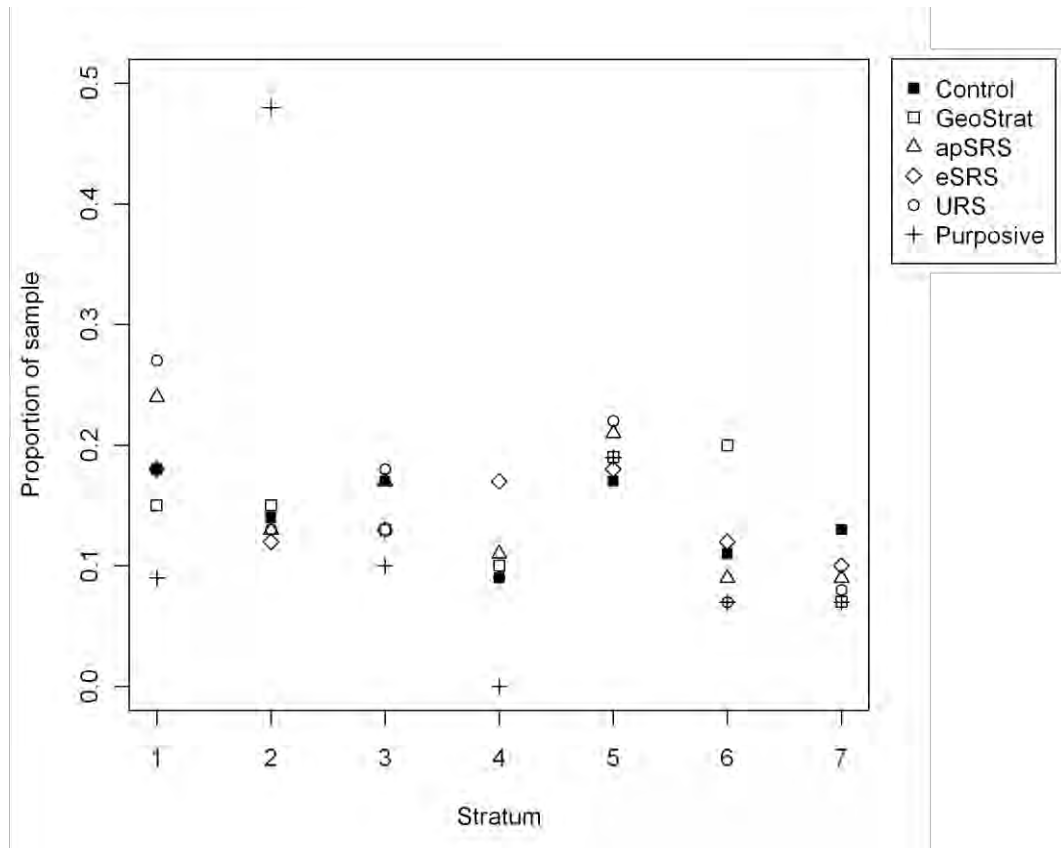
Fig. 5: Proportion of each stratum in the total sample of visited sites, according to the sampling design. In the legend, apSRS = area-proportional stratified random sample, eSRS = equal-stratified random sample, and URS = unrestricted random sample. Among the randomised sampling methods, the greatest variation was in s1, the largely bare mountain tops. These are not especially ecologically variable, and were over-sampled by URS and apSRS.

Plotting the proportion of sites in the vegetation classes mapped by Selkirk et al. (2000) showed close agreement between all the randomised samples and the control (Fig. 6). The sampling designs were most variable in the most common vegetation types. There was little difference among the sampling designs in the number of sites in rare vegetation types, such as mire and closed vegetation, but large variation in the number of sites in the common closed short herb and open short vegetation, which collectively cover almost 80% of the island. GeoStrat slightly over-sampled the tall closed vegetation in coastal areas, and under-sampled the open short vegetation, which represents the feldmark that dominates the high parts of the island. The purposive design included twice as many sites in closed tall vegetation and half as many sites in the open short vegetation than the randomised designs.
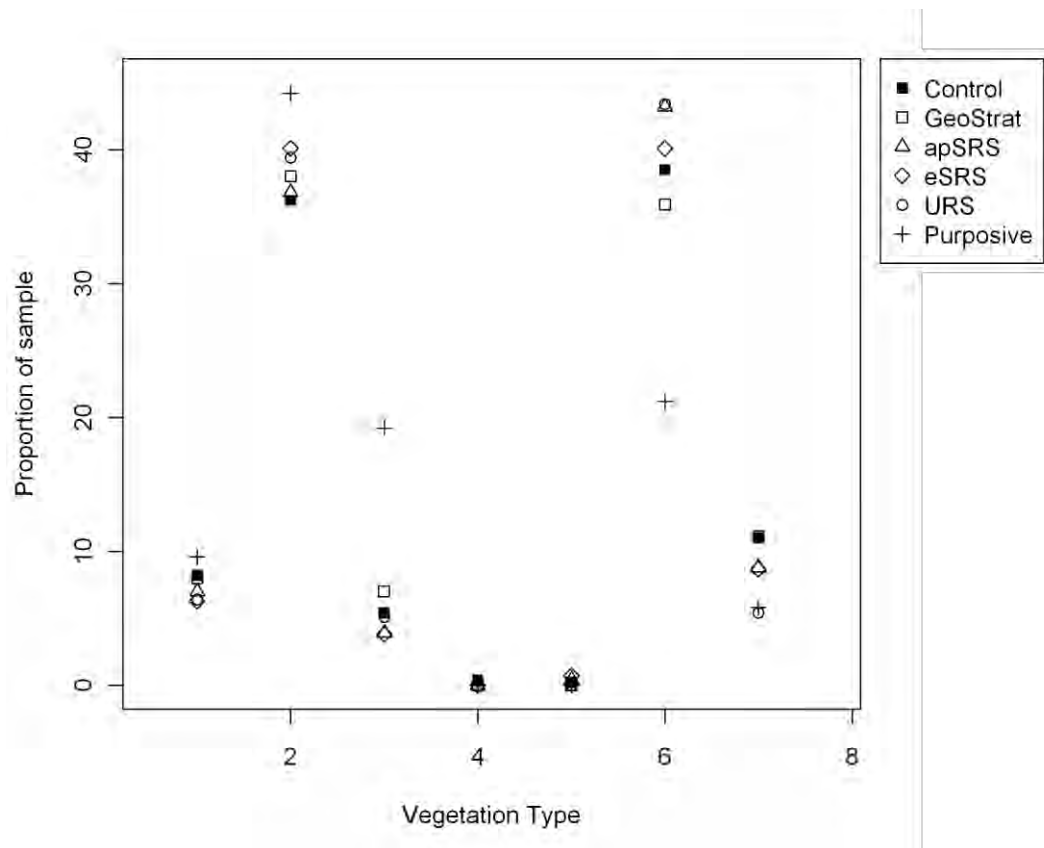
Fig. 6: Proportion of the sample in each vegetation type, as mapped by Selkirk et al. (2000). In the legend, apSRS = area-proportional stratified random sample, eSRS = equal-stratified random sample, and URS = unrestricted random sample. There was little difference among the randomised samples in terms of capturing these broad vegetation types.

Kolmogorov-Smirnov comparisons were used to compare the empirical cumulative distribution of each variable for each sample with the control (Table 5). These showed that equal-stratified random sampling was not significantly different from the control in terms of its distribution against all environmental variables (Table 3). The area-proportional stratified random sample was significantly different from the control when plotted against surface curvature; unrestricted random sampling was significantly different for two variables: elevation and solar radiation; and both GeoStrat and the purposive samples were significantly different from the control for four variables. The *D*-statistics showed that deviations from the control distribution were very small for GeoStrat (0.085-0.128) and large for the purposive sample (0.266 – 0.378).

Table 5: Kolmogorov-Smirnov tests of whether the distributions of the samples against the terrain and image variables were drawn from the same population as the control sample. The D statistic is

the distance between the empirical distribution functions of the control and experimental samples, and the P-value indicates significance (* denotes a significant difference from the control).

	Elevation		NDVI		Curvature		Slope	
	D	P	D	P	D	P	D	P
GeoStrat	0.112	0.002*	0.128	0.000*	0.086	0.037*	0.061	0.270
apSRS	0.054	0.394	0.031	0.952	0.075	0.085*	0.051	0.470
eSRS	0.072	0.111	0.061	0.247	0.071	0.122	0.052	0.428
URS	0.090	0.020*	0.038	0.807	0.050	0.469	0.068	0.147
Purposive	0.378	0.000*	0.352	0.000*	0.094	0.689	0.109	0.498

	Wind Speed		Solar Radiation		Wetness Index	
	D	P	D	P	D	P
GeoStrat	0.045	0.635	0.085	0.040*	0.062	0.239
apSRS	0.046	0.583	0.073	0.101	0.041	0.740
eSRS	0.054	0.396	0.04	0.775	0.067	0.159
URS	0.055	0.358	0.081	0.049*	0.057	0.322
Purposive	0.279	0.000*	0.266	0.001*	0.078	0.874

Empirical cumulative distribution function plots showed that all randomised designs follow similar distribution curves to the control sample (Fig. 7). The purposive sample is biased towards sites that are at low elevations, with moderate solar radiation and exposure to the prevailing winds and high NDVI (i.e. heavily vegetated). The GeoStrat curve lies between the purposive curve and the other randomised curves for elevation, solar radiation, and NDVI, indicating that GeoStrat, like purposive samples is deliberately biased towards the most vegetationally complex parts of the island.

The Kolmogorov-Smirnov tests and empirical cumulative distribution plots show that all the randomised samples are close to the control, while the purposive sample poorly captures the variation in terrain, NDVI, and mapped vegetation types. GeoStrat performs slightly less well than the other randomised samples in capturing the variation in terrain.

Although all the randomised sampling designs perform similarly in terms of capturing the variation in the environmental variables, we believe the bias of GeoStrat in the direction of the purposive sample demonstrates that it is functioning as a compromise between capturing environmental

variation and the perceived variation in vegetation types. The flexibility in setting the proportion of

sites in each stratum also allows researchers to increase the number of sites in areas where practical

considerations are likely to restrict access.
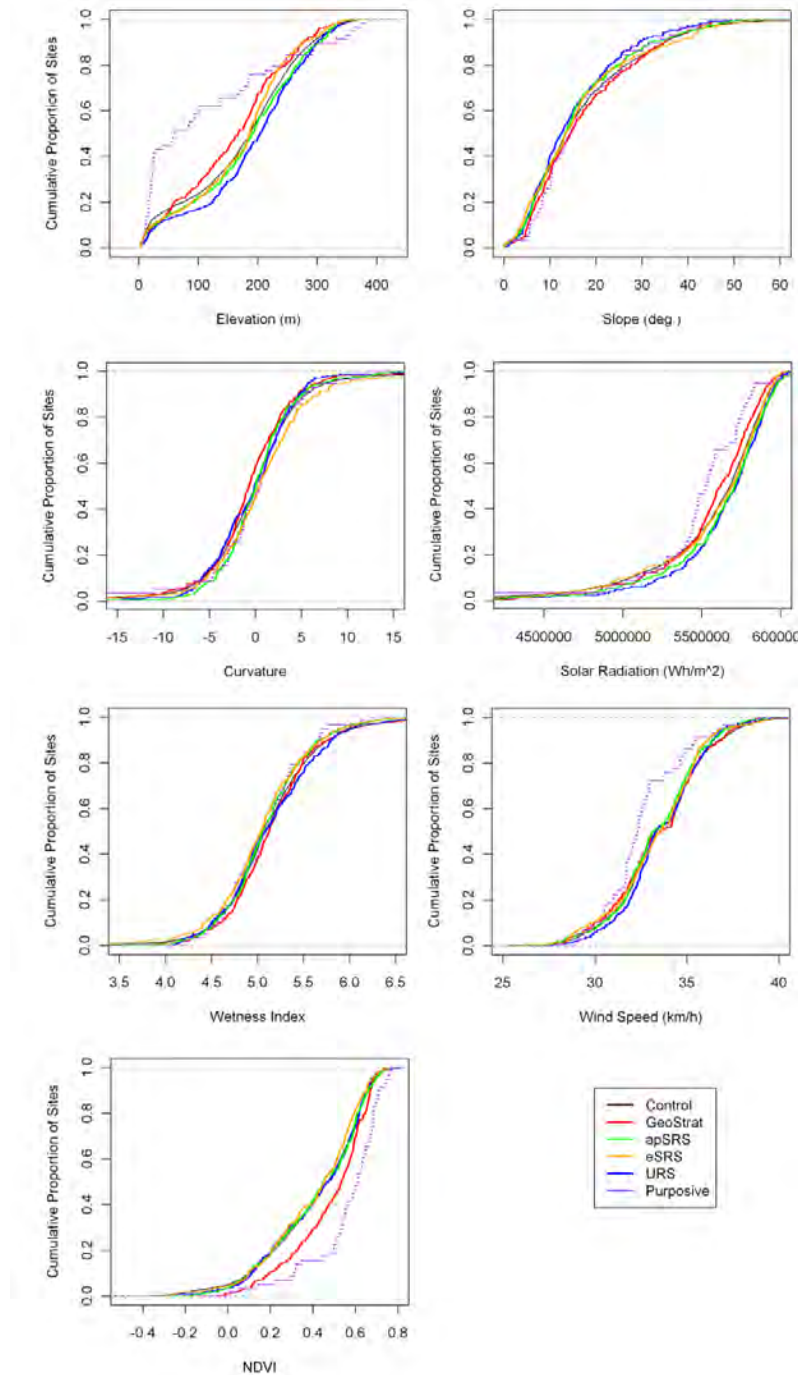
Fig. 7: Empirical cumulative distribution functions for the seven terrain and image variables used in the stratification, for each sampling design. In the legend, apSRS = area-proportional stratified random sample, eSRS = equal-stratified random sample, and URS = unrestricted random sample. These plots show considerable differences between the distribution of the purposive sample (purple line) and the others, including the control. GeoStrat (Red line) shows smaller differences, and lies between the purposive line and the other randomised samples.

3.4.   Discussion

Vegetation surveys are expensive and time-consuming to conduct, and it is important to ensure that the resulting data provides accurate information on the distribution of plants across the terrain. All the randomised sampling designs tested here performed similarly in capturing the true variation in the terrain variables and the broad vegetation structures mapped by Selkirk et al. (2000). The purposive sample was strongly biased towards the coastal areas where the vegetation is complex and away from the extensive but vegetationally simple plateau and mountain tops. Where the GeoStrat sample deviated from the true distributions of the terrain variables, it acted as a compromise between the other randomised samples and the purposive sample.

Landscape stratification for randomised survey design is generally employed to improve the sampling efficiency of randomised sampling, on the assumption that the variables used to stratify the sample are likely to affect the distribution of plant species. Terrain morphometric classification groups similar parts of the landscape in an objective and repeatable way, and is less prone to artefacts of stratification caused by converting continuous variables into discrete classes, then intersecting those classes, as occurs in GRADSECT-style stratifications. Such stratifications are limited in the number of variables that can be incorporated because of the exponential increase in the number of strata as the numbers of variables and classes increase (Guisan and Zimmermann 2000; Austin 2002; Chiarucci 2007). With GeoStrat, the number of continuous input variables that can be incorporated is limited only by data availability and the ecological judgement of the researcher. The number of classes produced is likewise determined by the researcher. In this study, we have demonstrated that GeoStrat can be implemented in remote, physically challenging conditions and that it comes close to capturing the full range of terrain variation whilst also capturing some of the vegetation patterns targeted by purposive sampling.

Using pairwise voting and weighted composite rankings allows us to incorporate more sophisticated measures of the relative "interest" in each stratum than simple area-proportional sampling. In this

case study, we used three separate variables to assign the proportion of sampling sites to each stratum: area; an objective quantitative measure of vegetation variability, NDVI; and a combined measure of perceived ecological variability and accessibility. This methodology in GeoStrat could also be applied to other measures more appropriate to specific vegetation sampling projects.

Most comparisons of the effectiveness of sampling regimes to date have either been based on simulated data (e.g. Hirzel & Guisan, 2002) or conducted in easily-accessible sites (e.g. Goedickemeier et al. 1997) and hence can be considered idealised scenarios. Here, we tested the effect of working in difficult environments on the planned sampling designs. After down-weighting for access limitations, all the randomised sampling designs, including GeoStrat, under-sampled the steepest coastal slopes (s7). Only GeoStrat, however, provided a mechanism to increase the number of samples in this stratum.

Fully testing the success of a vegetation sampling regime requires accurate knowledge of the true distributions of plant species and of the ecological variables, such as terrain or geology, which are under examination. Spatial datasets of terrain variables provide a good approximation of the true variation in these variables, which can be compared to the distribution in the sample (Goedickemeier et al., 1997; Franklin et al., 2001), as has been done here.

The tests presented here investigate the ability of the sampling designs to capture environmental variation as an imperfect proxy for the distribution of plant species. Better testing of the sufficiency and efficiency of sampling designs requires accurate knowledge of the true distribution of plant species in geographic space, as well as knowledge of the true distribution of environmental variables. This is unlikely to be possible in remote areas or for large regions, but such testing may be feasible in small areas and for conspicuous plant species. We recommend such experiments be undertaken to better understand the performance of the various randomised sampling designs. Further, we are unaware of any previous studies comparing the performance of sampling designs in areas where access is limited by safety concerns or the presence of wildlife. As access limitations are

unlikely to be randomly distributed across a study area, we believe it was important to examine the effect of spatially patchy access limitations on the performance of the sampling regimes. The strong performance of equal-stratified random sampling is likely to be an artefact of the coincidence of the hardest to access sites being in the smallest strata, and commensurately strong down-weighting of the sample in this stratum. On Macquarie Island, equal-stratified random sampling functioned similarly to GeoStrat, but in environments where the greatest access limitations occur in larger strata, this relationship is unlikely to hold.

3.5.   Conclusions

This study demonstrated a method for proportionally stratifying a random sample to account for topographic and ecological variation, as well as expected field limitations. GeoStrat uses fuzzy *c*-means classification of the landscape into homogeneous strata, and the proportion of planned sites in each stratum is determined by a composite ranking of three criteria: area, standard deviation in the NDVI, and the likelihood of ecological variation and access limitations. This method enabled the collection of quantitative species composition data for a large, ecologically representative sample in a remote environment. Down-weighting each of the samples on the basis of the access difficulties encountered in the field allowed a fair comparison between sampling designs in a physically challenging environment.

All of the randomised samples here including GeoStrat came close to sampling the full range of terrain variables, with equal-stratified random sampling performing best. GeoStrat, however, allows researchers to use their knowledge of the study area to account for likely field difficulties, and for likely hotspots of biological diversity. The flexibility in the design stage allows for flexibility in the field, and the use of fuzzy c-means clustering of environmental variables overcomes the limitations of existing landscape stratification methods for vegetation survey.

4. Are plant communities discrete mappable entities or fluid groupings? A test case using sub-Antarctic tundra vegetation

4.1.   Abstract

Question: Accurate maps for monitoring vegetation change at landscape scales require mapping

units that can be defined robustly and repeatably. This study examines whether unsupervised

classification of floristic data can produce stable definitions of communities suitable for change

detection mapping at landscape scales. It further examines the effect of various choices in cluster

analysis on the stability of the classification.

Location: Macquarie Island, sub-Antarctic

Methods: Researchers attempting to define communities using statistical clustering methods face

many choices in their analysis, including sampling design, the taxa for inclusion, data transformation,

the number of clusters and the clustering algorithm. Silhouette plots were used to measure the

effect of all these choices on the isolation of the clusters in floristic data from sub-Antarctic

Macquarie Island, an island with relative low species richness, no trees and simple vegetation

structure. The overlap between best-matched pairs of clusters was used as a further test of the

stability of the clusters to minor changes in classification method. Finally, we made qualitative

comparisons of existing community definitions for the island.

Results: No combination of methods resulted in well isolated clusters in the data. The best matched

pairs of sites had a mean overlap of 25.8%, with the most stable clusters containing sites dominated

by a single species.

Conclusions: Macquarie Island vegetation cannot be divided into stable communities for mapping

purposes using any of the methods examined here. While arbitrary communities could be defined,

we consider that maps of individual species are likely to provide fewer definitional uncertainties, and

are hence likely to provide a more robust framework for change detection mapping.

4.2.   Introduction

Vegetation maps are widely used in land management and conservation planning to monitor changes in ecosystems through time, assess impacts of management interventions or monitor the effects of environmental shifts such as climate change (Xie et al. 2008). Although simple in concept, mapping can be problematic as maps used for change detection must represent discrete, unambiguously defined units to ensure that successive surveys can distinguish ecological change from artefacts due to the mapping process (Zimmermann & Kienast 1999; Fortin et al. 2000; Hearn et al. 2011).

There are many ways to map vegetation. At regional scales, maps of plant communities are popular, especially among land managers, applied ecologists and remote sensing analysts (e.g. Vegetation Subcommittee 2002; Xie et al. 2008). However, mapping of plant communities for change detection is a task that depends on a very specific definition of the term "community", i.e. a homogeneous, discrete and recognisable vegetation unit that recurs across the landscape (Austin 1985). More than a century after the concept was first proposed (see Kendeigh 1954), a wide range of methods have been developed for defining plant communities and some dispute about whether they are a suitable entity for mapping purposes. Even where recurring assemblages of plants can be recognised by their most prominent species, closer inspection of most such communities reveals great internal heterogeneity (Whittaker 1951; Austin 2002) and mapped boundaries can often only represent best estimates of boundary zones in areas of uncertainty (Noy-Meir & van der Maarel 1987; Green & Hartley 2000). Under this view, communities form as the result of overlapping resource requirements and ecological tolerances of individual species, rather than as homogeneous entities. Many researchers now adopt an operational definition of plant community along the lines of "the living organisms present within a space-time unit of any magnitude" (Palmer and White 1994). There is general acceptance that communities are fluid entities, which raises questions about the purposes for which categorical maps can be useful.

Concerns have been raised about the scientific validity (Whittaker 1951), discreteness (Cushman et al. 2010), temporal stability (Guisan and Zimmermann 2000), definitional uncertainty (Franklin 1995), and even the reality (Dobrowski et al. 2008) of discrete plant communities. There is some doubt about the accuracy and usefulness of categorical community maps as opposed to individual species maps (Nilsen 1999; Zimmermann & Kienast 1999; Faber-Langendoen et al. 2007; Cushman et al. 2010).

In vegetation classification schemes and mapping projects, there is often limited testing of the assumptions that the communities are persistent and discrete entities, which can be reliably described independently of the partition method used (e.g. Vegetation Subcommittee 2002; Hill and Thomson 2005; Daemane et al. 2010). Most commonly, one or two partitioning methods are used, with TWINSPAN often used in isolation (Coker 2000; Dirnböck et al. 2003; Verrelst et al. 2009). Some researchers have compared phytosociological tables and a statistical clustering method (Grabherr et al. 2003; Černá & Chytrý 2005) or combined ordination with one or two clustering techniques (Bergstrom et al. 2002; Faber-Langendoen et al. 2007; Daemane et al. 2010).

Despite the lack of agreement over the appropriateness of the concept of crisply bounded plant communities for mapping, such maps are still popular. Cushman et al. (2010) suggested that the plant community paradigm persists largely due to a combination of historical legacy and management convenience. Maps of plant communities appear able to present complex ecosystems as a single data layer, are relatively cheap to produce, require less data than multiple maps of individual species distributions, and are easy to interpret (Franklin 1995; Hearn et al. 2011). Franklin (1995) noted that mapping of individual species distributions presents fewer definitional uncertainties than community mapping, and that community mapping is often employed because data limitations make individual species-mapping difficult, rather than because of loyalty to the community concept. Community mapping however, is only appropriate if the resulting maps are useful for their intended purpose. Increasingly, that purpose includes the capacity to predict the

distribution of resources or to accurately monitor changes in vegetation patterns (Cushman et al. 2010; Hearn et al. 2011) and hence requires precise definitions of mapping units.

To investigate the challenges of defining plant communities, we examined the vegetation of sub-Antarctic Macquarie Island. The island's tundra-like vegetation is relatively simple in both species richness and structure yet expected to change significantly in coming years due to a major pest (rabbits, rats and mice) eradication program (Shaw et al. 2011), changing climate conditions (Pendlebury and Barnes-Keoghan 2007) and recently discovered die-off in the endemic cushion plant (Bergstrom & Bricher unpublished data; Department of Primary Industries 2009). Fine resolution, accurate maps of the vegetation on Macquarie Island would thus allow significant change to be defined and monitored, and the impact of management action assessed.

To effectively monitor this change at an island-wide scale, the mapped entities must be objectively and repeatably defined. There have been several previous attempts to define the plant communities of Macquarie Island, using a range of techniques (Taylor 1955; Selkirk et al. 1990; Selkirk et al. 2000; Kirkpatrick and Scott 2002; Bergstrom et al. 2009a). We examine these classifications, and compare them for overlapping consistency. We then test the stability of plant communities on Macquarie Island as defined though statistical clustering of floristic data.

As there is no clear agreement in the literature on the best method for identifying clusters in floristic data, and there are a number of choices that must be made in cluster analysis (including sampling design, the clustering algorithm, the taxa included in the analysis, data transformations, and the number of clusters) we conduct multiple clustering attempts to test the effect of these choices on the stability of the clusters, make recommendations for future vegetation mapping efforts on the island and examine the significance of our findings to vegetation science more broadly.

4.3. Methods

4.3.1. **Study site**

Macquarie Island (54°30'S, 158°57'E) is located approximately equidistant between Australia, New Zealand and the Antarctic continent (Fig. 1). The 12,390 ha island is characterised by low-growing non-woody vegetation, with 44 vascular plant species and approximately 135 bryophyte species along with lichens, algae and fungi (Selkirk et al. 1990). A plateau of 200 m to 400 m elevation dominates the island, surrounded by steep escarpments and a fringe of coastal flats. The most exposed parts of the plateau are characterised by open feldmark, with an extremely variable mix of gravel, cushion plants and mosses, interspersed with small individuals of other vascular plant species. Less exposed parts of the plateau are characterised by short grasslands. Areas where the water table is close to the surface contain bryophyte- and sedge-dominated mires. The coastal slopes contain a complex mix of short grasslands, heavily disturbed or grazed sites and tall tussock grasslands and mega-herbfields (Selkirk et al. 1990).

Macquarie Island has a long history of disturbance by introduced vertebrate species, especially cats, rabbits, black rats, and mice, which have severely impacted the vegetation (Copson 1984). Cats have been removed from the island and a pest eradication program has begun to remove the rabbits and rodents (Parks and Wildlife Service 2007). Fluctuations in the rabbit population have been linked to changes in the distributions of several coastal plant species, most notably the megaherb *Stilbocarpa polaris*, tall tussock grasses *Poa foliosa* and *Poa cookii*, and the fern *Polystichum vestitum* (Copson & Whinam 1998; Bergstrom et al. 2009). The eradication is in an early phase, but there are already signs of coastal vegetation recovery (Shaw et al. 2011). On the plateau, die-off in the endemic cushion plant *Azorella macquariensis,* first noted by us in December 2008, has since been recorded in all areas of the species' range, and resulted in the species being listing as critically endangered under Australian legislation (Department of Primary Industries 2009). Measuring the impact of the eradication program and the cushion plant die-off requires, among other things, quantifying changes

in vegetation patterns. For this, the first step is to test whether plant communities can be defined in an objective and repeatable way.
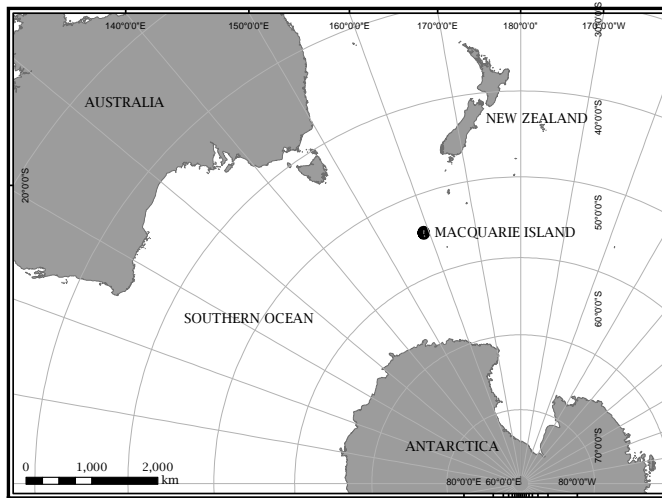
Fig. 1: Macquarie Island (54°30'S, 158°57'E) is located in the Southern Ocean. It is small (12, 390 ha), heavily disturbed, and has a small flora. This results in complex vegetation patterns that may pose challenges for unambiguous mapping of plant communities.
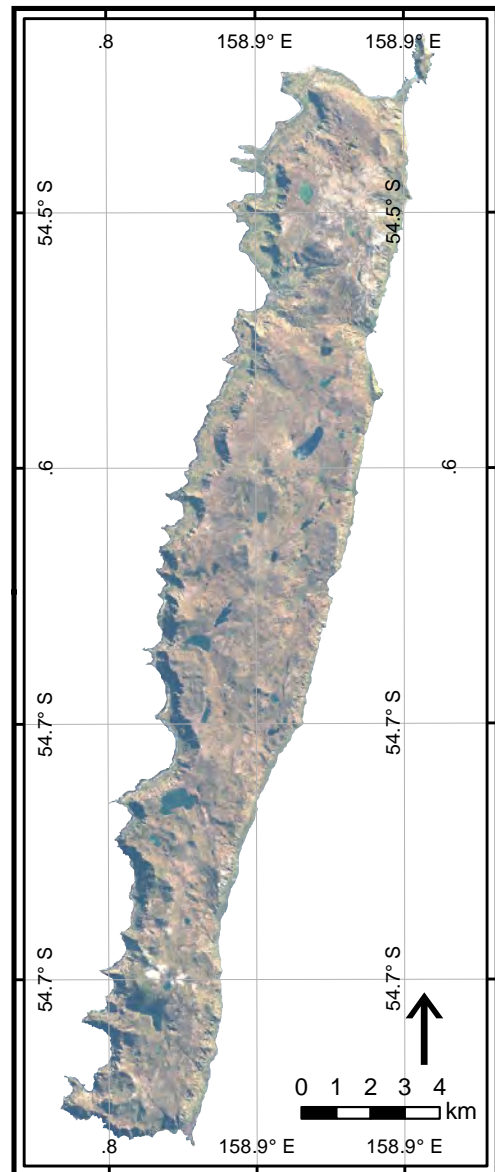
Of the existing community descriptions for Macquarie Island, four covered the entire island (Taylor 1955 ; Selkirk et al. 1990; Selkirk et al. 2000; Bergstrom et al. 2009), while Kirkpatrick and Scott's classification (2002) was restricted to undisturbed tall vegetation on the coastal slopes. Taylor (1955) and Selkirk et al. (1990) attempted to subjectively describe the main communities on the basis of the dominant species and environmental variables. Selkirk et al. (2000) mapped vegetation structure by visually classifying the foliage density and canopy height observable in a SPOT 3 satellite image

(panchromatic resolution: 10 m; multispectral resolution: 20 m). Bergstrom et al. (2009) used

multivariate clustering to explore changes in vegetation cover at 18 sites between 2001 and 2007.

The sites for that study were purposively chosen in 2001 to represent a range of subjectively defined

plant communities. Kirkpatrick and Scott (2002) used TWINSPAN to divide 30 undisturbed sites on

coastal slopes into four groups on the basis of their floristic composition.

We qualitatively compared the published descriptions of the community descriptions for the tall

coastal vegetation, by drawing links among described communities with at least partial overlap.

4.3.2. **GeoStrat**

To choose the sampling sites, we developed GeoStrat, a geographically-explicit proportional

stratified random sampling design (see chapter 3 for details). Six terrain variables (elevation, slope,

surface curvature, topographic wetness index, solar radiation and prevailing wind speed) were

calculated from a 5 m resolution digital elevation model (DEM). A normalised difference vegetation

index (NDVI) was calculated from a cloud-free QuickBird satellite image that was acquired on 15

March 2005. These seven spatial data layers were classified using fuzzy c-means clustering to

generate a map of eight landscape strata. The proportion of random points in each stratum was

allocated according to rankings of the surface area, standard deviation of NDVI, and likelihood of

complex vegetation and access problems for each stratum. Accounting for accessibility in the design

inflated the number of potential sites in the vegetationally complex coastal slopes, where access was

limited due to safety and conservation concerns. Stratification for vegetation sampling assumes that

the strata correspond with actual vegetation types.

During the summers of 2008/09 and 2009/10, we visited 349 sites. From the sites chosen through

GeoStrat, 43 sites were discarded and a further 26 planned sites were relocated by up to 30 m into

similar vegetation due to inaccessibility. The sample was supplemented by 54 purposively located

sites used for other ecological research projects that were deemed to be representative of perceived

plant communities (Bergstrom et al. 2009; Bergstrom, unpub. data), and eight sites were purposively

located to sample very rare vegetation types (e.g. the formerly common *Polystichum vestitum* fernbrakes or *Stilbocarpa polaris* dominated herbfields), as recommended by Roleček et al. (2007).

4.3.3. **Field Methods**

At each site, a 10 x 10 m plot was marked out, and the location of the south-east (magnetic) corner of the plot was recorded using a Trimble ProXH single-frequency differential GPS receiver (horizontal precision mean = 0.22 m, SD = 0.3 m). The size of the field plots was chosen to match four multispectral pixels in QuickBird satellite imagery, so that the plots could be used to train image classifications. An 8.75 m^2 vertical photograph was taken in each corner of the plot, covering a total area of 35 m^2 or 1/3 of the plot. The photographs were taken from a height of approximately 2.5 m using a Nikon D80 digital camera with a 14 mm wide-angle lens mounted on a dipod (Fig. 2). A species list was recorded for each site.

(a)

(b)
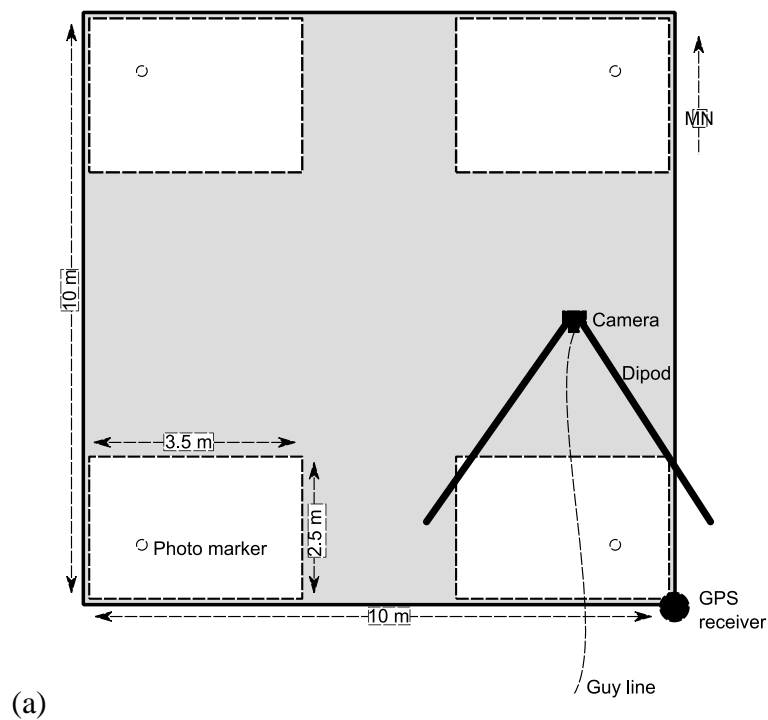
Fig. 2: (a) Site layout, showing the location of the GPS logger in the south-east corner, and the photography equipment over the north-west quarter of the plot. A guy-line pegged to the ground or held by a field operator is used to stabilise the dipod. (b) A field operator photographs the south-west corner of a feldmark site.

4.3.4. **Photographic interpretation**

Cover values for each study plot were estimated using point intercept photo-interpretation. One hundred points were randomly placed over each photograph using Coral Point Count software (Kohler and Gill 2006), and the ground cover class beneath each point identified. For each site, the results from the four photographs were summed. Points were excluded where the ground cover class could not be confidently identified or was obscured by research equipment (e.g. photo markers and tape).

Point intercept interpretation methods frequently underestimate the species richness of a site, as rare or understorey species have a reduced chance of being sampled (Foster et al. 1991). To avoid this, we added a token 0.5% cover for all taxa that the field species list recorded as present, but that were not represented in the point intercept data. We tested the effect of this addition in the statistical analysis (as recommended in chapter 2).

Most vascular plants were identified to species level, and non-vascular taxa were identified to broader levels of biological organisation (e.g. algae and fungi were treated as single taxa). Those

vascular plant species that could not be reliably identified to species level in photographs (*Deschampsia spp.*, *Acaena spp.*, *Epilobium* spp., and two species of *Colobanthus*, *C. apetalus* and *C. affinis*) were identified to genus level. Of these genera, only the two *Acaena* spp. have different ecological ranges, with *Acaena minor* restricted to high altitude sites, while *Acaena magellanica* occurs in all environments, including most of the range of *A. minor*. We do not believe that identifying these genera to genus level would significantly cloud floristic classifications.

Two non-vegetated ground cover classes were also quantified – water and bare ground/dead vegetation.

4.3.5. **Statistical Analysis**

Every analysis of plant community data involves a number of choices. These include which taxa to include, how strongly the data are transformed, which clustering algorithms to use, and the optimal number of clusters. We tested the effects of these choices by repeating the clustering analysis with different inputs. By definition, clustering algorithms find groups in data regardless of the presence of clear structures in the dataset. Cluster stability provides an objective criterion to test whether the clusters are artefacts rather than true groups (Pillar 1999). Silhouette plots were used to measure the isolation of the clusters in each clustering attempt (Maechler et al. 2002).

4.3.5.1. **Data subsets**

We tested the sensitivity of the classification results to the inclusion of different subsets of the data. First, we examined the sensitivity of the clustering to the sampling design by comparing randomly and purposively sampled sites. We then examined the effect of including two groups of taxa in the analysis: the broad taxa used for non-vascular plants (e.g. bryophytes and lichen); and very small, widely dispersed vascular plant species. The broad non-vascular taxa often make up a large proportion of the vegetation at a site, but these taxa include ecologically diverse species, and may thus blur the boundaries between communities. There are a number of very small vascular plant species on Macquarie Island, which occur as isolated individuals in a wide range of habitats, and

which are easily overlooked. To minimise the possibility that our results were an artefact of selection

criteria, we conducted partitioning around medoids (PAM) clustering on multiple subsets of the data

(Table 1) and tested the isolation of the resulting clusters (Kaufman and Rousseeuw 2005).

Table 1: The data subsets used for the cluster analysis, including subsets of sites and of taxa. Cluster stability was tested for each of these subsets.

	Subset	Description
1	All sites, all taxa	Data from all sites. All taxa recorded in the point intercept method.
2	All sites, few interstitials	As in subset 1, but with those taxa that were extremely rare or difficult to identify in photographs removed. These taxa are: *Deschampsia* spp., *Puccinellia macquariensis*, *Isolepis aucklandica*, *Cardamine corymbosa*, *Cerastium fontanum*, *Colobanthus apetalus/affinis*, *Coprosma perpusilla*, *Crassula moschata*, *Nematoceros* spp., *Stellaria media*, *Stellaria parviflora*, *Grammitis poepiggiana* and *Huperzia australiana*.
3	All sites, no interstitials	As in subset 2, but with additional taxa excluded. These taxa generally occur as small, isolated individuals that may be overlooked in a photograph. Their inclusion in the analysis has the potential to add noise. This subset was the strictest in terms of taxon-exclusion. The additional excluded ground cover classes were: *Callitriche antarctica*, *Epilobium* spp., *Montia fontana*, *Ranunculus crassipes* and bare ground.
4	All sites, all taxa, with field observations	As in subset 1, but with 0.5% cover listed for all taxa recorded in the field species list that were not recorded in the point intercept observations.
5	Random sites, all taxa, with field observations	As in subset 4, but only data from randomly located sites were included.
6	Purposive sites, all taxa, with field observations	As in subset 4, but only data from purposively located sites were included.
7	All sites, no broader taxa, with field observations	As in subset 4, but a number of taxa were excluded. Those species that do not occur in more than one site in the dataset (*Anthoxanthan odoratum*, *Poa cookii*, *Poa litorosa*, *Carex trifida*, *Galium antarcticum*, *Myriophyllum triphyllum*, *Rumex crispus*, *Stellaria parviflora*, *Hymenophyllum falklandicum*, *Huperzia australiana*) were removed. In addition, those broad taxa that include ecologically diverse species (fungi, algae, lichens, bryophytes) were excluded to minimise the potential for the creation of clusters that represent biologically/ecologically diverse vegetation types.

4.3.5.2.   Data transformations

The strength of the data transformation in ordination and clustering analyses affects the degree to which common and rare species dominate the clustering. We analysed the data using the Bray-Curtis distance metric applied to three transformations: untransformed, square-root transformed, and Wisconsin transformed data. Stronger data transformations were not used, as they have been shown to produce poor clustering results in similar data collected on Macquarie Island (see chapter 2 for details).

4.3.5.3.   Ordination and clustering

Non-metric multidimensional scaling (nMDS) was used to explore the distribution of the sites after each transformation. Two- and three-dimensional nMDS were calculated using the metaMDS tool in the R package *Vegan* (Oksanen, 2006).

The data were clustered using two clustering algorithms. PAM clustering is a form of fuzzy *c*-means clustering that allows the use of non-Euclidean distance metrics, including Bray-Curtis distance (Maechler et al. 2002). Hclust is an agglomerative hierarchical clustering algorithm, and was applied with an average weighted linkage method (R Core Development Team 2010)

One measure of cluster stability is the isolation between clusters, as measured using silhouette plots (Maechler et al. 2002). Silhouette widths are calculated by equation [1]. The mean silhouette width (MSW) was used to validate single clustering results as well as to compare the results of multiple clustering efforts.

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \qquad\qquad [1]$$

Where, for *i*, a site belonging to cluster *a*, $a(i)$ = average Bray-Curtis dissimilarity of site *i* to all other sites in cluster A and $b(i)$ = average Bray-Curtis dissimilarity of *i* to all sites in the nearest neighbouring cluster. Fielding (2007) suggests that an MSW of 0.7 indicates strongly structured clusters, while a value around 0.5 suggests a weak clustering structure, and a score below 0.25 indicates no evidence of clustering.

Previous classifications of the vegetation on Macquarie Island have defined between four and 35 communities, so there was no *a priori* reason to choose any particular number of clusters. To guide this decision, we applied PAM to all combinations of data subsets and transformations. MSWs were used to choose an optimal number of clusters. The combinations of data subsets, transformations, number of clusters and clustering algorithms tested are summarised in Table 2.

Table 2: Summary of the analysis methods tested in this study. All combinations of data subset, transformation and cluster numbers were assessed for stability using the mean silhouette width of the clusters generated using PAM. Once the cluster number had been set at eight, the choice of clustering algorithm was tested for all data subsets and transformations.

Data subset		Transformation	No. Clusters	Clustering Algorithm
1	All sites, all taxa	Untransformed	2	Partition Around Medoids (PAM)
			3	
2	All sites, few interstitials		4	
			5	
3	All sites, no interstitials	Square Root	...	
4	All sites, all taxa, with field observations			
5	Random sites, all taxa, with field observations			
6	Purposive sites, all taxa, with field observations	Wisconsin	46	Hierarchical Clustering
			47	
			48	
7	All sites, no broader taxa, with field observations		49	
			50	

4.3.5.4. Measures of overlap

To further explore cluster stability, we paired up the best-matching individual clusters from each pair

of clustering attempts. For each cluster-pair, the percentage of sites that was common to both

clusters was calculated. This measure of site overlap among clusters was calculated by equation [2]

and is analogous to the Jaccard coefficient for species overlap among sites.

$$J = \frac{A \cap B}{B \cup B} * 100 \qquad \text{for sites } A \text{ and } B. \qquad [2]$$

4.4. Results

4.4.1. Existing community definitions

The qualitative comparisons of the existing community descriptions by other researchers showed

that overlaps among the various descriptions were only partial (Fig. 3) (Taylor 1955; Selkirk et al.

1990; Selkirk et al. 2000; Kirkpatrick & Scott 2002; Bergstrom et al. 2009). No pair of research teams

defined a single community in the tall coastal vegetation the same way, which illustrates the

complexity of even this vegetation on a small island, and the subjectivity involved in deciding where

the boundaries among communities should lie. The structural classes defined for mapping purposes

by Selkirk et al. (2000) are not connected to the community definitions of other authors in Fig. 3,

because each of these classes contain a complex mosaic of almost all the communities as defined by

the other authors.
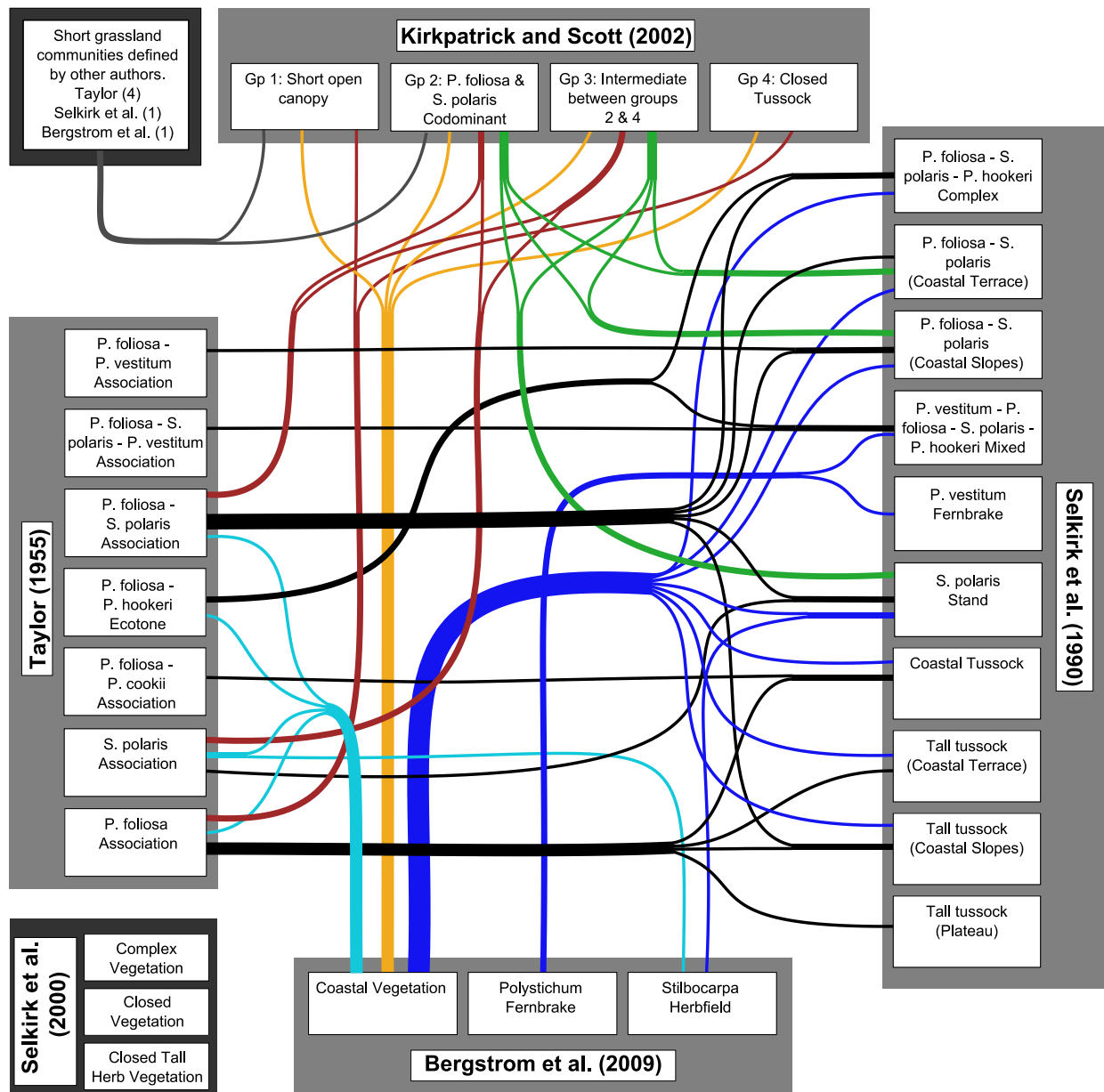
Fig. 3: Relationships between the definitions of plant communities described by previous authors. The structural classes defined by Selkirk et al. (2000) are not connected to any other communities because they all represent complex mosaics of almost all the communities defined by the other researchers.

4.4.1.1. Clustering results

This study examined the impact of multiple choices on the clustering of vegetation data, and no combination of choices resulted in stable groupings of sites. Using the MSW as the measure of stability, all clustering attempts indicated there was either no or very weak clustering structure in the data (Fielding 2007).

In choosing an appropriate data transformation, data subset and number of clusters, the MSW was low for all clustering attempts (mean = 0.18, SD = 0.05) (Fig. 4). The highest MSW of 0.32 was found for 18 to 21 clusters applied to untransformed data from the purposive sites only. There was thus no clear evidence for an optimal number of clusters, and all clustering efforts produced very low MSW.
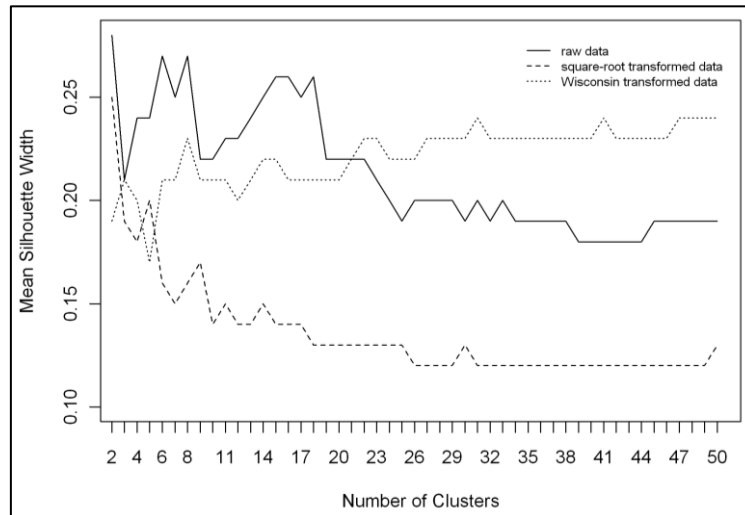
Fig. 4: Mean silhouette widths of PAM clusters of data subset 7 with varying numbers of clusters. Similar results were obtained with all other data subsets. A mean silhouette width below 0.25 indicates that there is little evidence of cluster structure in the data, while a value above 0.5 is needed to indicate a reasonable cluster structure.

For most subsets of the data, however, there was a small peak in MSW around eight clusters, with a slightly smaller peak around 14-17 clusters. All further analyses were therefore conducted with eight clusters. The choice of data subset made little difference to the MSW (range 0.13-0.3, mean = 0.21). The increase in the MSW for purposive sites (0.3) was minor relative to randomly-sampled sites (0.26), so there was no evidence for stable cluster structures in the dataset even when targeting sites to sample homogeneous areas of perceived plant communities.

We performed all further analyses on subset seven. This subset included data from all sites visited in the 2008/09 or 2009/10 field seasons. It included a token 0.5% cover for all species observed in the field but that were missed by the point intercept photo analysis (as recommended in chapter 2) and excluded the broad non-vascular taxa. This data subset theoretically should produce the best

opportunity for clear clusters, but it produced similar results to those obtained using other subsets

of the data.

4.4.1.2.   Comparing clustering algorithms

The two clustering algorithms were applied to data subjected to the three transformations

(untransformed, square-root and Wisconsin). Due to the voluminous nature of the analysis, plots are

presented here only for the untransformed data, which produced the strongest clustering results.

PAM clustering resulted in more equal-sized clusters than hclust. Hclust produced two very large

clusters and a number of very small clusters (largest cluster: 178 sites) (Fig. 5a), and an MSW of 0.3

(Fig. 5b). The clusters produced by PAM were more equal in size (smallest cluster: 13 sites; largest

cluster: 78 sites). The MSW for the PAM clusters was low (0.27) and five of the eight clusters

produced by PAM on untransformed data contain sites that are closer to the medoids of their

nearest neighbouring clusters (Fig. 5c) than to the medoid of their assigned cluster. This suggests

there was little or no cluster structure in the dataset (Fielding 2007)*.*
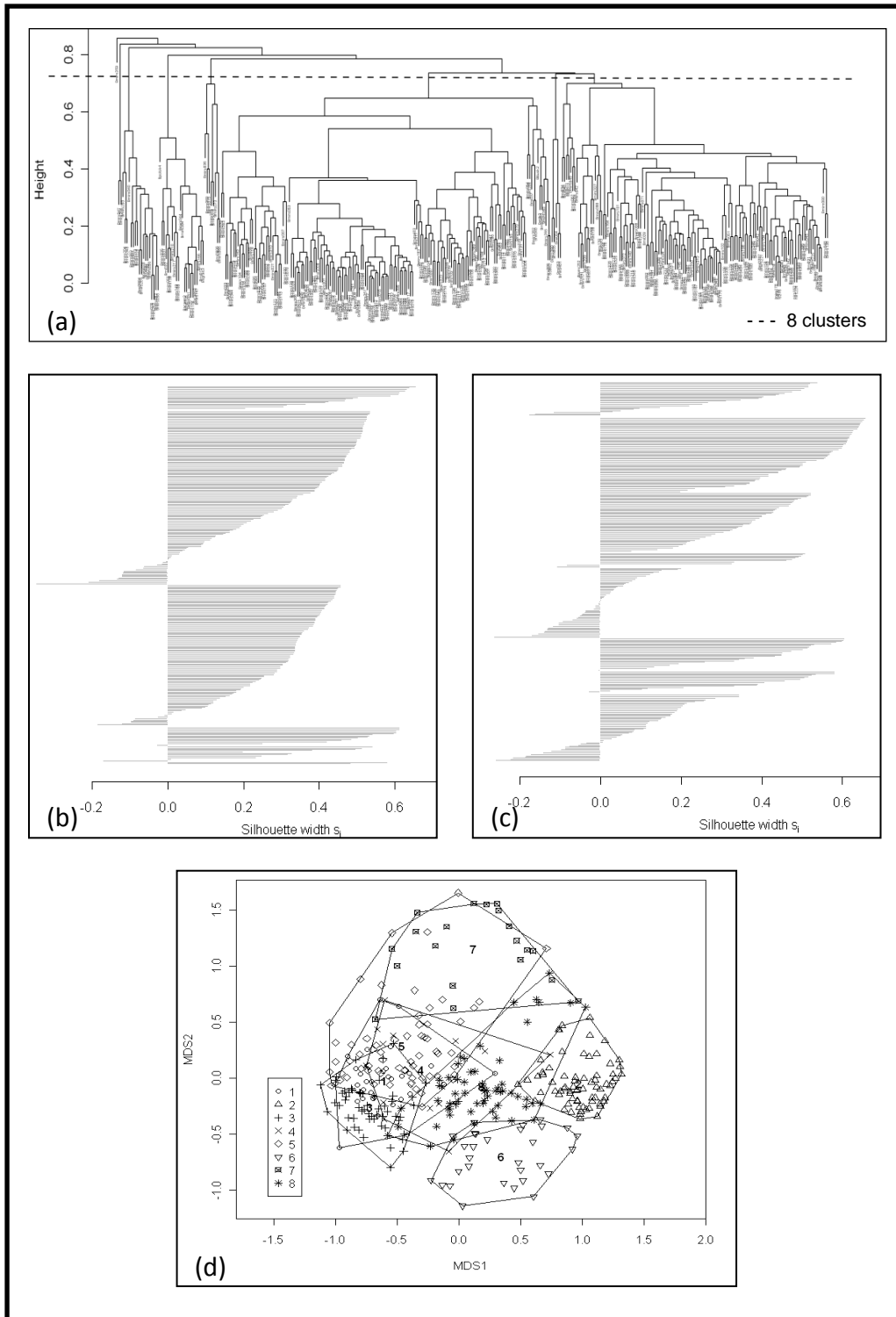
Fig. 5: Clustering of untransformed data in subset 7 into eight clusters. (a) The hclust dendrogram, with cluster sizes ranging from 1 site to 178 sites. (b) The silhouette plot for the hclust, with an MSW of 0.3. (c) The silhouette plot of the PAM clustering, with an MSW of 0.27. (d) The first two dimensions of a 3-dimensional nMDS, showing highly overlapping convex hulls around the eight clusters found using PAM (Stress = 0.14). Clusters 2 and 6, the most clearly separated clusters in the nMDS plot contain feldmark sites, which are characterised by extensive bare ground and the cushion plant, Azorella macquariensis.

101

The clusters identified by the two clustering algorithms for the three data transformations had only weak overlaps (Table 3). The overlaps for best-matched pairs of clusters ranged from 0.6% to 48.6% with a mean overlap of 25.8%. That is, almost three-quarters of the sites that occurred in either of a best-matched pair of clusters did not occur in both. Thus, changes in data transformation or clustering algorithm had large effects on the partitioning. This suggests that any grouping masks great internal heterogeneity and the boundaries are very fuzzy (Pillar 1999).

Table 3: Mean percentage (%) overlap of sites between the best pairs of clusters in all six clustering attempts. This demonstrates that changing either the data transformation or cluster algorithm resulted in a mean of at least 65% of sites changing clusters.

Cluster Set	hclust (wis)	hclust (sqrt)	hclust (untrans)	PAM (wis)	PAM (sqrt)
PAM (untrans)	22.0	25.1	27.3	23.5	34.7
PAM (sqrt)	25.8	25.7	26.1	25.9	
PAM (wis)	21.3	21.2	20.9		
hclust (untrans)	28.7	31.2			
hclust (sqrt)	27.6				

The few pairs of clusters with reasonable overlaps among the various clustering attempts contained sites with very high coverage of a single taxon. For example, cluster 7 from PAM (untransformed) had a relatively high overlap with a cluster from each of the other clustering attempts (mean overlap 41.2%). Eighteen sites occurred in at least three of the six matching clusters, out of a total of 33 sites identified by at least one of those clusters. Those 18 sites had more tussock grass *Poa foliosa* than all other sites in the dataset. The other group of matching clusters with overlaps above 40% (PAM (untransformed) cluster 1 and its closest matches) mostly comprises sites with extensive mats of *Acaena spp*. In this group, 29 sites occur in at least three clusters of the six best matching clusters, out of a total of 42 sites identified by at least one of these clusters. These 29 sites are among the 34 sites with the highest percentage cover of *Acaena spp.*, and all have at least 15% cover. *Acaena spp.* occurs in a wide range of habitats, but mostly occurs as isolated small individuals. It forms extensive mats in some lower altitude sites, and it is these sites that dominate the clusters in this group. This

suggests that the most stable clusters in each clustering attempt were dominated by a single species with a restricted ecological range. All other matched groups of clusters had less overlap, and their relationships to individual species were harder to discern.

Plotting the results of the clustering attempts with non-metric multidimensional scaling (nMDS) further demonstrated the lack of clear separation among many clusters. Three-dimensional nMDS plots for all data transformations, subsets and clustering algorithms showed that most clusters were not clearly separated in ordination space, although feldmark plots with large amounts of bare ground, cushion plants, and bryophytes (clusters 2 and 6 in Fig. 5d) showed reasonable separation from the other clusters.

4.5. Discussion

The success of a cluster analysis lies in the stability of the clusters that are produced and the robustness of the clustering to the choices that are made in the analysis. While it was possible to generate clusters in the data collected here, the clusters were not well isolated. The small MSW for all number of clusters, sampling methods, included taxa, data transformations, and clustering algorithms showed that clustering was unstable and sensitive to any methodological changes. The low overlap of sites between matched pairs of clusters provides further evidence of instability, with few exceptions. The most stable clusters were those in which the sites with highest coverage of a single dominant taxon. If the only stable clusters are those dominated by a single taxon, it is likely that mapping individual species distributions will be more useful for monitoring changes in vegetation patterns than mapping communities, because the entities being mapped will be so much less ambiguous (see Franklin 1995). One limitation of the approach used here is that bryophytes were excluded from the analysis. Much of the ecological variation on Macquarie Island occurs within the bryophytes, which were excluded on pragmatic grounds in this study due to challenges of identifying bryophyte species in photo-plots.

A wide variety of statistical methods has been developed for defining plant communities, in addition to traditional phytosociological techniques such as comparison tables. We tested a number of the options involved in unsupervised classification, in addition to comparing the results of previous attempts to define the communities on the island (Taylor 1955; Selkirk et al. 1990; Selkirk et al. 2000; Kirkpatrick and Scott 2002; Bergstrom et al. 2009). If the vegetation of Macquarie Island does form homogeneous, discrete, and recognisable units (*sensu* Austin 1985) then it could also be expected that attempts by different research teams to define those communities would reach a reasonable level of agreement, at least for those research teams working at similar times. As it is, no pair of research teams produced identical definitions of a single plant community in the tall coastal vegetation. It is impossible to know the degree to which the differences among the community definitions are due to the variations in scope, techniques used to define the communities, and subjective judgements of the researchers, as opposed to genuine temporal changes in the vegetation patterns between the dates of the classifications. Copson (1984) noted that the vegetation alliances and associations described by Taylor (1955) had become increasingly difficult to identify in the intervening 30 years. The community definitions were created up to 44 years apart, and the vegetation in coastal areas has changed especially rapidly in the past decade due to the large increase in rabbit numbers after the removal of cats and reduction in the *Myxoma* control program (Bergstrom et al. 2009a&b; Dowding et al. 2009). The observation that the combinations of species are shifting through time adds weight to concerns that community mapping may not be most useful for tracking changes through time.

The basic nature of plant communities has been contested since the concept was first proposed (Kendeigh 1954). Many ecologists now use a strictly operational definition of communities and hence avoid the question of whether the communities are discrete, homogeneous and bounded, but applied ecologists and land managers still frequently use a categorical patch model to describe vegetation patterns, although questions remain about its appropriateness (Palmer & White 1994). Cushman et al. (2010) found large discrepancies among independently produced community maps,

and between the maps and field observations of individual species distributions, for forest inventory in the USA. Similarly, Hearn et al. (2011) found low levels of repeatability in the application of Britain's National Vegetation Classification scheme to a small area by experienced surveyors, though the poor results in their study may have been exacerbated by the application of a national scheme to a relatively small area. Not all studies have been so negative (e.g. Stevens et al. 2004) and it is unclear what drives the variation in repeatability among studies. We therefore believe that testing the stability and repeatability of community definitions is an important preliminary step in vegetation mapping exercises. Because the question of whether plant communities can be sensibly defined as having discrete boundaries for mapping purposes is so contested, we believe it was important to test the choices made in any attempt to classify data. We are not aware of any vegetation classification studies that have explicitly tested the effect of so many of those choices.

The use of a large, geographically stratified proportional random sample, supplemented with purposive samples for core and rare vegetation types increased our confidence that the data were statistically representative and captured the full range of environmental and floristic variation (Goedickemeier et al. 1997; Franklin et al. 2001; Lájer 2007; Roleček et al. 2007). Few researchers have collected real-world data using both purposive and stratified random sampling designs to test for the effect of sampling design on the outcome of ecological research. Goedickemeier et al. (1997) tested the effect of sampling design on a biodiversity assessment, with mixed results. Hirzel and Guisan (2002) used simulated data of a hypothetical species distribution to test the effect of sampling designs on distribution predictions, and again obtained mixed results. Neither of these studies, however, looked at plant community descriptions or tested the purposive sampling that is so common in phytosociological studies (Lájer 2007). One might expect the data from the purposively located sites to form more stable clusters than data from randomly located sites, which include sites located in ecotones. Clustering just the purposive sites, however, resulted in only a small increase in the cluster stability.

 Unstable partitions in the data indicate that the data has a fuzzy cluster structure at the given level of clustering being examined (Pillar 1999). By this measure, the vegetation of Macquarie Island is fuzzy at all levels of groupings, and hence any attempt to delineate boundaries in geographic space will involve subjective and uncertain assessments about where to place boundaries and how to define the communities. As Fortin et al (2000) noted, identifying the dynamics of ecotones is not possible if they are located subjectively by researchers.

If the geographically-stratified random sampling design is representative of the topographic variation on Macquarie Island and the vegetation could be divided along discrete boundaries in geographic space, then we would have expected to see relatively stable groupings in attribute space. Instead, these results indicate that if plant communities can be said to exist on the island, it is only as fluid groupings that result from overlapping resource requirements, ecological tolerances, historical disturbance, and interaction among species that respond to environmental gradients individualistically (Zimmermann and Kienast 1999; Cushman et al. 2010). Such a definition sees communities as fluid entities rather than stable groupings suitable for change detection maps. Treating fluid communities as hard classes for categorical mapping may be appropriate for understanding general patterns in the vegetation, so long as both the map creators and its users understand and accept the necessary level of generalisation (Palmer and White 1994).

There are a number of features of the Macquarie Island flora that may complicate problems with the plant community paradigm. Due to its isolation, the island contains relatively few vascular plant species and many of these are ecological generalists. A similar situation has made it difficult for researchers to agree on community definitions for the arctic Svalbard archipelago, though Nilsen et al. (1999) were able to define mappable communities with reasonable accuracy. Macquarie Island has a relatively small extent and is less topographically diverse than the study areas in many such classification studies, which may also contribute to the lack of discontinuities in the floristic composition of sites. This may partially explain why classification appears to be more successful over

larger, more diverse areas, such as the Austrian alps (Grabherr et al. 2003). Macquarie Island is also a heavily disturbed environment and this disturbance has interrupted succession processes (Selkirk et al. 1990). The temporal instability of plant communities adds an additional level of uncertainty that is not addressed in the present study (see Palmer and White 1994).

4.6. Conclusions

Despite testing a wide range of assumptions, including sampling design, choice of taxa for inclusion in the analysis, data transformations and clustering algorithms we were unable to find stable groupings in a large floristic dataset for Macquarie Island. We tested a reasonable subset of the methods commonly used to partition floristic data into plant communities, and the lack of stability in our results corresponds with the lack of agreement among other researchers on the definitions of plant communities on Macquarie Island.

Given that the aim of this study is to develop methods for change detection over time, we consider it important to have tested the assumption that plant communities can be delineated unambiguously before any vegetation mapping was conducted. This paper does not examine temporal changes, but the ability to clearly identify homogeneous and repeated vegetation units at a single time-step is a necessary first step. Other researchers have warned of the risk that real ecological changes may be confounded by artefacts of the mapping process (Cushman et al. 2010; Hearn et al. 2011).

Although mapping the distributions of individual species may not capture the complexity of an ecosystem, this approach has the advantage of minimising the ambiguity in defining the entities to be mapped. The less ambiguity there is in the definition of the mapping units, the greater the chance that apparent change is real, rather than an artefact of the mapping process. For change detection on a small, heavily disturbed island with low species-richness, like Macquarie Island, single species mapping is likely to be more informative than community mapping.

5. Mapping sub-Antarctic cushion plants: using random forests to combine very high resolution satellite imagery and terrain modelling

5.1.   Abstract

Detecting changes in the distribution and density of plant species requires accurate and high-resolution baseline maps of those species. The need for accurate and objective mapping methods has led to the development of many novel techniques in the fields of species distribution modelling and satellite image classification. Within the field of satellite image classification, there are major divisions between object-based and pixel-based training methods and between spectral classification and species distribution modelling. In this study, random forest classification was used to explore the effectiveness of these approaches on maps of the distribution of the critically endangered cushion plant *Azorella macquariensis* on sub-Antarctic Macquarie Island. These models were created using pixel- and object-based classifications using terrain variables, spectral data from very high resolution WorldView-2 satellite imagery, and a combination of both (hybrid). The hybrid classifications produced higher accuracies than either the terrain or spectral variables in isolation, with classification accuracies of up to 90.9%. The multiple pixel based classifications had the highest accuracies, and there was little difference in the performance of the object and single pixel based classifications, perhaps because of the apparently patchy nature of *A. macquariensis* distributions at all spatial scales. Highly accurate maps of absent, sparse and moderate *Azorella* cover indicate that these maps provide a suitable baseline for monitoring expected change in the distribution of the cushion plants.

5.2.   Introduction

There is increasing interest in monitoring landscape-scale changes in the distributions of plants due to a range of impacts, including climate change, species invasions, and management actions (Wallentin et al. 2008; Stohlgren et al. 2010). Monitoring changes in the distribution of individual species or communities is dependent on the creation of accurate high-resolution maps. Such maps can be used to monitor responses to environmental changes at regional or landscape scales, and hence complement plot-level studies. The production of these maps is time-consuming and expensive, and extensive research has been directed at improving mapping methods (Lu & Weng 2007; Xie et al. 2008). Traditionally, species distribution maps have been produced by manually drawing boundaries on the basis of field surveys or photo-interpretation. In addition to the cost, manual mapping involves subjective judgements, and the degree of repeatability may hence be low (Cherrill & Mcclean 1999; Hearn et al. 2011; but see Stevens et al. 2004 for an example of high repeatability).

The field of remote sensing has produced a wide range of semi-automated image interpretation methods to improve the repeatability and objectivity of the resulting maps (Xie et al. 2008). As part of efforts to improve accuracy, mapping technologies are expanding rapidly on a number of fronts, including the development of object-based image analysis techniques (Blaschke 2010); the merging of terrain and spectral data in image classifications (e.g. Dirnböck et al. 2003); the use of texture measures to provide contextual information (e.g. Mallinis 2008; Dobrowski et al. 2008; Grossmann et al. 2010); and more sophisticated classification algorithms such as random forests (RF) (Breiman 2001; Pal 2005; Cutler et al. 2007).

There have been two key technology-driven approaches to investigating species distributions: species distribution modelling (SDM) and satellite image classification. SDM uses available terrain and other environmental data to derive maps of potential habitat (Guisan and Zimmermann 2000). SDMs rarely incorporate temporal disturbances, such as landslips or fires, and are hence limited to

defining the potential distribution of a species, rather than its actual distribution at a given point in time (Dobrowski et al. 2008). In contrast, satellite image classification produces thematic classes on the basis of spectral values (e.g. Laliberte et al. 2004; Lawrence et al. 2006; Mallinis 2008). This general approach includes a wide range of methods, including pixel-based and object-based image analyses and both supervised and unsupervised classification. Such classification may be based on raw spectral data (Laliberte et al. 2004), include indices created from combinations of the available spectral bands (Dobrowski et al. 2008; Hüttich et al. 2011), or include texture measures calculated from one or more spectral bands (Grossmann et al. 2010). These approaches have proved especially useful for classifying images on the basis of the vegetation structure and for identifying disturbance patterns, but to be less successful at identifying individual species or vegetation types (Dirnböck et al. 2003). To compensate for these inherent problems, some researchers have used a hybrid approach that incorporates both ecological and spectral data (e.g. Franklin 1995; Dirnböck et al. 2003; Dobrowski et al. 2008). Such hybrid approaches combine the structure and disturbance information from the satellite imagery with the potential habitat information for individual species from the SDM. One of the limitations of increasing the number of variables in a classification analysis is that it increases the size of the sample needed to train the classification. One solution to this is to use a binary classifier to target the analysis on a single class of interest. By reducing the number of classes of interest, it requires smaller training sets than classifying multiple classes (Boyd et al. 2006).

As with traditional SDMs and satellite image classification, there is a large number of statistical and remote sensing methods that can be used to generate the hybrid models. One of the major divisions in remotely sensed image classification techniques is between pixel-based and object-based image analysis. Pixel-based classification has limited ability to incorporate information from neighbouring pixels into the classification. When applied to very high resolution imagery, the increased variability of the pixels often results in speckled classification results (Dobrowski et al. 2008; Blaschke 2010). In contrast, Geographic Object-based image analysis (GEOBIA) starts by segmenting the image into areas of contiguous, spectrally similar pixels, known as objects, and the classification is then applied

110

to the objects rather than the pixels. GEOBIA is becoming increasingly popular as a method for managing very high resolution (VHR) satellite imagery, because the pixels are often smaller than the individual entities to be mapped (Blaschke 2010). However, it is not always obvious *a priori* which level of analysis would be most appropriate for a given mapping application. The richest spectral data and finest spatial resolution are retained when analysing data at the pixel-level. In contrast, GEOBIA makes the classification more efficient by generalising the spectral distinctions among neighbouring pixels. The price of this generalisation is the risk of incorporating pixels that genuinely belong in different classes into a mixed object (Wang et al. 2004). Where the pixel size is smaller than the entities being mapped, as often occurs with VHR imagery, there are clear advantages to an object-based approach, which disappear when the entities being mapped are typically smaller than an individual pixel (Blaschke 2010). It is less clear which of the two approaches is likely to be most useful for mapping entities that vary in size from smaller than a single pixel to larger than multiple contiguous pixels (Addink et al. 2007).

5.2.1. **Random Forests for image classification**

Ecological datasets are often complex, unbalanced, and noisy. Collecting field samples is expensive and time consuming, while extracting derivatives from geographic data such as satellite images and digital elevation models (DEMs) is comparatively cheap and easy. Thus, generating predictive variables is often simpler than collecting extra field samples, and it is hence difficult to acquire sufficiently large field samples to conduct reliable statistical regressions of all the potential predictive variables. Further, when the relationships among the variables are complex and non-linear, traditional statistical classification techniques often fail to find meaningful ecological patterns (De'Ath and Fabricius 2000).

To improve the interpretation of such data, researchers have begun using tools from the field of machine learning, including random forests (RF) classification (Cutler et al. 2007). RF is an ensemble classifier that builds a forest of classification trees, using a different bootstrapped training sample

111

and randomly selected set of predictor variables for each tree. Unweighted voting is then used to produce an overall prediction for each site in the sample (Breiman 2001; Liaw and Wiener 2002). RF has been shown to perform well in comparison to decision trees and other ensemble classifiers (Pal 2005) and is able to capture complex, non-linear interactions among noisy, non-normal predictor variables (Gislason et al. 2006; Cutler et al. 2007). In addition, RF provides measures of variable importance that can be used for exploratory ecological interpretation (Pal 2005; Gislason et al. 2006; Cutler et al. 2007). RF has been used to classify vegetation with very high accuracy in a number of mapping applications, including mountain forest communities (Gislason et al. 2006); cropping (Pal 2005); invasive species (Lawrence et al. 2006; Cutler et al. 2007); and predicting rare species distributions (Cutler et al. 2007).

5.2.2. **Contributions of this paper**

In this study, we examine the use of several of these new tools for mapping the distribution of an endangered cushion plant that is endemic to sub-Antarctic Macquarie Island. *Azorella macquariensis* was listed as critically endangered under Australia's Environment Protection and Biodiversity Conservation Act 1999 and endangered under equivalent Tasmanian legislation, after we discovered widespread dieback in late 2008 (Department of Primary Industries 2009; Bergstrom and Bricher, unpublished data). To monitor and understand changes in the distribution of *A. macquariensis*, a first step is to produce a high-accuracy fine-scale map of its distribution at the time the dieback was first discovered.

We extract a vegetation index and several texture measures from VHR WorldView-2 satellite imagery and derive several environmental variables from a fine-resolution digital elevation model (DEM). We examine the effect of three training methods (two pixel extraction methods and object-based analysis) on the accuracy of the image classifications. We then compare the performance of each of these. Finally, we test whether the best-performing classification is able to accurately predict sparse and moderate cover of *Azorella*. The primary aim of the study is to maximise the accuracy of

the mapping, with a secondary aim of improving the understanding of the autecology of *A. macquariensis*. Additionally, we explore whether these methods can reliably separate sparse and moderate cover of *Azorella* or improve understanding of the autecology of *Azorella.*

5.3.   Methods

5.3.1. Study Site and Datasets

We investigated the distribution of an endemic cushion plant on subantarctic Macquarie Island (54°30'S, 158°57'E). The 12,390 ha island is dominated by a plateau at 200 – 400 m altitude, with the endemic cushion plant *Azorella macquariensis* dominating the highest parts of the island. *Azorella* occurs as small cushions on the fringes of mid-altitude plateau grasslands, forms extensive cushions on east-facing higher slopes, and then becomes progressively smaller and sparser in feldmark (vegetation with less than 50% cover) and in the extreme polar desert zones (Fig. 1). In feldmark and polar desert, *Azorella* is usually the dominant vascular plant species, with other species occurring as epiphytes or between cushions. The distribution of cushions tends to be patchy at all spatial scales ranging from tens of centimetres to the entire plateau. There is therefore no 'natural' scale of analysis, and both image pixels and objects are likely to contain mixtures of *Azorella* and other species. Thus, the entity being mapped is a patch that contains *Azorella*, and the minimum size of that entity is uncertain and variable.
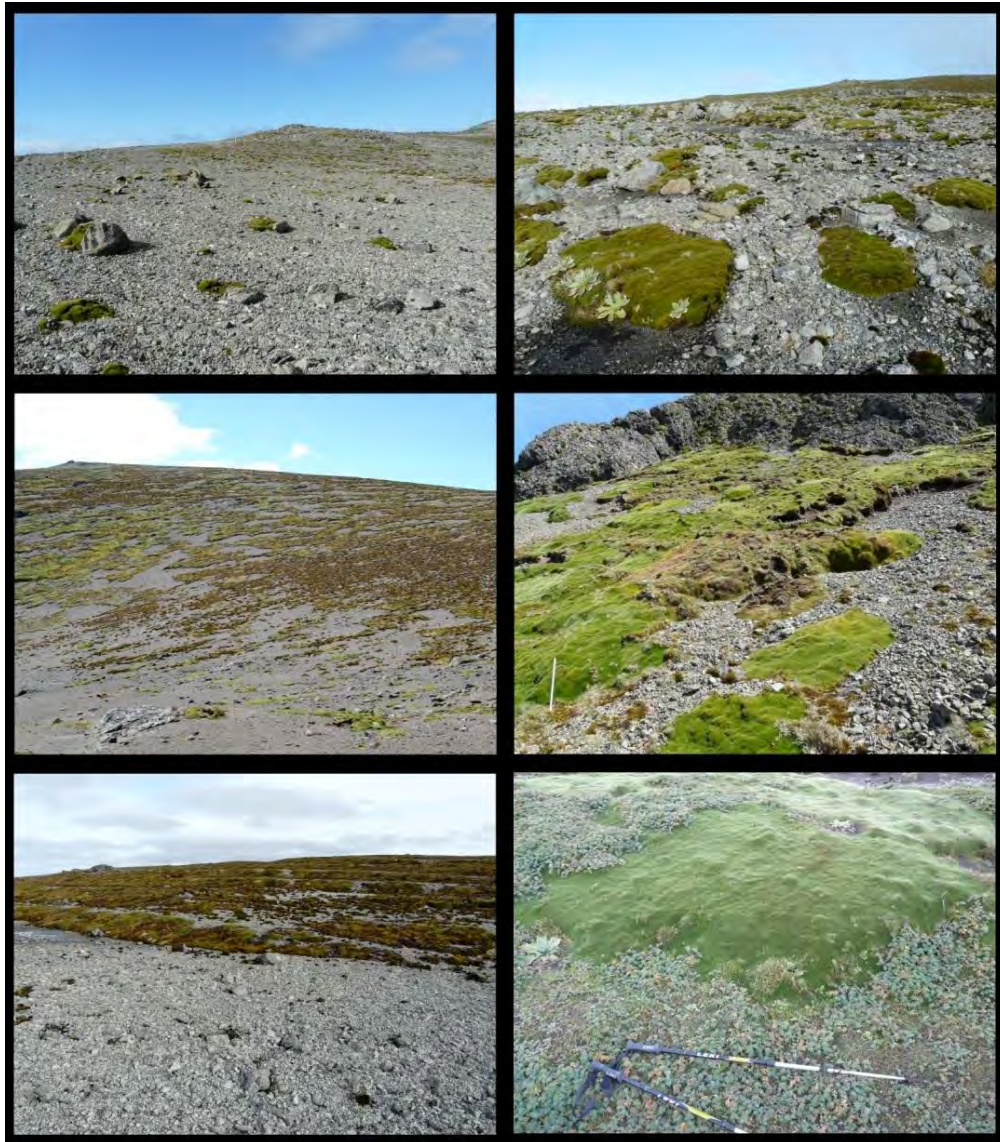
Fig. 1: Azorella *exhibits a range of growth patterns on Macquarie Island, from sparse polar desert (top left) to dense herbfields (bottom right). This variability increases the challenges involved in modelling its distribution.*

On sub-Antarctic islands, *Azorella spp*. are widely considered to be keystone species because they

provide a large proportion of the biomass in feldmark environments and create microhabitat for

other species (Le Roux et al. 2005). The sub-Antarctic region has experienced significant climate

change in the past 50 years (Bergstrom and Chown 1999; Pendlebury and Barnes-Keoghan 2007) and

sub-Antarctic *Azorella* species are expected to shift their distribution in response to that change,

due to their vulnerability to dry conditions and shading from neighbouring plants (le Roux et al.

2005) . On Macquarie Island, *A. macquariensis* has recently been observed to undergo extensive die-

back, leading to it being listed as critically endangered (Department of Primary Industries 2009;

Bergstrom and Bricher, unpublished data). The reasons for the dieback are still not conclusive

despite three years of research. As much dead cushion is now being blown away leaving bare peats

exposed to the wind, the future of the species is bleak, at least in the short term. There is a need for

fine-scale and accurate spatial data for monitoring changes in *Azorella* distribution to help

researchers understand the causes of the dieback; develop and assess the effectiveness of

conservation efforts; and quantify the impact of climate change.

5.3.1.1.   Field data

Macquarie Island was visited over two summers in 2008/09 and in 2009/10. During these visits 349

sites were visited, of which 201 were in the cloud-free portion of the available satellite image (Fig.

2). Most sites were located using a geographically stratified random sampling design we developed,

named GeoStrat, with some additional sampling purposively located in homogeneous patches of

perceived plant communities, as part of an ongoing study of vegetation on the island (see Bergstrom

et al. 2009a). GeoStrat used an unsupervised fuzzy *c*-means classification of the island based on six

terrain variables that were anticipated to affect microclimate (elevation, slope, solar radiation,

surface curvature, topographic wetness index, and topographically-deflected mean wind speed) and

a normalised difference vegetation index (NDVI) derived from WorldView-2 satellite imagery (see

chapter 2 for details).   Of the 201 sites in the cloud-free portion of the image, 172 were positioned

by GeoStrat and 29 were purposively located.

At each site, a 10 x 10 m plot was laid out and an 8.75 m^2 vertical photograph was taken at each

corner of the plot (i.e. four photographs per plot). Point-intercept analysis was used to estimate the

percent cover of all vascular plant species, using Coral Point Count software (Kohler and Gill 2006) to

lay 100 randomly located points over each photograph. The results were averaged for each site.

Azorella occurred in one-third of the sites, forming 0.25 – 26.6% of the canopy at those sites. Initial

attempts to apply regression models to the percent-cover data proved unsuccessful, as evidenced by

a random forest multiple regression, using all input variables that explained 29.1% variance with a

mean square residual of 10.2, We subsequently switched focus to classifying the data into two or

three classes. The two-class classification divided the sites into present and absent. The three-class

classification divided the sites into absent, sparse and moderate cover, with the boundary between

the sparse and moderate classes set at 5% cover.

An independent validation dataset was acquired from another ongoing study of *Azorella* on

Macquarie Island (Bergstrom, unpub. data). This dataset comprised 88 randomly located sites that

contained *Azorella,* showing presence only.
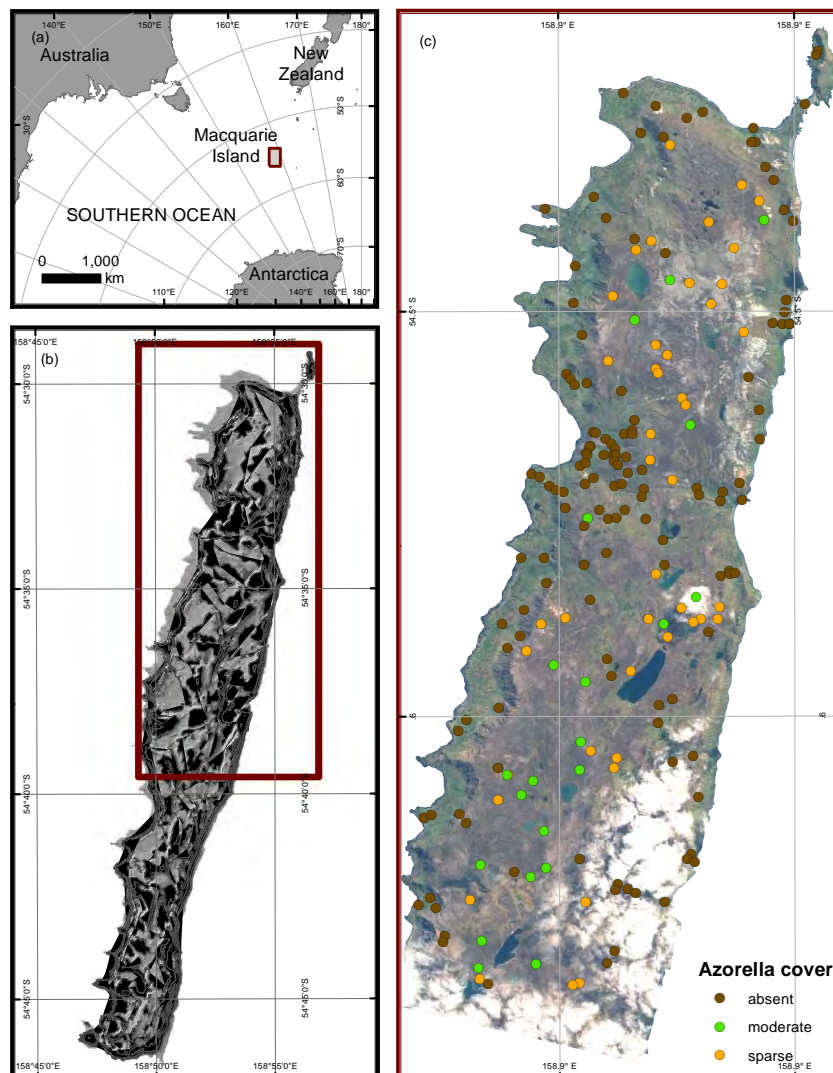
Fig. 2: Field sites on northern Macquarie Island, showing the observed presence or absence of Azorella. Panel (a) shows the location of Macquarie Island in the Southern Ocean; (b) shows the extent of the island; and (c) shows the extent of the WorldView-2 image and location of the training samples. Those field plots in areas obscured by cloud in the image were excluded from the analysis.

5.3.1.2. Satellite imagery

For northern Macquarie Island, two recent images very high resolution (VHR) satellite imagery were available for analysis. A cloud-free QuickBird image of the entire island with 2.4 m pixels obtained in March 2005 was used for the GeoStrat sampling design. A mostly cloud-free WorldView-2 image, with 2 m multi-spectral pixels and 8 spectral bands, was obtained in December 2009. From this image, we calculated an NDVI layer and a set of texture measures, which have been shown elsewhere to improve the accuracy of image classification (e.g. Murray et al. 2010). NDVI is a widely used index that captures the relative proportions of red and near infrared (NIR) reflectance in a satellite image, as a proxy for the amount of live vegetation in a pixel (Sellers 1985). Texture measures provide information about the spatial context of a pixel (Murray et al. 2010). Here, we used grey-level co-occurrence matrix (GLCM) texture measures (Haralick et al. 1973; Franklin et al. 2001). The GLCM was calculated for the NDVI image using an 11 x 11 cell kernel. GLCM computes a matrix that compares the greyscale values of neighbouring cells in a moving window. From the GLCM, we calculated eight texture measures: mean, variance, homogeneity, contrast, dissimilarity, entropy, angular second moment and correlation (Haralick et al. 1973).

5.3.1.3. Terrain data

Elevation values were taken from a 5 m resolution digital elevation model (DEM) of Macquarie Island, with heights accurate to 5 m, derived from Airborne Synthetic Aperture RADAR data acquired in 2000 by the NASA PACRIM Mission 2 (Tapley et al. 2004; Brolsma 2008). From this DEM, we derived the following terrain variables, which we considered likely to affect the distributions of plants on the island: slope, aspect, topographically-deflected wind speed, solar radiation, topographic wetness index, surface curvature (including planar and profile curvature), ridgeness, valleyness and distance from the coast (see Chapter 3 for details).

To explore the controlling environmental parameters influencing the distribution of *Azorella* we incorporated a suite of terrain variables, as proxies for direct environmental variables (van Niel et al. 2004), into the analysis. The vegetation on Macquarie Island exhibits strong altitudinal gradients,

although elevation only indirectly affects plant physiology (Guisan & Zimmermann 2000; Selkirk et al.1990). *Azorella* species in the sub-Antarctic are known to be affected by wind and precipitation patterns (Le Roux et al. 2005). Spatial data on precipitation were unavailable, but some terrain variables are known to affect water flow patterns, namely surface curvature (including both planar and profile curvature). Additionally, the topographic wetness index models potential areas of water accumulation (for details, see chapter 2). The prevailing wind on Macquarie Island comes from the west and north west, at a mean speed of 35.1 km/h (Bureau of Meterorology 2011), and we used a topographically-deflected wind speed model to estimate the wind speed across the island (Liston and Sturm 1998; Wallace 2006). Solar radiation, as the source of energy for photosynthesis, is likely to have a direct impact on vegetation distributions. Distance from the coast correlates with the distributions of several plant species on Macquarie Island, including Azorella, so we calculated the surface distance from the coastline. *Azorella* appears to be more common in higher, more exposed sites than in gullies, even at high elevations. To test this, we calculated two multi-scale measures of topographic position, namely ridgeness and valleyness, using the multi-scale landform classification algorithm in the LandSerf software (Wood 1996). This calculated the proportion of scales at which each cell occurred in a ridge or valley for neighbourhood sizes ranging from 3 x 3 to 49 x 49 cell neighbourhoods. For species distribution modelling at broad scales, measures of climate variability are often incorporated (Guisan and Zimmermann 2000). Macquarie Island is sufficiently small and isolated that there is only a single meteorology station, and it is hence not possible to interpolate climate variables across the island, although a lapse rate of around 0.8°C per 100m has been recorded (Tweedie and Bergstrom 2000) and this is included by way of proxy data in the measure of elevation.

5.3.1.4. Random Forest Classification

Random forest classification (RF) is increasingly used for mapping vegetation from remotely sensed data (e.g. Pal 2005; Gislason et al. 2006; Cutler et al. 2007; Grossmann et al. 2010), and was used in this study to predict the presence of *Azorella* on the basis of high resolution terrain and spectral

data. There are only two user-determined variables to set: the number of variables in the random subset at each node, and the number of trees in the forest. RF is claimed to be relatively insensitive to the first of these (Liaw and Wiener 2002). We used the *randomForest* package in R (Liaw and Wiener 2002) to construct 5000 trees per classification, and plotted the error rates as a function of the number of trees. The number of input variables for each forest varied, but for all forests, a random subset of three input variables was used to split the data at each node of each tree. When the classes in an RF classification are unbalanced, the error rates are highest in the rarest classes (Evans and Cushman 2009). A few techniques have been suggested to correct for this (e.g. Liaw and Wiener 2002; Evans and Cushman 2009; Grossmann et al. 2010) but we chose not to apply them, as the dataset was only moderately unbalanced. Instead, we chose to use the predictive accuracy for the smaller 'present' class as the primary measure of model performance, rather than overall accuracy.

RF produces multiple outputs to aid in interpreting the results. In addition to the hard class predictions, it calculates the probability of membership for each class, out-of-bag (OOB) accuracy estimates, variable importance measures, and partial dependence plots. The OOB accuracy estimate is an internal accuracy measure that is reported to be unbiased and slightly more conservative than independent validation (Liaw and Wiener 2002; Gislason et al. 2006). It is calculated by withholding a random subset of 36% of the training set (the out-of-bag set) from the construction of each tree, then applying the out-of-bag set to that tree. For each data point, its class for the entire forest is calculated by counting the votes from each tree for which it was in the out-of-bag set. The OOB error estimate for the forest is the proportion of data points for which the voting classified them incorrectly (Liaw and Wiener 2002). The OOB accuracy is the inverse of the OOB error estimate.

The variable importance measures and partial dependence plots aid in interpreting the impact of the input variables. To calculate the variable importance, the mean decrease in accuracy is calculated according to the increase in prediction error when OOB data for that variable is permuted and all

other variables are left unchanged (Liaw and Wiener 2002). The partial dependence plots show the relationships between individual predictor variables and the predicted probability of the presence of *Azorella* (Cutler et al. 2007).

5.3.1.5. Experimental design

This paper examined the role of several analysis approaches in improving the accuracy of RF classification for mapping the distribution of *Azorella*. These comparisons were as follows:

- Pixel-based versus object-based image analysis

- Within pixel-based classification, training the classification using values from a single image pixel in the centre of a field plot, versus multiple contiguous pixels that intersect with the field plot boundaries

- The role of terrain, spectral, and hybrid sets of input variables

- Producing reduced classifications using a hypothesis-driven subset of the input variables versus using the variable importance measures to select input variables for reduced classifications.

- The number of classes: a binary classification versus a three-class classification

We first tested the reliability of RF OOB accuracy estimates and plotted the error rates against the size of the forests. We then built a series of forests, using all combinations of training methods (single pixel, multiple pixel, and object); input variables; variable reduction methods; and class number (Table 1).

Table 1: The combinations of sampling approach, input variables and the number of classes for the classifications compared in this study.

Sampling Approach	Input Variables	Number of Classes
Single pixel	Terrain	Absent/Present
Multiple pixel	Spectral	
	Hybrid	Absent/Sparse/Moderate
Object	Hypothesis-driven	

A total of 27 input variables were available for the classifications, including ten terrain variables and 17 spectral variables (eight spectral bands, eight texture measures, and an NDVI layer). For each combination of training method and sets of input variables (i.e. terrain, spectral and hybrid) a forest was first built using all available input variables. Then the variable importance plots were used to indicate which variables contributed most to the classification. Trial-and-error was used to find a minimum set of input variables that maintained the accuracy of the full-model classification. This process showed that including variables with lower importance often decreased the model accuracy by introducing noise. In addition to the improvement of model accuracy, reduced models make ecological interpretation of classifications simpler (Evans and Cushman 2009). For each variable in the reduced models, a partial dependence plot was created to show the marginal effect of that variable on the class probability.

In addition to reduced models with the inputs selected on purely statistical grounds, we created RFs using a subset of variables that we considered *a priori* to be most likely to maximise the ecological and spectral separation of *Azorella* from other vegetation. These variables were the blue, green, yellow, red edge and near-infrared 2 spectral bands; NDVI; the mean, homogeneity and entropy GLCM texture measures; and the elevation, slope, solar radiation, curvature, ridgeness, valleyness, and wetness index terrain variables.

The major purpose of this study was to develop a baseline map for change detection. As the distribution of *Azorella* on Macquarie Island is expected to contract, it was therefore important to test whether these classifications simply identified the high elevation, bare areas of the island or whether they can distinguish these from areas that contain small amounts of *Azorella*. We therefore repeated the classifications using three classes – absent, sparse and moderate cover of *Azorella*.

We assessed the classification accuracy using confusion matrices of the OOB accuracy estimates, a kappa statistic, independent validation of the present class, and visual inspection of the resulting

maps. As OOB accuracy tends to be lowest for the smallest class, the primary OOB measure of interest was the producer's accuracy for the present and moderate cover classes.

5.3.1.6. Testing pixel- and object-based classifications

The most appropriate scale of analysis can be difficult to discern in advance, especially for a species where individuals range in diameter from a few centimetres to several metres, and which is patchy at multiple spatial scales. We therefore classified both pixels and small objects. Pixel-based analysis allows mapping at the finest resolution, but limits the amount of contextual information that can be incorporated (Dobrowski et al. 2008).

We tested two pixel-based training methods. For the single pixel approach, we extracted values for the image pixel at the centre of the study plot at each site. For the multiple pixel approach, we extracted values for each image pixel that intersected with the plot boundaries at each site. The former approach ensured that the selected pixel was within the boundaries of the study plot, regardless of errors in GPS positioning or in georeferencing of the image. The latter approach captured the variability of spectral values within a plot, at increased risk of erroneously including pixels from outside the plot boundaries. This latter approach is also vulnerable to spatial autocorrelation, but is commonly used in remote sensing applications (Mannel et al. 2011).

For the object-based RFs, the WorldView-2 image was divided into objects using the multi-resolution segmentation algorithm in eCognition Developer 8 software. This segmentation algorithm divides an image into homogeneous regions by grouping neighbouring pixels based on their Euclidean distance in multivariate attribute space. The size of the objects is determined by the scale parameter, which sets the threshold for homogeneity within each object. The value of the scale parameter is generally determined by trial and error, because there are no objective methods available to choose an appropriate value (Luscier et al. 2006). Here, we used the pixel values in the eight spectral bands of the WV-2 image to identify image objects. We set the scale parameter to 35, and the shape parameter to 0 so that the objects could take any shape. We considered this to be most appropriate

for cushion plants, which mix with mosses to form clumps ranging in shape from almost circular to long, thin, inter-connected terraces. At this scale, objects were observed to follow the edges of these terraces, and had a mean area of 208.4m^2. The mean values for the terrain derivatives, spectral bands, NDVI, and texture measures were calculated for each object. These mean values were then subjected to the same RF classification procedure as the pixels.

In addition to reduced models with the inputs selected on purely statistical grounds, we created RFs using a subset of variables that we considered *a priori* to be most likely to maximise the ecological and spectral separation of *Azorella* from other vegetation and bare ground. These variables were the blue, green, yellow, red edge and near-infrared 2 spectral bands; NDVI; the mean, homogeneity and entropy GLCM texture measures; and the elevation, slope, solar radiation, curvature, ridgeness, valleyness, and wetness index terrain variables.

5.3.1.7.   Accuracy Assessment

The major accuracy assessment tool for RF is the out-of-bag (OOB) accuracy estimate, which was calculated for each class. OOB errors are claimed to be accurate, so long as enough random forest trees have been grown (Liaw and Wiener 2002). This approach allows all data to be used to train the models, improving the efficiency of field data collection (Lawrence et al. 2006). To test this assumption, we randomly divided the dataset into 50/50 training/test datasets. RF classifications were built on the training datasets and then applied to the test datasets, to assess the validation accuracy. We compared the validation accuracy with the OOB accuracy estimate for a classification built on the entire dataset. Additionally, Kappa statistics were calculated for each classification, using the *epiR* package, to measure the whether the agreement between predicted and observed values beyond the level expected from chance alone.

Once the classification had been finalised, we applied it to all pixels/objects in the image to produce a predictive map. The present class was then validated against the independent set of sites that are known to contain *Azorella*.

5.4.   Results

5.4.1. **Testing random forest classification**

RF classification is claimed to be relatively insensitive to the number of trees in the forest. In this study, the error rates stabilised before the forests reached 1000 trees with values between 0.005 and 0.2, and the highest errors in the presence class (Fig 3).
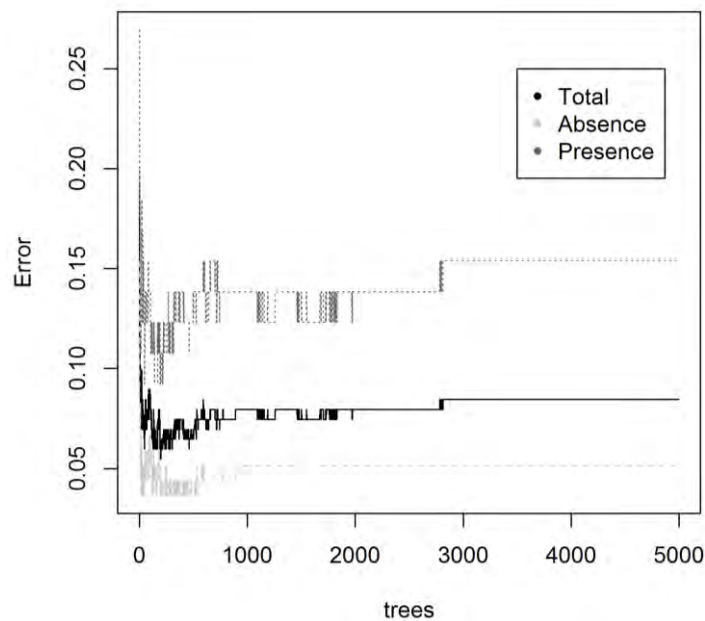
Fig. 3: Error rates plotted against the number of trees for single pixel RF based on all input variables. Black shows the overall error rate, light grey shows the error rate for the presence class and dark grey shows the error rate for the absence class.

Validation of the out-of-bag (OOB) accuracy estimates showed that where the number of observations was low, the OOB accuracy estimate was conservative in comparison to test validation, underestimating the classification accuracy by a small amount in most cases, but by 12% for the smallest present class in the single-pixel classification, and with the kappa statistic varying by up to 0.084. This discrepancy disappeared for the object and multiple-pixel approaches (Table 2). We therefore considered that OOB errors provide a suitable but conservative accuracy measure for comparing models within this study. Comparison with accuracy values from other studies, however, should be done cautiously. Other researchers have also found that the OOB error underestimates

accuracy compared with validation (Gislason et al. 2006; Hüttich et al. 2011), but have not reported

such a large offset in the error estimates as in the present study (Lawrence et al. 2006). In the

present study, the disparities between the OOB and validation accuracy measures were typically

small, so we considered it to be still useful for comparing models. Additionally, because of the

relatively small size of the field sample for the single-pixel and object-based approaches and the

availability of an independent dataset for the present class, we chose not to sacrifice data for an

independent validation.

Table 2: Confusion matrices of the accuracies of the single pixel-based (A); object-based (B); and multiple pixel-based (C); classifications of field sites on the basis of all input variables. The Test rows show the accuracy scores from validation of an RF classification built on a 50% training set. The Full model OOB rows show the accuracy scores for an OOB estimate of model accuracy for an RF classification built on the entire dataset.

(A)		Absent	Present	Accuracy	Kappa
Test	Absent	64	4	94.1%	
	Present	1	31	96.9%	
	Accuracy	98.5%	88.6%	95%	0.888
Full model OOB	Absent	129	7	94.9%	
	Present	10	55	84.6%	
	Accuracy	92.8%	88.7%	91.5%	0.804

(B)		Absent	Present	Accuracy	Kappa
Test	Absent	65	2	97.0%	
	Present	4	28	87.5%	
	Accuracy	94.2%	93.3%	93.9%	0.859
Full Model OOB	Absent	128	7	94.8%	
	Present	8	57	87.7%	
	Accuracy	94.1%	89.1%	92.5%	0.828

(C)			Absent	Present	Accuracy	Kappa
Test Validation		**Absent**	1718	0	100.0%	
		Present	2	828	99.8%	
		Accuracy	99.9%	100.0%	99.9%	0.998
Full model OOB		**Absent**	3435	1	100.0%	
		Present	2	1658	99.9%	
		Accuracy	99.9%	99.9%	99.9%	0.999

Across all the classification models that we compared, both the OOB and validation accuracy measures showed several consistent patterns. The statistically-driven hybrid model and hypothesis-driven subset of input variables had the highest accuracies, while classifications based on spectral variables alone had the lowest. The multiple pixel training method consistently had the highest accuracies, though its inflation of the OOB accuracy estimates due to spatial autocorrelation, caused by extracting data from multiple adjacent pixels in each plot limited interpretation. The object-based classifications consistently had the lowest accuracies. The effect of training method was the most important predictor of model accuracy; such that multiple pixel based classifications had equal validation accuracy, regardless of which subset of input variables was used (Fig. 4 and Table 3).

The multiple hypotheses tested in this study created voluminous results. Here we present detailed results for the most important models, but for completeness present all classification results in Appendix B.
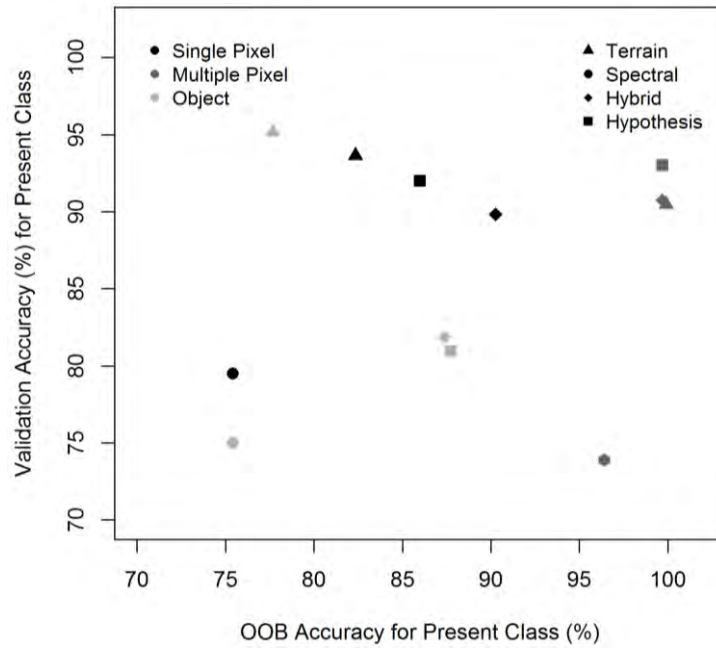
Fig. 4: The OOB and validation accuracy measures for all two-class classification models. OOB over-estimated the accuracy of the multiple pixel-based models. This was most obvious for the multiple pixel-trained spectral classification, in which the OOB predicted 96.4% accuracy for the present class, while independent validation showed that the accuracy of that class was 73.9%.

Table 3: Summary of the performance of the two-class classification models. We used four methods of assessing the reliability of the classifications, which need to be considered jointly. The OOB and its resulting kappa were unreliable for the multiple pixel-trained classifications, due to spatial autocorrelation. Validation data were only available for the present class, so could not assess whether the models over predicted the distribution of Azorella. *Visual inspection acted as a final, if subjective, check of the models' performance.*

Input Variables	Training Method	OOB Reliability	OOB (Present)	Kappa (from OOB)	Validation (Present)	Visual Inspection	Overall Performance
Hybrid	Single Pixel	High	89.2	0.862	89.8	Excellent	Excellent
	Multiple Pixel	Low	99.9	0.999	90.9	Excellent	Best
	Object	Moderate	89.2	0.851	81.8	Good	Good
Hypothesis-driven	Single Pixel	High	86.5	0.817	92	Excellent	Excellent
	Multiple Pixel	Low	99.8	0.996	93.2	Excellent	Best
	Object	Moderate	87.7	0.839	80.7	good	Good
Terrain	Single Pixel	High	81.5	0.717	93.2	Good	Good
	Multiple Pixel	Low	100	0.9996	90.9	V. good	V. good
	Object	Moderate	76.9	0.689	95.5	Good	Good
Spectral	Single Pixel	High	75.4	0.666	79.5	Poor	Poor
	Multiple Pixel	Low	96.4	0.949	73.9	Poor	Poor
	Object	Moderate	75.4	0.666	75	Poor	Poor

5.4.1.1. Testing the training method

Using a combination of spectral and terrain variables, we tested the effect of the training method on

the accuracy of the image classifications, using two measures of accuracy. Both OOB accuracy

estimates and independent validation of the present class showed that the multiple pixel approach

provided the highest accuracy, though it provided only a small improvement over the single pixel

sampling method (Fig. 5). The object-based method was the least accurate by both measures. The

OOB accuracy estimate tended to over-estimate the accuracy of the multiple pixel training method,

producing accuracies near 100% for every subset of input variables we examined. It appeared to be

vulnerable to the effects of spatial autocorrelation in the sampling method. In contrast, the

independent validation estimate was not affected by this.

The random forest classifications based on a combination of terrain and spectral input variables

accurately predicted the presence of *Azorella* with very high accuracy. Independent validation of the

present class showed that the single pixel (89.8%) and multiple pixel (90.9%) approaches had very

similar accuracy, with both performing better than the object-based classification (81.8%). The OOB

accuracy estimate could not separate these three models.

 After inspection of the variable importance plots, the same input variables were selected for both

the single pixel- and object-based classifications: elevation, solar radiation, coast distance, NIR 1, NIR

2, red edge, NDVI and the GLCM Mean texture measure. These variables were also selected for the

subsets used in the terrain and spectral data classifications. The multiple pixel classification required

most of the same variables, but replaced the two NIR spectral bands with ridgeness.

The partial dependence plots (Fig. 6) showed that *Azorella* presence typically occurs at elevations

above 200 m, with NIR 1 reflectance values below 550, NDVI values below 0.55, GLCM Mean values

less than 50, distances from the coast greater than 600 m, the highest values for solar radiation (>

5.6 MWh/m^2), red edge reflectance below 650, and NIR 2 reflectance less than 850 (out of 2048 DN

values).

On Macquarie Island, *Azorella* typically grows on the highest parts of the island (i.e. those areas with high elevation, solar radiation and distance from the coast); with low to medium reflectance in the NIR and red edge portions of the spectrum, and hence to be in open vegetation; and with low values in the NDVI and the mean values of the GLCM matrix based on the NDVI layer (Fig. 6). The slope, shape of the terrain, topographic position and wetness index had little effect on the classifications, as shown by low variable importance values, and by the fact that excluding them from reduced classification models did not decrease the accuracy of the classifications, and often increased the accuracy marginally. The maps produced by the three classifications appeared very similar, though the multiple pixel training method produced a slightly more fragmented pattern of *Azorella* distribution than either of the other two hybrid classifications. In conjunction with the high validation accuracy, this indicates that the multiple pixel approach was better able to capture the variability in the dataset.

Relying on the variable importance measures to select the variables for inclusion in the reduced model showed that terrain variables were most important to the classification, with the role of spectral data largely being confined to locating areas with sparse vegetation. Spearman rank correlation coefficients showed that the spectral variables incorporated into these models were strongly correlated with NDVI (Table 4), indicating that the major role of spectral variables in these classifications was to select areas with sparse vegetation.

Table 4: Spearman rank correlation coefficients between NDVI and other spectral variables selected in the hybrid classifications based on terrain, spectral and hybrid sets of input variables. Strong correlations between input variables need not result in the variables being excluded from the RF models, but they do complicate the interpretation.

Spectral Variable	Correlation with NDVI
GLCM Mean	0.94
Red Edge	0.82
NIR 1	0.89
NIR 2	0.87
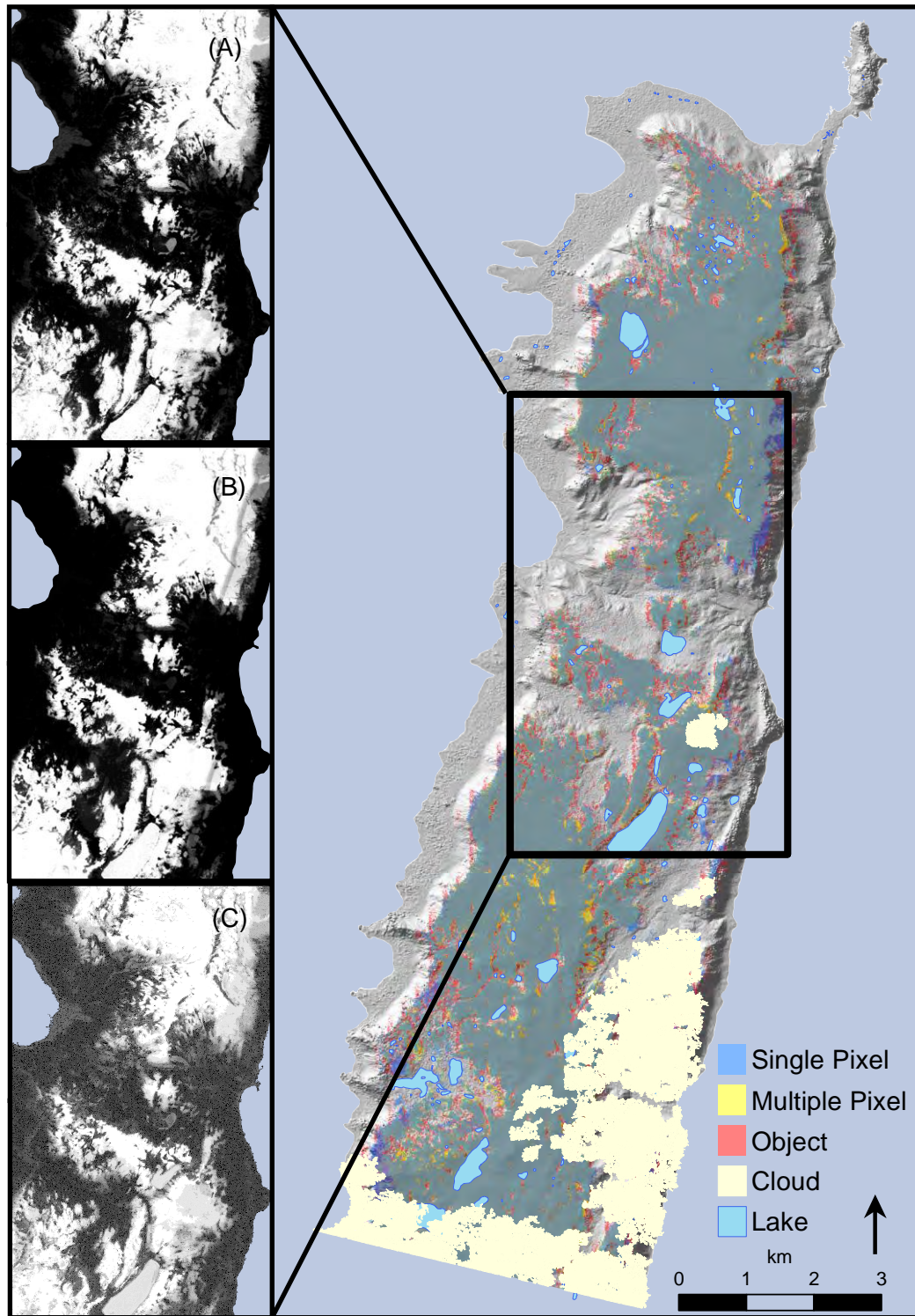
Fig 5: Predicted Azorella *presence on northern Macquarie Island, from single pixel, object and multiple pixel based classifications of the hybrid variables. The main map shows the hard classes for* Azorella *presence, with the individual prediction layers made partially transparent to demonstrate the overlaps in the predicted distributions. The inset maps show the probability of* Azorella *presence based on single-pixel (A); multiple pixel (B); and object-based (C) classifications. The differences among the three image training methods were subtle.*
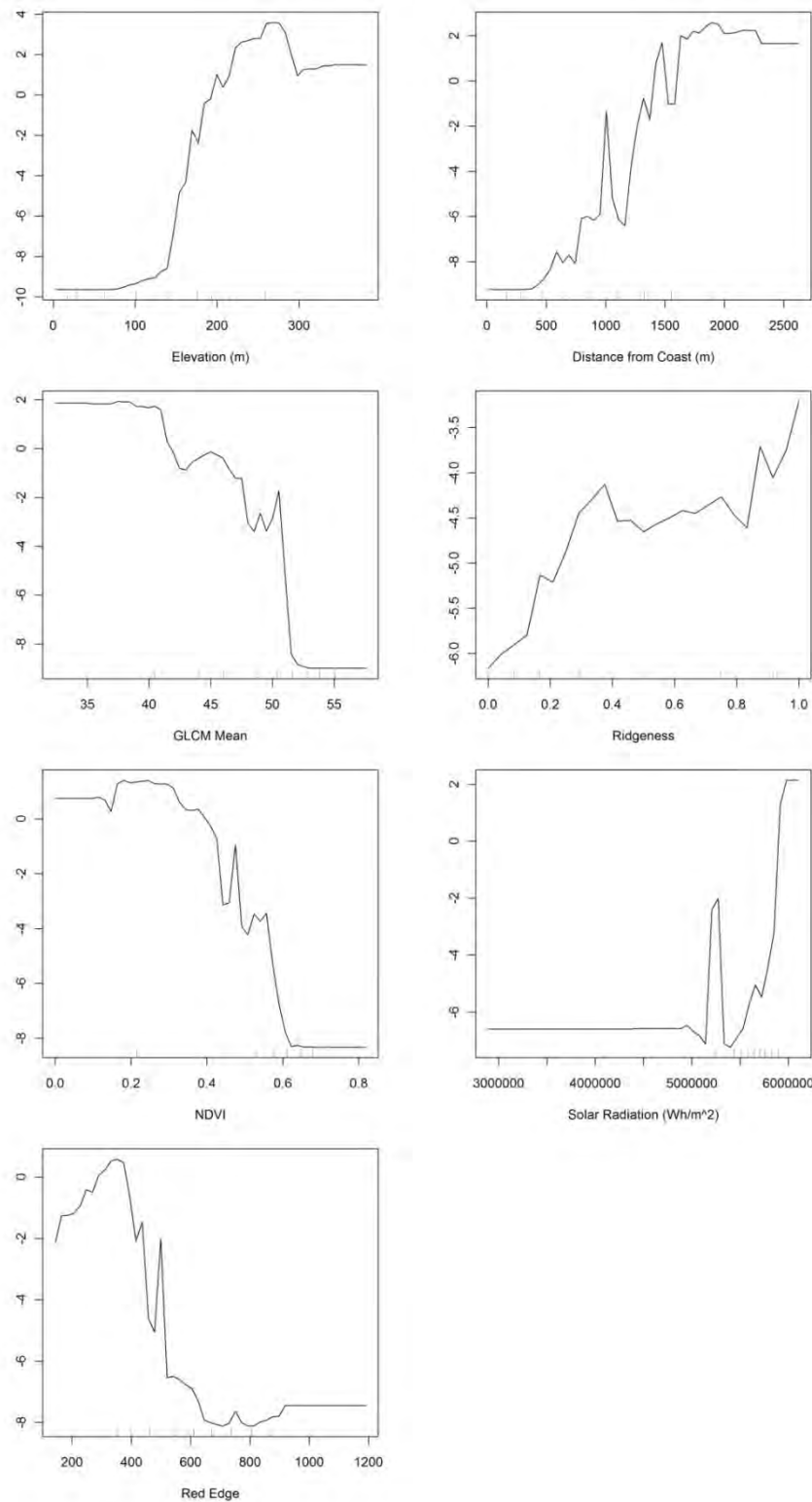
Fig. 6: Partial dependence plots for the variables selected for the reduced hybrid multiple pixel-based classification of Azorella *presence/absence. The variables included were chosen on the basis of the variable importance measures. The y-axis shows a relative measure of the marginal effect on the probability of Azorella presence. Azorella is associated with high values for elevation, distance from coast, ridgeness, and solar radiation; and with low values for GLCM Mean, NDVI and red edge reflectance.*

To further explore the respective roles of terrain and spectral variables in predicting the presence of

Azorella, we produced classifications using the terrain and spectral variables in isolation.

5.4.1.2. Terrain-based classification

Independent validation of the terrain-based classification of the presence or absence of *Azorella*

using the multiple pixel training method showed that terrain-based classification was equally

accurate at predicting *Azorella* presence as the hybrid classification (90.9%).

The OOB and independent validations for the single pixel and object-based classifications conflicted

in their assessments of model fit. OOB indicated that the hybrid classifications were 7.7 – 12.3%

more accurate than the terrain-based ones, while independent validation showed that the terrain-

based classifications out-performed the hybrid ones by 3.4 – 13.7%. These contradictory results are

difficult to interpret, though comparisons of the terrain-based maps indicated that all three training

methods produced very similar spatial distributions. We therefore present full details here for the

multiple pixel training method, and present full results for the other models in Appendix B.

The variable importance measures used to choose the most important variables for inclusion in a

reduced model revealed that the classification was primarily based on elevation, distance from the

coast, solar radiation, aspect, and ridgeness. The partial dependence plots (Fig. 7) showed that

Azorella typically grows above 200 m elevation, more than 1000 m surface distance from the coast;

in areas with the highest annual solar radiation (> 5 600 000 Wh/m^2); and with the lowest

topographically-deflected mean wind speed (<33 kts). The relationship between *Azorella* distribution

and aspect was not clear. The relationships between *Azorella* presence and elevation, coast distance

and solar radiation are very similar to those in the hybrid classifications.

These spatial rules describe the mountain tops and highest parts of the central plateau that

dominates Macquarie Island (Fig. 8). The highest areas have the highest solar radiation levels

because there are no surrounding hills to provide shade from the morning or afternoon sun.

Elevation also correlates with distance from the coastline because the island is part of an oceanic

ridge, with a plateau running along the spine of the island. The slope and curvature measures were not included in any reduced RF classifications, indicating that they had little predictive value for *Azorella* presence. Topographic position was included in one of the terrain-based classifications in the form of the ridgeness variable, but it was not consistently chosen and the valleyness measure of topographic position was never included in a reduced model.

Terrain-based classification of *Azorella* presence is a form of SDM, and hence maps potential, rather than actual, *Azorella* distribution. It might therefore be expected to over-estimate the area of the present class, unless *Azorella* is currently occupying its entire potential habitat (Franklin and Miller 2009). High accuracies in the OOB estimates for the absent class indicated that the terrain-based classifications were not badly over-predicting *Azorella* distribution. Inspection of the resulting maps, however, showed that the terrain-based classifications predicted presence in large homogeneous blocks, rather than the fragmented patterns seen with the other classifications in this study, and that are seen in the field (Fig. 8 and Appendix B).
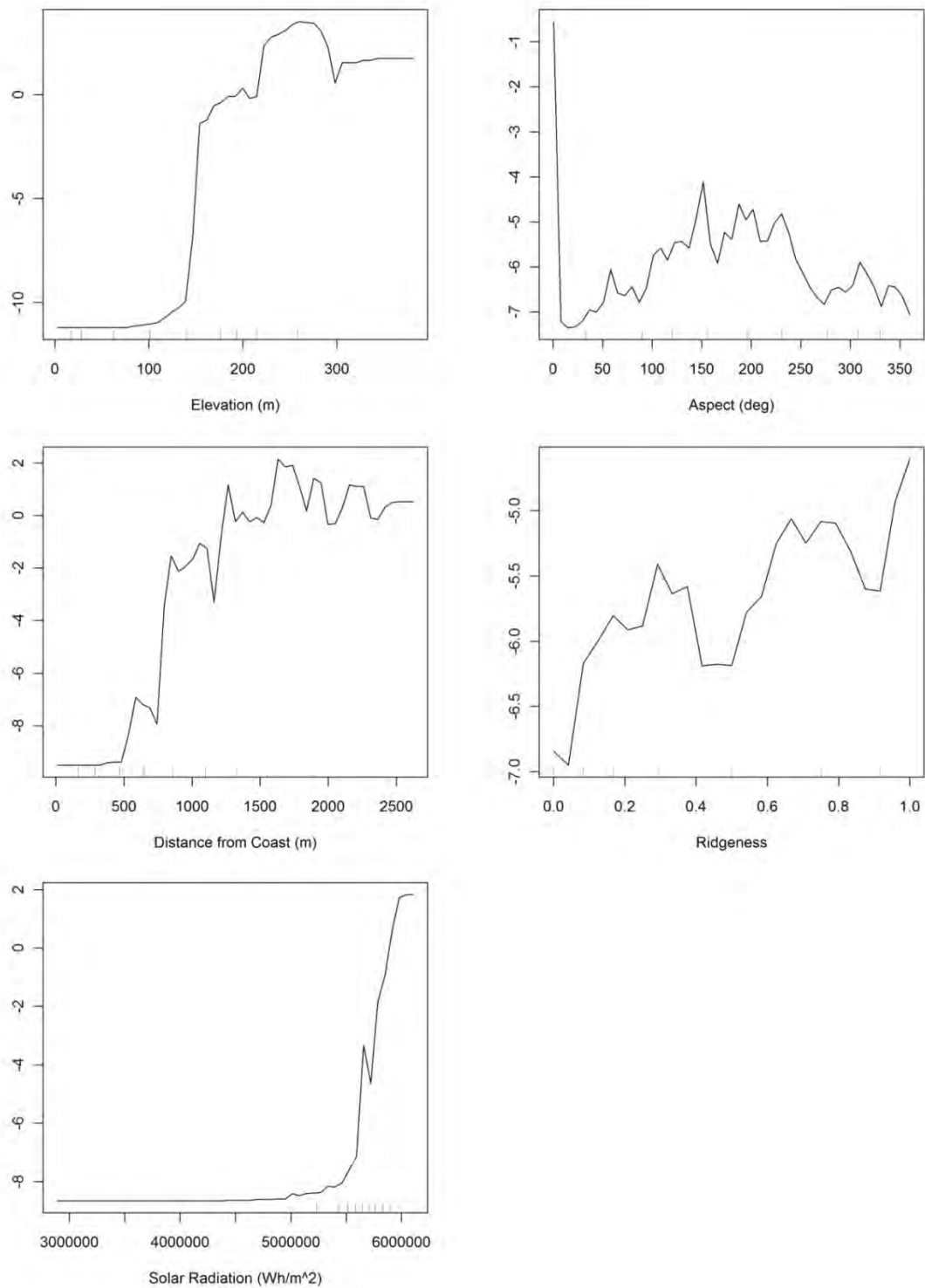
Fig. 7: Partial dependence plots for the random forest classification of terrain variables for the presence and absence of Azorella*, using a multiple-pixel based classification, with the variables chosen on the basis of variable importance measures. The y-axis shows a relative measure of the marginal effect on the probability of* Azorella *presence. Azorella presence is associated with high values for elevation, distance from the coast, ridgeness, and solar radiation; and with mixed values for aspect.*
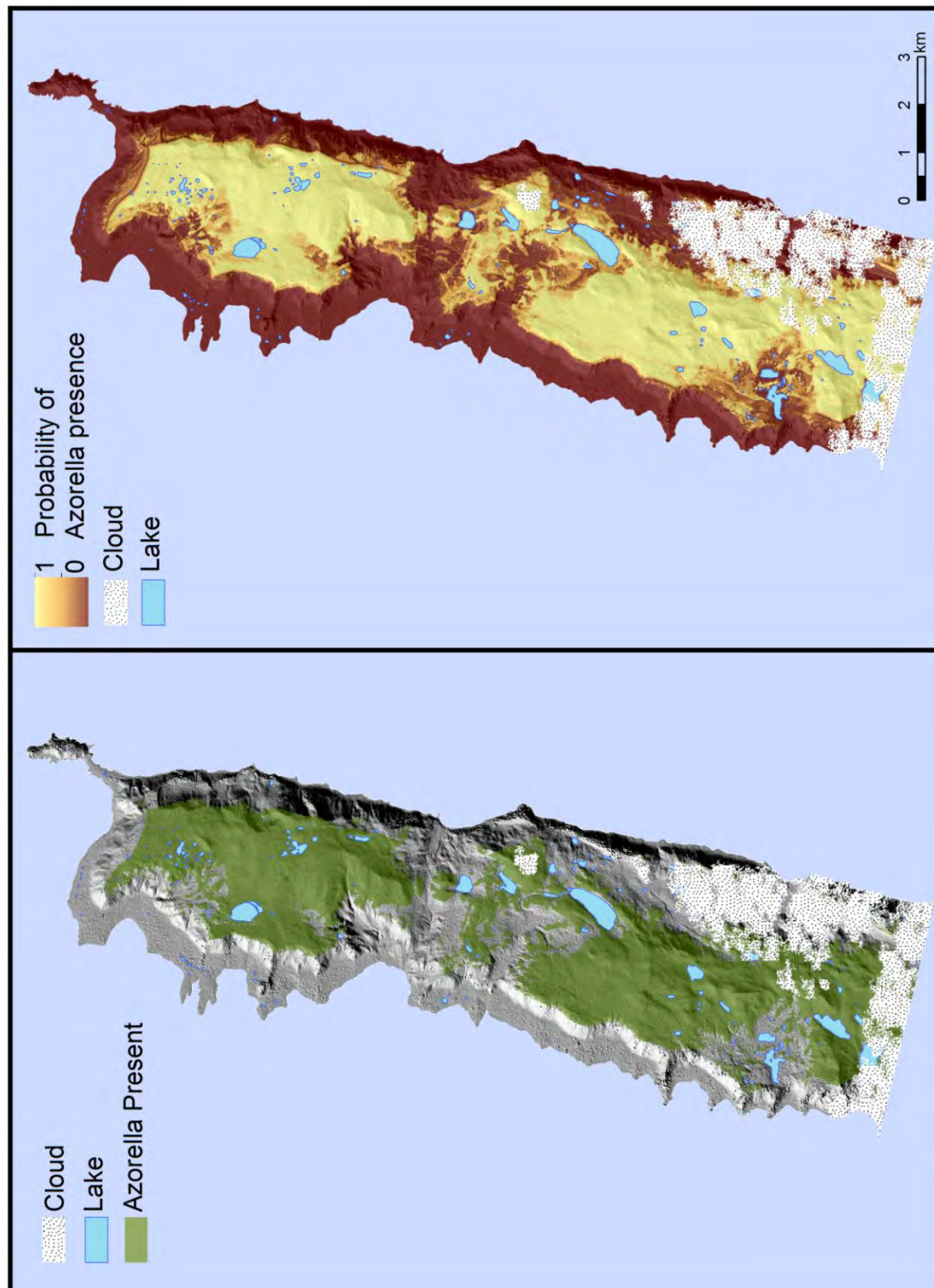
Fig. 8: Terrain-based prediction of Azorella *presence, using the multiple pixel training method. Panel (A) shows the hard class for* Azorella *presence and panel (B) shows the probability of* Azorella *presence. These maps show potential, rather than actual,* Azorella *habitat, and hence predict solid areas of* Azorella *on the highest parts of the island.*

135

5.4.1.3. Spectral-based classification

Classification on the basis of spectral variables and the multiple pixel training method also achieved

90.9% accuracy in the present class, when tested with independent validation. This is the same as

the hybrid and terrain based classifications of the multiple pixel image samples. Classifications of the

objects and single pixel image samples had much lower validation accuracy (75% and 79.5%,

respectively). Thus, the multiple pixel approach appeared to largely compensate for the reduced

accuracy of basing the classification on spectral data only, by capturing the variation in spectral

response within each field site. All three spectral-based classifications produced very similar maps,

so we will discuss the multiple pixel classification here, for consistency. See Appendix B for details of

the single pixel and object-based classification results.

Variable importance measures showed that almost all spectral variables were required to maintain

the accuracy of the classification based on the full set of spectral variables. Four of the GLCM

texture measures were excluded (correlation, dissimilarity, entropy and homogeneity) along with

yellow reflectance. The partial dependence plots (Fig. 9) showed that *Azorella* presence was

associated with low values for GLCM variance and mean, red edge, NDVI, NIR 1 and NIR 2; high

values for GLCM angular second moment and reflectance in the blue and red bands; and extreme

values in the GLCM contrast texture measure and the UV and green reflectance bands. The strong

correlations between the NDVI, mean GLCM, and the red edge and near infrared reflectance bands

reinforce the importance of the sparseness of the vegetation in predicting *Azorella* presence.

The map derived from the spectral-based classification included areas that, for ecological reasons,

could not contain *Azorella*. Most notably, it predicted the presence of *Azorella* on beaches and bare

coastal rocks. This indicated that spectral data alone was unable to adequately describe the

distribution of *Azorella*, despite the high validation accuracy for the present class for the multiple

pixel training method.
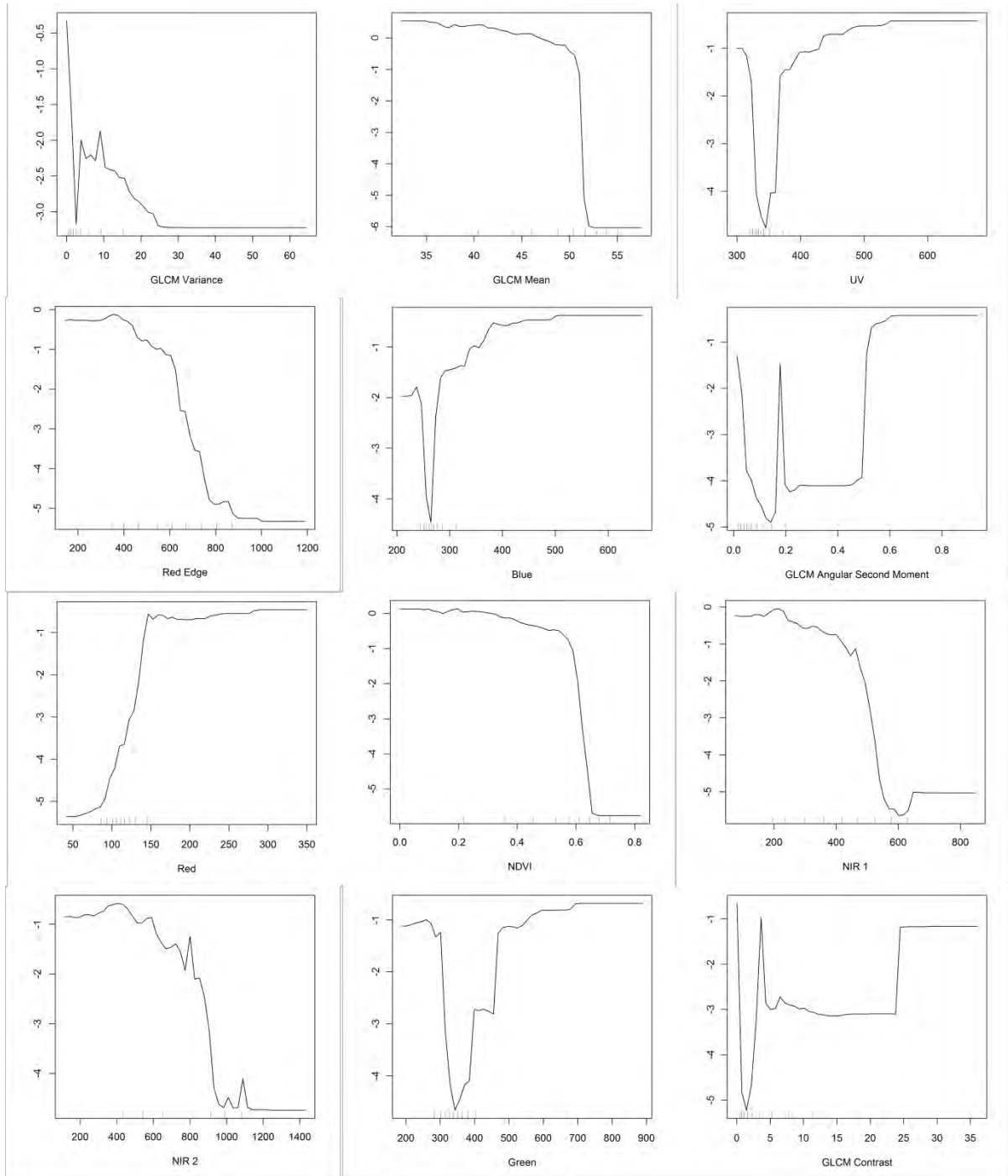
Fig. 9: Partial dependence plots for the multiple pixel-based classification of Azorella *presence and absence based on the spectral variables. The y-axis shows a relative measure of the marginal effect on the probability of* Azorella *presence.* Azorella *presence is associated with high values for GLCM Angular Second Momentum, blue and red reflectance; with low values for GLCM variance, GLCM mean, NDVI, red edge, NIR1, and NIR2 reflectance; and with mixed values for GLCM contrast, and green and UV reflectance.*
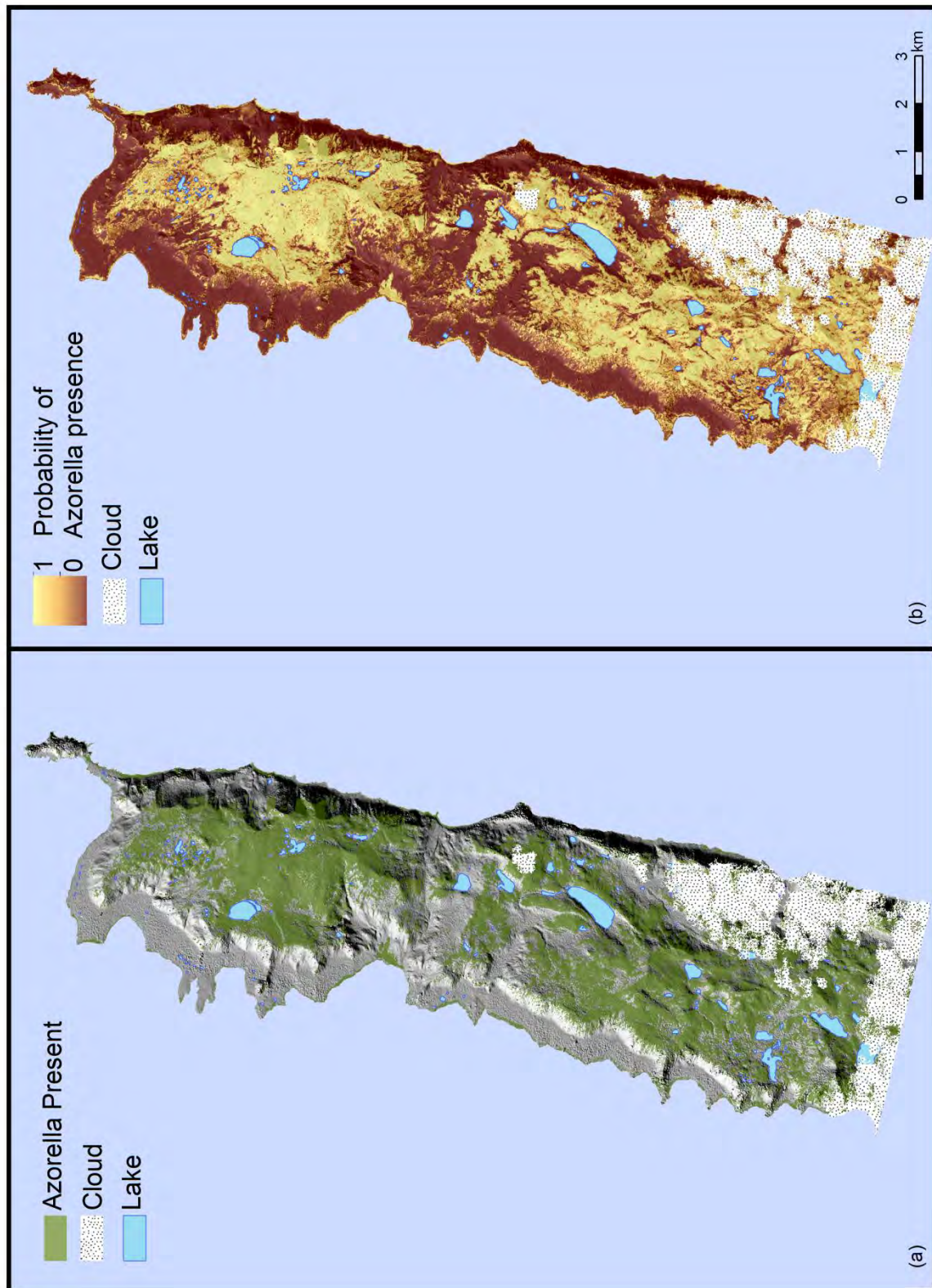
Fig. 10: Spectral-based prediction of Azorella *presence, using the multiple pixel training method.*
Panel (A) shows the hard class for Azorella *presence and panel (B) shows the probability of* Azorella
presence. The predicted presenceclass includes small patches of beach and rocky headlands, that are
known to not contain Azorella*.*

138

5.4.1.4. Three-class hybrid classification

The terrain and spectral classifications appeared to predict the presence of *Azorella* principally by mapping indirect variables that correlate with the presence of *Azorella*, rather than direct observations of the *Azorella* itself. This is inherent to SDMs like the terrain modelling presented here. It also occurred in the spectral-based classifications, in which it became apparent that the classification was principally mapping areas with sparse or absent vegetation. One of the major purposes for mapping *Azorella* cover on Macquarie Island was to provide a baseline for change detection, and if current trends continue, it is more likely that *Azorella* distribution will shrink rather than expand. Thus, maps of bare ground in potential *Azorella* habitat would be of limited use for change detection.

To explore whether the RF classification could reliably separate sparse and moderate cover of *Azorella* from bare ground, we divided the sites into three classes and repeated the hybrid classification.

Three class classifications resulted in very unequal class sizes. *Azorella* was absent from 136 of the field sites, the sparse class (< 5% cover) was found at another 44 sites, and the moderate class (> 5% cover) occurred at 21 sites. The low number of observations in the moderate cover class resulted in extremely low accuracy for these classes when using the single-pixel or object-based classifications (Fig. 11 and see Appendix B for full results). OOB accuracy estimates for classifications based on multiple pixels per site remained high (>89.6% accuracy for the present class, regardless of input variables).
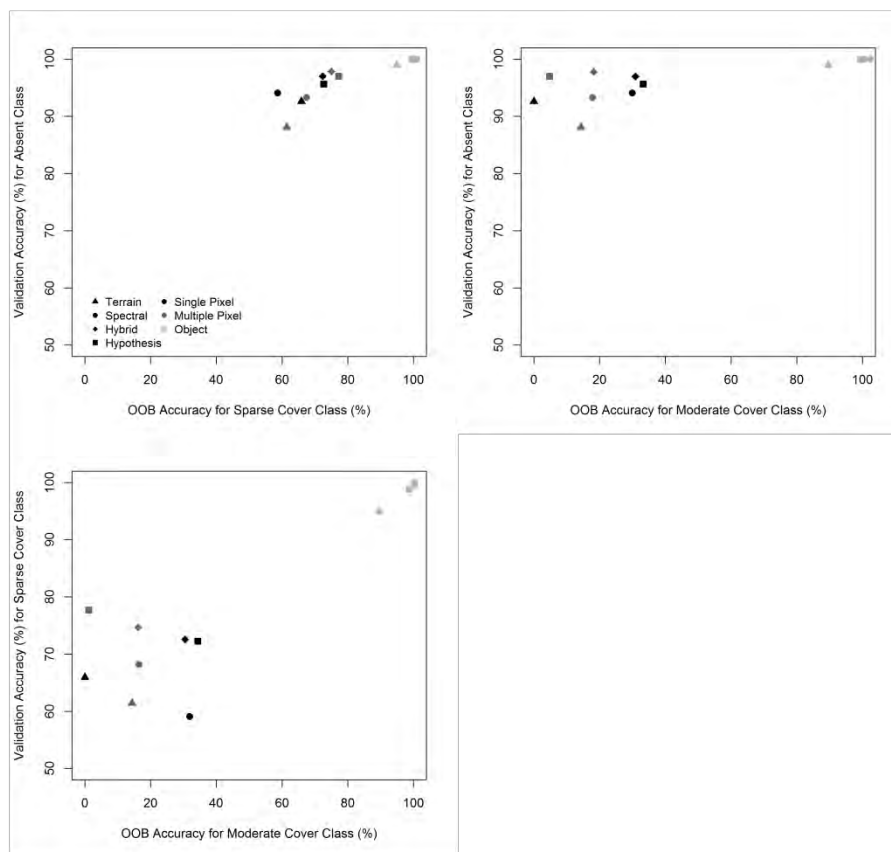
Fig. 11: Comparison of OOB accuracy estimates for three-class classifications of Azorella *cover on northern Macquarie Island. All image training methods and sets of variables perform well for the absent class, reasonably for the sparse class, and with the worst performance in the moderate cover class, which had the smallest training data set.*

As with the presence-absence classifications, the OOB accuracy assessment appeared to over-

estimate the accuracy of the classification based on a hybrid set of input variables, with only two

pixels misclassified (Table 5).

Table 5: Confusion matrix for the multiple pixel classification of the hybrid set of variables. The extremely high accuracy values are largely a function of selecting multiple adjacent pixels from each field plot, resulting in severe spatial autocorrelation.

		Predicted			
		Absent	Present	Sparse	Accuracy
Observed	**Absent**	3436	0	0	100.0
	Present	0	528	0	100.0
	Sparse	2	0	1130	99.8
	Accuracy	99.9	100.0	100.0	100.0

The independent validation dataset did not include estimates of *Azorella* cover that could be used to separate the sparse and moderate cover classes, but it could be used to identify sites that were falsely predicted to contain no *Azorella*. Using this more robust accuracy measure, 84.1% of the independent validation sites contained some amount of *Azorella*. Inspection of the maps showed that these sites were on the fringes of the predicted sparse cover class, in areas where the classes were heavily fragmented. Inspection of the map (Fig. 12) showed that the moderate cover class was largely restricted to east-facing slopes and sheltered depressions on the plateau. The solid band of moderate *Azorella* cover running down the centre of the bottom third of the predictive map appeared to be at least partially artefactual, largely influenced by distance from coastline. Although there is a large amount of moderate cover *Azorella* in this area, field observations show that it does not form a solid area of the form seen in the map. The sparse cover class included both extremely exposed west-facing slopes and areas where the feldmark intergrades with surrounding short grasslands. This corresponded with field observations.

The input variables chosen based on their high variable importance measures for the reduced hybrid classification were: elevation, distance from the coast, GLCM mean, topographically-deflected wind speed, aspect, ridgeness, wetness index, and NDVI. The partial dependence plots for the moderate cover class (Fig. 13) showed that this class was most likely at altitudes above 250 m, more than 1500 m from the coast, at higher wind speeds, at NDVI values below 0.6, GLCM mean values less than 50, and wetness index scores below seven. The relationships with aspect and ridgeness were unclear. This indicates that higher *Azorella* cover was associated with the highest, driest, most exposed sites with sparse vegetation.
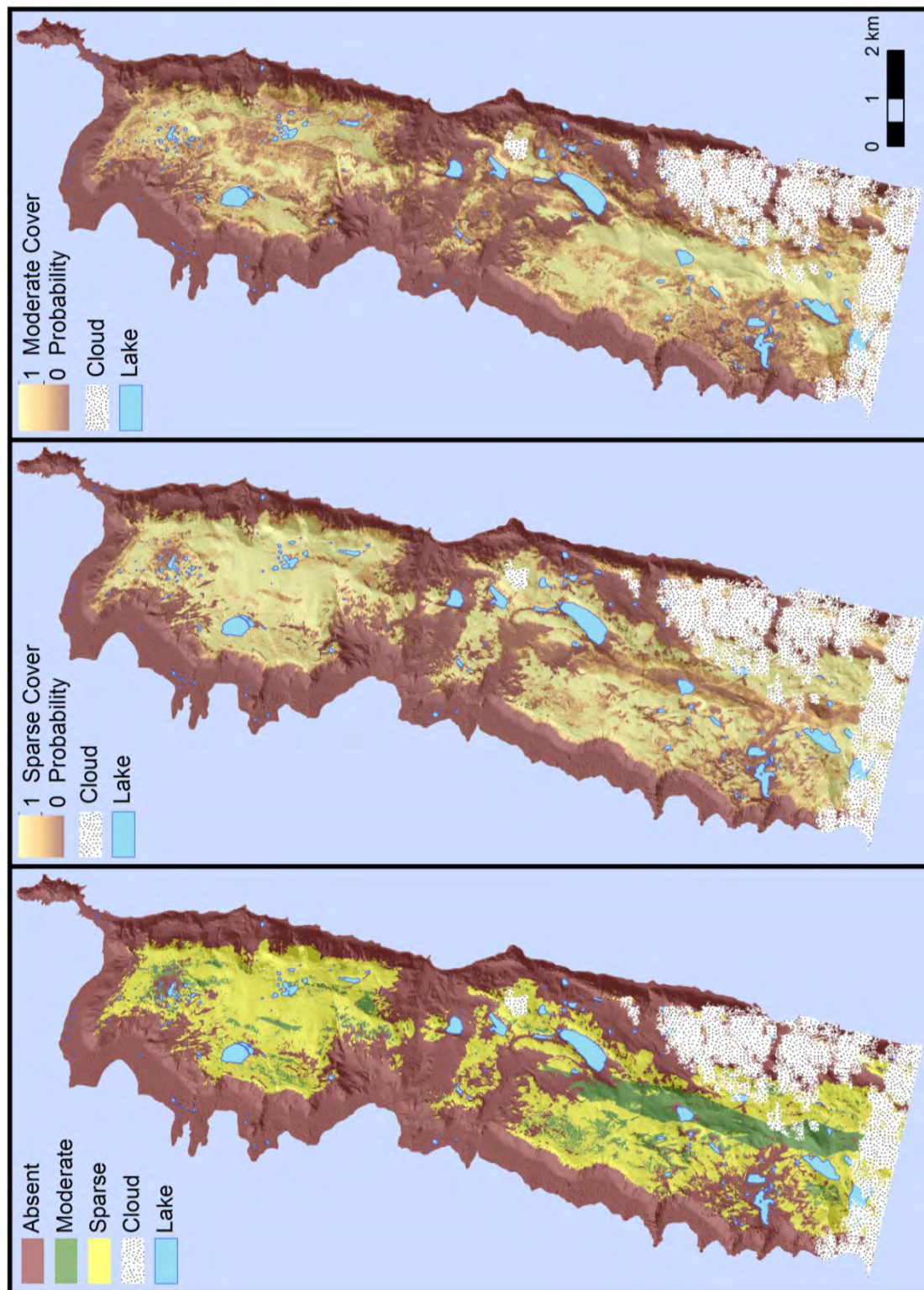
Fig. 12: Three-class Azorella *distribution map, created using a combination of terrain and spectral variables selected according to the variable importance measures. The sparse class occurs on the highest and most exposed western-facing sites, and the moderate class occurs on east-facing flanks of the mountains, and in protected hollows, in line with current understanding of the species' ecology.*
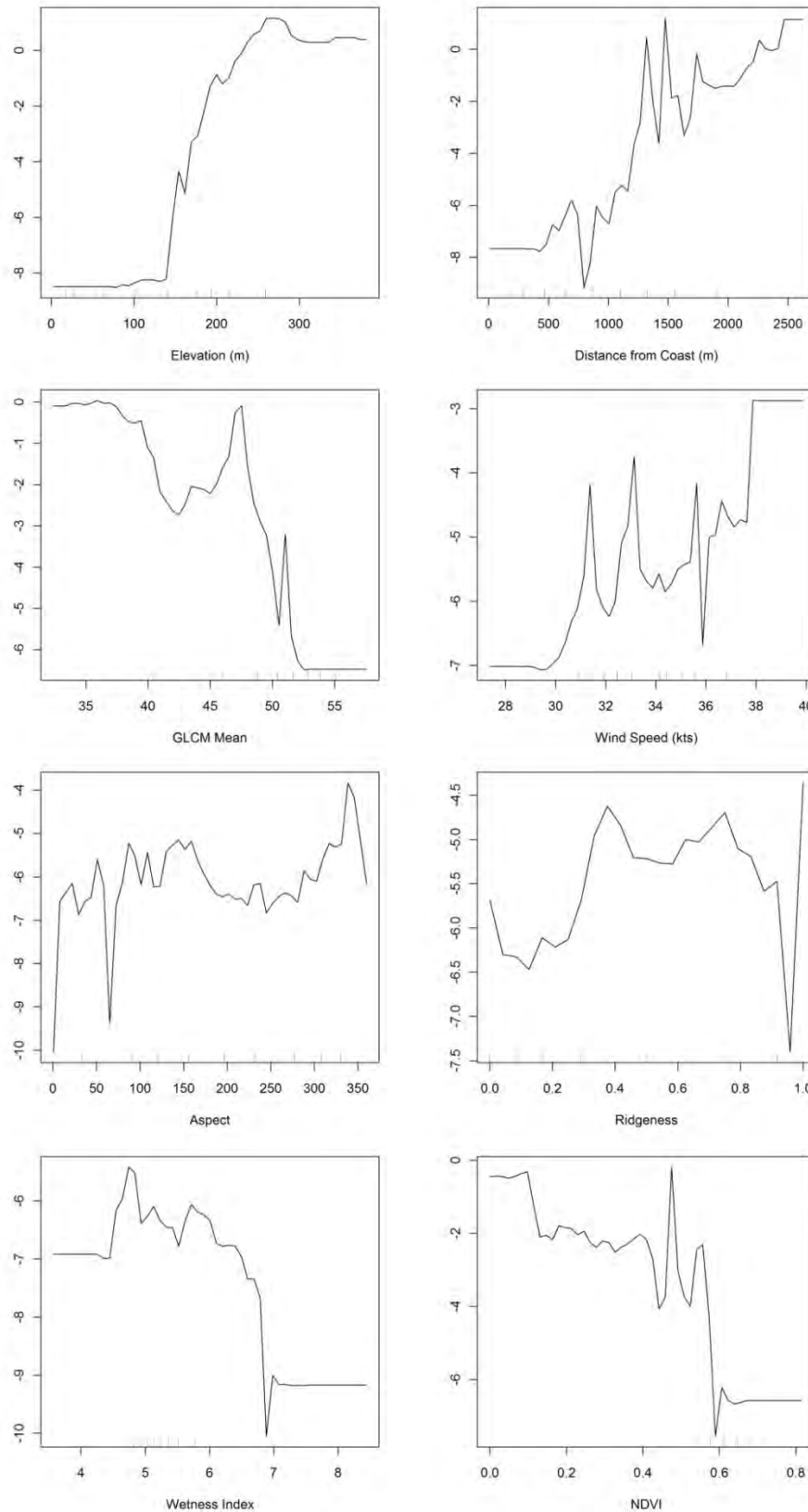
Fig. 13: Partial dependence plots for the moderate cover class of the multiple pixel-based classification of Azorella *cover based on a hybrid set of input variables. The moderate cover class was associated with high values for elevation, distance from the coast, and wind speed; with low values for the GLCM mean, wetness index, and NDVI; and with mixed values for aspect and ridgeness.*

5.4.1.5.   Hypothesis-driven 3-class classification

Allowing variable importance and OOB measures to guide the choice of input variables excluded

several variables that might *a priori* be expected to play a role in separating *Azorella* from bare

ground. Using the hypothesis-driven subset of input variables, rather than statistically-chosen input

variables, resulted in a three-class map with higher independent validation accuracy. The

hypothesis-driven three-class map correctly predicted *Azorella* presence in 89% of the independent

validation sites, and again produced extremely high OOB accuracy scores (Table 6).

Table 6: Confusion matrix for the multiple pixel classification of the hypothesis-driven set of variables. As with the other multiple pixel-trained classifications, spatial autocorrelation in the training pixels led to an inflated accuracy assessment.

		Predicted			
		Absent	Present	Sparse	Accuracy
Observed	Absent	3431	0	5	99.9
	Present	0	525	3	99.4
	Sparse	3	2	1127	99.6
	Accuracy	99.9	99.6	99.3	99.7

Inspection of the resulting map showed that areas of moderate *Azorella* coverage predicted with this

classification were more fragmented than in the statistically chosen set of inputs, though the

moderate class was again more likely to occur on the eastern flanks of the island's peaks and in

depressions on the highest parts of the island (Fig. 14). The predicted moderate cover class was

slightly more extensive than in the statistically-driven hybrid 3-class classification (731 ha, compared

with 692 ha). The solid band of moderate cover class that was predicted by the hybrid classification

as much more fragmented in this hypothesis-driven classification. The predictions of this

classification more closely correspond with field observations and inspection of landscape

photographs.
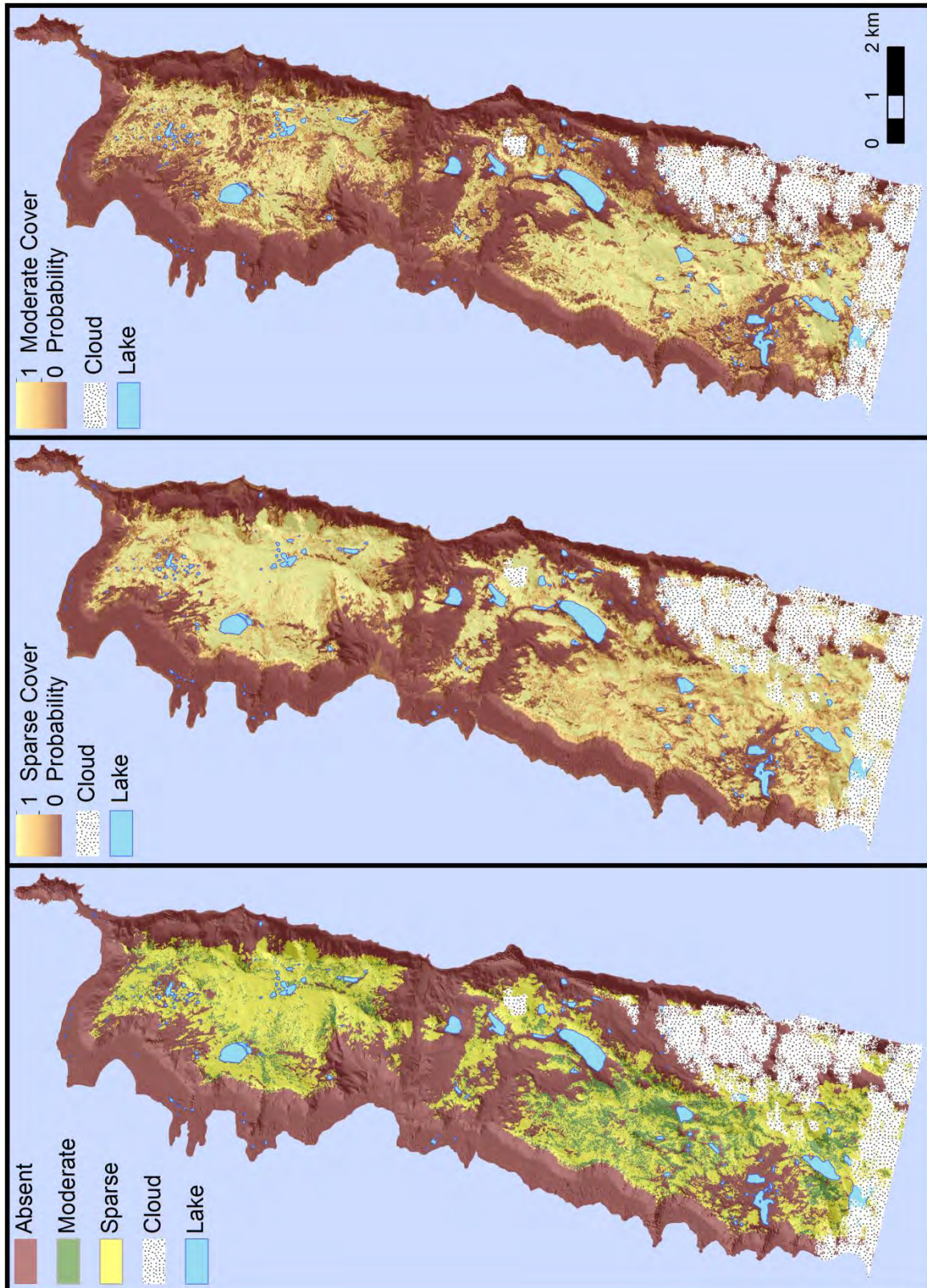
Fig. 14: Three-class Azorella *distribution map, created using a hypothesis-driven subset of input variables. The predicted moderate cover class is more fragmented and less 'blocky' than that produced by the statistically-driven hybrid classification. This resulted in a higher validation accuracy and more realistic appearance to the map.*

5.5.   Discussion

This study explored the effectiveness of several emerging tools in the field of vegetation mapping to produce high-resolution maps of the distribution of an endangered sub-Antarctic cushion plant. Of the three training methods tested here, extracting multiple pixels per field plot resulted in the highest validation accuracies. Validation of the present class indicated that the object-based training method was generally the least accurate. This is probably because the image segmentation acts as a smoothing or averaging function, and hence captures less spectral variability. For all three training methods, the hybrid and hypothesis-driven classifications were most accurate, while the spectral-only classification was least accurate. In contrast to this general pattern, the classifications based on multiple-pixel sampling were all equally accurate. Overall, the statistically-derived and hypothesis-driven sets of input variables were generally most accurate and the multiple pixel training method independently improved the accuracy of the models. For the two-class classifications, there was no way to separate the performance of the statistically-driven hybrid classifications from those based on the hypothesis-driven subset of input variables, but the latter performed slightly better in the three-class classification.

In summary, all two-class classifications were able to identify sites that contain *Azorella* with very high accuracy, regardless of the training method or the set of input variables, but the best performance was associated with multiple pixel training method and the hybrid sets of input variables that included both terrain and spectral information. Only multiple pixel training provided high accuracies for the three-class classification, and in this case, the best performance was found from the hypothesis-driven subset of variables.

In addition to the improved classification accuracy of the hypothesis- and statistically-driven hybrid classifications, there are theoretical reasons to favour them. Classifications based purely on spectral data exclude ecological data about the potential habitat for species, while those based purely on

environmental (i.e. terrain, climatic or geological) data capture potential rather than realised habitat (Dobrowski et al. 2008).

Using the variable importance measures in random forest classification to reduce the variables incorporated in the final classification often improved the OOB accuracy estimate and aided in ecological interpretation of the classification models. The OOB accuracy estimate showed that the statistically-driven classifications performed slightly better than those based on the hypothesis-driven subset of variables. In contrast, independent validation showed that the hypothesis-driven classification performed slightly better than the statistically-driven hybrid classifications.

Other researchers have found that OOB was a conservative measure of the accuracy of random forest classification (Gislason et al. 2006). Here, we found disagreements between OOB and validation accuracies of up to 13.7%, with disagreements in both directions. We are hence cautious about relying on OOB as a sole measure of model performance. The independent validation is a more robust measure of the classifications, but here we were only able to use it to assess the prediction accuracy of the present class.

One of the limitations of RF in comparison to regression-based classification methods is that it does not produce an equation with slope and intercept coefficients that can be used for direct ecological interpretation. Ecological interpretation of RF classification is, therefore, limited in comparison to statistical regression models (Cutler et al. 2007). For studies such as this one, where the primary task is classification, ecological interpretation is of secondary importance. The variable importance measures can be combined with visualisation tools to provide basic post-hoc ecological interpretation, though such interpretation must be considered hypothesis-generating rather than hypothesis testing. The ecology of *Azorella* is already quite well described (e.g. Selkirk et al. 1990), but the spatial rules revealed in the partial dependence plots from these classifications go some way to quantifying the existing descriptions of it as a species that typically grows in feldmark and intergrades into high-altitude short grasslands.

RF classification proved to be a useful tool for mapping the distribution of *Azorella*, despite a comparatively small sample size of 201 field sites; collinear input variables that made weak contributions to the classification; complex and non-linear interactions between the input variables; and noisy, non-normally distributed data. The accuracies of the classifications in this study are on a par with those found in other image classification mapping applications using a range of input data (e.g. Gislason et al. 2006; Cutler et al. 2007; Evans and Cushman 2009). Mapping projects have differing goals, extents, spatial and spectral resolutions, and accuracy measures, so comparisons with other studies should be done cautiously (Dobrowski et al. 2008).

This study highlighted some major challenges in mapping vegetation in high-latitude areas. Spatial data on environmental variables that are likely to affect plant habitat, like soil composition, are often missing. On Macquarie Island, the presence or absence of peat is likely to be a significant predictor of the distribution of *Azorella*, but such data were not available. Furthermore, field data were gathered from 349 sites over two summers, of which just 201 sites were within the cloud-free area of the available WorldView-2 image. In attempting to collect contemporaneous field data and satellite imagery, a gamble must be made that the imagery will cover the same area as the field data. In the sub-Antarctic, cloud cover often interferes with image acquisition, increasing the risk that field sites will be obscured. In this study, the exclusion of 42% of field sites left us with insufficient data to divide into training and test datasets. The independent validation data from another vegetation study provided some remedy for this, but this approach is not ideal because it could not be divided into absent, sparse and moderate classes to match the training data.

This study is part of a small but growing body of work using semi-automated image interpretation to monitor vegetation in the sub-Antarctic. Other projects have focussed on plant communities rather than single species (Lucieer 2008; Murray et al. 2010). Mostly, remotely sensed mapping of high latitude vegetation has mapped broad vegetation types or vegetation indices to capture general vegetation patterns (Stow et al. 2004; Fretwell et al. 2011). The methods presented here

demonstrate that semi-automated classification of satellite imagery can be used to map individual species in sub-Antarctic tundra.

5.6.   Conclusions

Despite the limitations on the available environmental data and the reduced number of field sites that could be used to train the classifications, they provided highly accurate maps of *Azorella* distribution, in either two or three classes. There are three characteristics of *Azorella* that may make it better suited to mapping from satellite imagery than other species: *Azorella* has a clearly defined ecological range on the island; it is located in sparsely vegetated areas, which have a distinct spectral response; and it is often the dominant species in a structurally simple vegetation type. Further, this study has demonstrated that the OOB measure in the *randomForests* package is inappropriate for assessing the accuracy of multiple pixel-trained image classifications, and alternative cross-validation measures should be used.

This study has demonstrated that a critically endangered plant species can be reliably mapped using random forest classification to combine very high resolution satellite imagery and terrain modelling. It has shown that the multiple pixel training method and hypothesis-driven subsets of input variables resulted in the most accurate maps. These maps provide a reliable baseline for monitoring expected changes in the distribution of *Azorella* on Macquarie Island. Accurate monitoring of these changes may in turn help improve understanding of the causes of the die-off.

6. Mapping multiple plant species

6.1.   Introduction

Chapter 5 introduced and tested a suite of techniques to map a single plant species on Macquarie Island, by classifying very high resolution spectral and terrain data. This resulted in highly accurate maps, especially when the random forest (RF) classifier was trained on multiple pixels per field plot and a hypothesis-driven subset of the available input variables. Identifying the distribution of a single species is computationally simpler than multi-class classification (Boyd et al. 2006) and was hence appropriate for testing analysis methods, but the ability of these tools to accurately distinguish multiple classes was not tested. This chapter presents preliminary results for testing whether these methods could be applied to a multi-class classification of the vegetation on Macquarie Island.

Earlier in this thesis, it was established that stable communities, suitable for change detection mapping, could not be defined (see chapter 4), so the classification in this chapter is based on the distributions of individual species. Mapping individual species rather than exclusive vegetation classes requires the capacity to produce overlapping classes to capture the intergraded distributions of those species. RF classification is suitable for this task because it produces measures of the probability of membership for each class (Liaw and Wiener 2002). In contrast with the maximum likelihood classification used to determine which hard class should apply at a given location, the presence or absence of the overlapping classes can be optimised by choosing a likelihood threshold that maximises the accuracy of that class.

6.2.   Methods

6.2.1. **Datasets**

Two WorldView-2 satellite images covering parts of Macquarie Island were available. An image

covering the northern half of the island was acquired on December 26 2009, and was used for the

analysis in chapter 5. A second image, acquired on January 9 2011, covered the southern half of the

island and was almost entirely cloud-free. Fast Line-of-sight Atmospheric Analysis of Spectral

Hypercubes (FLAASH) atmospheric correction (Cooley et al. 2002), using the sub-arctic summer

atmospheric model and maritime aerosol model, was applied to both images before they were

mosaicked to a single image for analysis using ENVI 4.7 software. The cloud-free portion of the

mosaicked image covered 91.2% of the island. The cloud-obscured area and a narrow strip between

the two images were masked out of the analysis. The spectral derivatives introduced in chapter 5,

normalised difference vegetation index (NDVI) and grey level co-occurrence matrix (GLCM) texture

measures, were calculated for the mosaicked image and the 5 m resolution terrain data used in

earlier chapters were also incorporated into this analysis. The visible portion of the mosaicked image

coincided with 312 of the 349 field sites visited during the summers of 2008/09 and 2009/10 (see

Chapter 3 for details of the field data collection), and an assumption was made for this exercise that

no substantial change in vegetation occurred between the two seasons.

6.2.2. **Training the classifier**

The ground cover classes chosen for the classification were those that formed large proportions of

canopy at many sites. These classes were *Agrostis magellanica*, *Luzula crinita*, bryophytes, *Azorella*

macquariensis, *Acaena* spp., *Pleurophyllum hookeri*, tall vegetation (comprising *Poa foliosa* and

Stilbocarpa polaris), bare ground, and water. The last two classes were sampled by manually

selecting points on the image that contained no vegetation, and generating 10 x 10 m squares

around those points, using Geospatial Modelling Environment software (Beyer 2009-12). The bare

ground was divided into dark, medium and light-coloured ground. The two large-structured species,

P. foliosa and *S. polaris*, were combined into a single tall class because they are structurally and ecologically similar and there were too few *S. polaris* sites to adequately train a separate class. *Stilbocarpa polaris* cover has been heavily grazed by rabbits in recent years, such that few sites retain dense canopies.

To train the classifiers, we selected sites with a high percentage of a single cover class of interest, avoiding sites with more than 25% cover of more than one of the classes, to exclude mixed pixels. The training set was made up of 71 field plots and 42 digitised bare and water plots, leaving 241 field plots for use as validation sites.

For each site, the layer values for all pixels that intersected the field plot were extracted using the multiple pixel training method presented in chapter 5. RF classification was applied to the hypothesis-driven subset of variables used for mapping *Azorella* in chapter 5: *i.e*. the blue, green, yellow, red edge and near-infrared 2 spectral bands; NDVI; the mean, homogeneity and entropy GLCM texture measures; and the elevation, slope, solar radiation, curvature, ridgeness, valleyness, and wetness index terrain variables. For the RF, the number of trees was set at 2000, after exploratory analysis showed that classification errors stabilised at 0.002 at approximately 1000 trees.

The default output of RF classification is a series of hard classes that do not overlap, with each observation assigned to the cover class with the highest probability. The values for the probability of membership for each class can be interrogated to find an optimal probability-threshold that will produce the maximum accuracy for each class, regardless of overlaps with other classes. We adapted the approach taken by Grossmann et al. (2010) to adjust the cut-off boundaries for the presence of each class. Here, the probability of class membership was plotted against the presence or absence of that species of interest for all validation sites, to find a probability-threshold that maximised the separation between true presences and absences.

6.2.3. **Testing the classification accuracy**

Once the forest had been trained, the RF was applied to the mosaicked image. Several accuracy measures were used to test the accuracy of the classification. First, a confusion matrix was used to explore the relationships between species presences in the validation plots and the hard classes. In addition to dominating the canopy at some sites, many plant species on Macquarie Island grow as isolated individuals in most environments across the island, so they were deemed to be present in the validation sites when there was more than 10% cover. The exception to this was *Azorella*, which has a typically sparse growth habit. For this class, it was deemed to be present at a site whenever it occurred in a site. We calculated both the producer's accuracy, which assesses the sensitivity of the classification, or errors of omission; and the user's accuracy, which assesses the specificity of the classification, or errors of inclusion (Foody 2010).

As many of the species under examination grow in mixed stands and ecotones, there was likely to be high levels of uncertainty in the classification. This was quantified using standardised entropy after the approach of Fisher et al. (2006).

$$E = \frac{-\sum_{n=1}^{i} P(Cn) \cdot \log[P(Cn)]}{-\log\left(\frac{1}{i}\right)}$$

Where *E* is the entropy of a pixel, and ranges between 0 and 1; *n* is the number of classes; and P(*Cn*) is the probability of membership to class *n*. Low entropy indicates an unambiguous classification, while high values indicate confusion and high classification uncertainty.

To further examine the relationship between the probability of class membership and species cover in the validation plots, Spearman rank correlation coefficients were calculated.

6.3.   Results and Discussion

6.3.1. Hard Classes

Multi-class random forest classification produced variable accuracies. For the spectrally and ecologically distinct *Pleurophyllum hookeri*, both the producer's and user's accuracies were very high. In contrast, the *Agrostis*, *Azorella*, bryophytes, *Luzula* and bare ground classes had very high user's accuracies, but low producer's accuracies. The hard classes identify the core areas for these species, but exclude the ecotones where they intermingle. The tall vegetation class had a very high producer's accuracy (91.3%) but strongly over-estimated the true distribution of these two plant species (Fig. 1 and Table 1). Overall, the multi-class classification presented here had much lower accuracies than the binary classifications presented in chapter 5, though it varied by class.

Macquarie Island Dominant Species (Hard Classes)

Acaena monostand

Dense *Azorella* dominated feldmark

Agrostis grassland

Bare ground
with bryophytes in background

Bryophyte dominated grassland

Luzula dominated grassland

Pleurophyllum dominated grassland

Tall Vegetation

Stilbocarpa monostand

Poa foliosa tussock grassland

0 1 2 3 4
km

	Unclassified
	Acaena
	Agrostis
	Azorella
	Bryophytes
	Dark Bare
	Lake
	Light Bare
	Luzula
	Mid Bare
	Pleurophyllum
	Tall Vegetation
	No Data

Fig. 1: Map of hard classes of dominant species' distributions on Macquarie Island.

155

6.3.1.1. *Acaena*

The *Acaena* class (958.1 ha) was concentrated at the base of the coastal escarpments, in low – mid

altitude bowls, and along coastal streams. These are areas where water and nutrients tend to

accumulate, and below petrel nesting localities. Of the validation sites classed as *Acaena*, 58%

contained >10% *Acaena* cover. A similar proportion of sites contained significant amounts of

Agrostis and bryophytes. This class contained 50% of the validation sites that had significant cover of

Acaena. There was no correlation (ρ = 0.03) between observed *Acaena* cover and the probability of a

plot being assigned to this class (Table 2). Thus, this class simultaneously under- and over-estimated

the true distribution of *Acaena* spp. The probability map showed that this class had intermediate

probability in all vegetated areas, with the highest values on coastal slopes and flats (Fig. 2).

6.3.1.2. *Agrostis*

The *Agrostis* class (2421.3 ha) mostly covered patches of the coastal slopes and the lower parts of

the plateau. Of the validation sites in this class, 88% had significant *Agrostis* cover, and a similar

proportion of the sites contained significant bryophyte cover. This class contained 38% of all

validation sites with significant *Agrostis* cover. There was some correlation (ρ = 0.47) between

Agrostis cover in a site and membership of this class. *Agrostis* grows on all parts of the island and in

a wide range of growth habits; ranging from isolated individuals on beaches and in feldmark to

dense monostands in mid-altitude grasslands. This class appeared to capture the dense *Agrostis*

grasslands, but excluded the ecotones where *Agrostis* intermingled with other species. Probability

mapping revealed moderate to high probability for this class in all vegetated areas, though the

probability was somewhat lower on coastal flats.

6.3.1.3. *Azorella*

The *Azorella* class (2408.4 ha) in this map (Fig. 1) was somewhat different to that produced in

chapter 5. The classification was trained on the dense *Azorella* beds that were more common in the

southern half of the island. The *Azorella* class covered extensive areas of the southern half of the

plateau and was more fragmented in the north. This general pattern corresponded with field

observations of the distribution of *Azorella*. Of the validation sites in this class, 89.1% contained

some *Azorella*. More than 80% of these sites also contained significant cover of bryophytes and bare

ground. The *Azorella* class contained 35.5% of all validation sites with some *Azorella* cover, and there

was a correlation between the percent-cover of *Azorella* in a site and the probability of class

membership (ρ = 0.63). This class, trained as it was on dense *Azorella*, underestimated the

distribution of Azorella, especially in sparse feldmark. There were few false positives, suggesting that

the class was quite specific to *Azorella*. The probability of Azorella presence was moderate in the

sparsely vegetated high altitude parts of the island.

6.3.1.4. Bare Classes

The bare ground classes (light: 1009.7 ha; mid: 560.4 ha; and dark: 283.8 ha) were fragmented

across the beaches and highest peaks. All the validation sites in the three bare classes contained

significant bare ground. These classes, however, represented just 9.2% of the validation sites with

significant bare ground. There was reasonable correlation between the cover of bare ground in a site

and the probability of membership to one of these classes (ρ = 0.43 – 0.69). These classes thus

severely underestimated the total area of bare ground on the island. Probability maps showed that

the three bare ground classes largely targeted separate parts of the non-vegetated areas. The light

bare ground was focussed on the highest peaks; the dark bare ground mostly occurred on steep

slopes and in penguin colonies; and the mid-shaded bare ground targeted intermediate areas across

the plateau.

6.3.1.5. Bryophytes

The bryophyte class (2494.9 ha) typically occurred between the Azorella and bare ground classes on

the higher parts of the plateau. Of the validation sites assigned to this class, 95.2% contained

significant bryophyte cover. This class, however, underestimated the total distribution of

bryophytes, as it contained 24.9% of the validation sites with significant bryophyte cover. Extensive

bryophyte cover occurred in all the vegetated classes, but the percent-cover of bryophytes in a site

was correlated with the probability of membership (ρ = 0.44) to the bryophyte class. Thus, the

bryophyte class represented areas with dense bryophyte cover, but excluded mixed vegetation.

Inspection of the probability maps showed a moderate to high probability of bryophyte presence for all parts of the plateau, apart from lakes and the highest peaks, where vegetation was extremely sparse. Modelling the distribution of bryophytes as a single taxon on Macquarie Island is complicated by the wide variation in ecological and spectral characteristics of the 180 bryophyte species included in this taxon. It is unlikely to be practical to collect species-level data for bryophytes from a sufficiently large sample to train and test an image classification.

6.3.1.6. *Luzula*

Luzula grows in a wide range of habitats on Macquarie Island, often mixed with *Agrostis*. The *Luzula* class (126.3 ha) occurred as small patches (mean 160.8 m^2) on the coastal slopes and low altitude grasslands, often surrounded by the *Agrostis* class. Of the validation sites classed as *Luzula*, 88.9% contained significant *Luzula* cover, indicating that the class was specific to areas with high *Luzula* cover. In contrast, this class represented just 12.7% of the validation sites that contain significant amounts of *Luzula*. There was a weak correlation between *Luzula* cover and the probability of class membership ($\rho= 0.29$). Probability of presence was moderate to high for all but the steepest coastal slopes. Together, these results indicated that the hard class severely underestimated the extent of *Luzula*, particularly the areas where it intermingles with other species.

6.3.1.7. *Pleurophyllum*

The *Pleurophyllum* class (145.1 ha) was concentrated near the base of the escarpments on the west coast and on the lower parts of the northern plateau. These areas typically occurred between the *Agrostis* and *Acaena* classes. Of the validation sites assigned to this class, 88.9% contained significant amounts of *Pleurophyllum*. This class very closely matched the true distribution of *Pleurophyllum*, representing 88.9% of all sites with significant *Pleurophyllum* cover. Thus, this class was both sensitive and specific to the distribution of the species, despite there being no correlation between the percent-cover and probability of class membership ($\rho = 0.09$). It is likely the lack of correlation was due to the small number of plots with moderate cover of *Pleurophyllum* in the validation

dataset, and this should be investigated further. The probability maps showed moderate probability

of *Pleurophyllum* cover on coastal slopes, especially on the western side of the island.

6.3.1.8. Tall Vegetation

The tall class (1171 ha) covered almost all of the coastal flats between a thin strip of beach and the

Acaena class that characterised the base of the escarpments. Of the validation sites in this class,

44.7% were observed to contain significant cover of at least one of the two tall species (*P. foliosa*

and *S. polaris*). Significant cover of bare ground (68%) was more common and bryophytes (44.7%)

were as likely as the tall species to grow in these sites. This class contained 91.3% of all validation

sites with significant cover of tall vegetation. Thus, this class significantly over-estimated the

distribution of the two tall plant species. Most obviously, the featherbed mire that characterises the

north-west coastal shelf was classified as containing tall vegetation. This area had a high probability

of belonging to this class. In reality, however, this area contained small patches of tall vegetation

interspersed with bryophyte- and sedge-dominated mires. The mires contain an extremely complex

mixture of species and were not well captured by a classifier trained on plots with high cover of a

single species. Training a separate mire class would likely improve the accuracy of the tall vegetation

class.

6.3.1.9. Standardised Entropy

The classifications in highly vegetated areas were uncertain (Fig. 2). The standardised entropy layer

showed low uncertainty for the lake and bare ground classes. The *Azorella* class also had low

entropy, especially in the highest parts of the southern plateau where the class formed large

homogeneous patches. Uncertainty was highest on the steep coastal slopes and the low-altitude

grasslands, where multiple species grow in complex mixtures. Thus, it was likely that the areas of

high entropy in the classification reflected real world complexity, rather than poor classification

results.

1   *Table 1: Confusion matrix of the hard map classes and validation sites. The user's accuracy (sensitivity) was high for all classes except the tall vegetation. The*
2   *producer's accuracy (specificity) was much lower for most classes.*

| | | Map Class | | | | | | | | | | | | |
		Acaena	*Agrostis*	*Azorella*	Light Bare	Mid Bare	Dark Bare	Lake	Bryo-phyte	*Luzula*	*Pleuro-phyllum*	Tall	Total	Accuracy (%)
Validation	*Acaena* **> 10%**	22	13	3	0	0	1	0	4	0	0	1	44	50.0
	Agrostis **> 10%**	23	73	19	1	0	2	0	45	9	4	16	192	38.0
	Azorella **> 0%**	2	12	41	4	6	1	1	46	1	0	2	116	35.3
	Bare Ground > 10%	13	32	38	6	6	5	1	44	4	3	32	184	9.2
	Bryophytes > 10%	22	73	39	3	1	4	1	59	7	7	21	237	24.9
	Luzula **> 10%**	8	26	1	0	0	1	0	8	8	1	10	63	12.7
	Pleurophyllum **> 10%**	0	1	0	0	0	0	0	0	0	8	0	9	88.9
	Tall Vegetation > 10%	1	1	0	0	0	0	0	0	0	0	21	23	91.3
	Total	38	83	46	6	6	5	1	62	9	9	47		
	Accuracy (%)	57.9	88.0	89.1	100.0	100.0	100.0	NA	95.2	88.9	88.9	44.7		

3
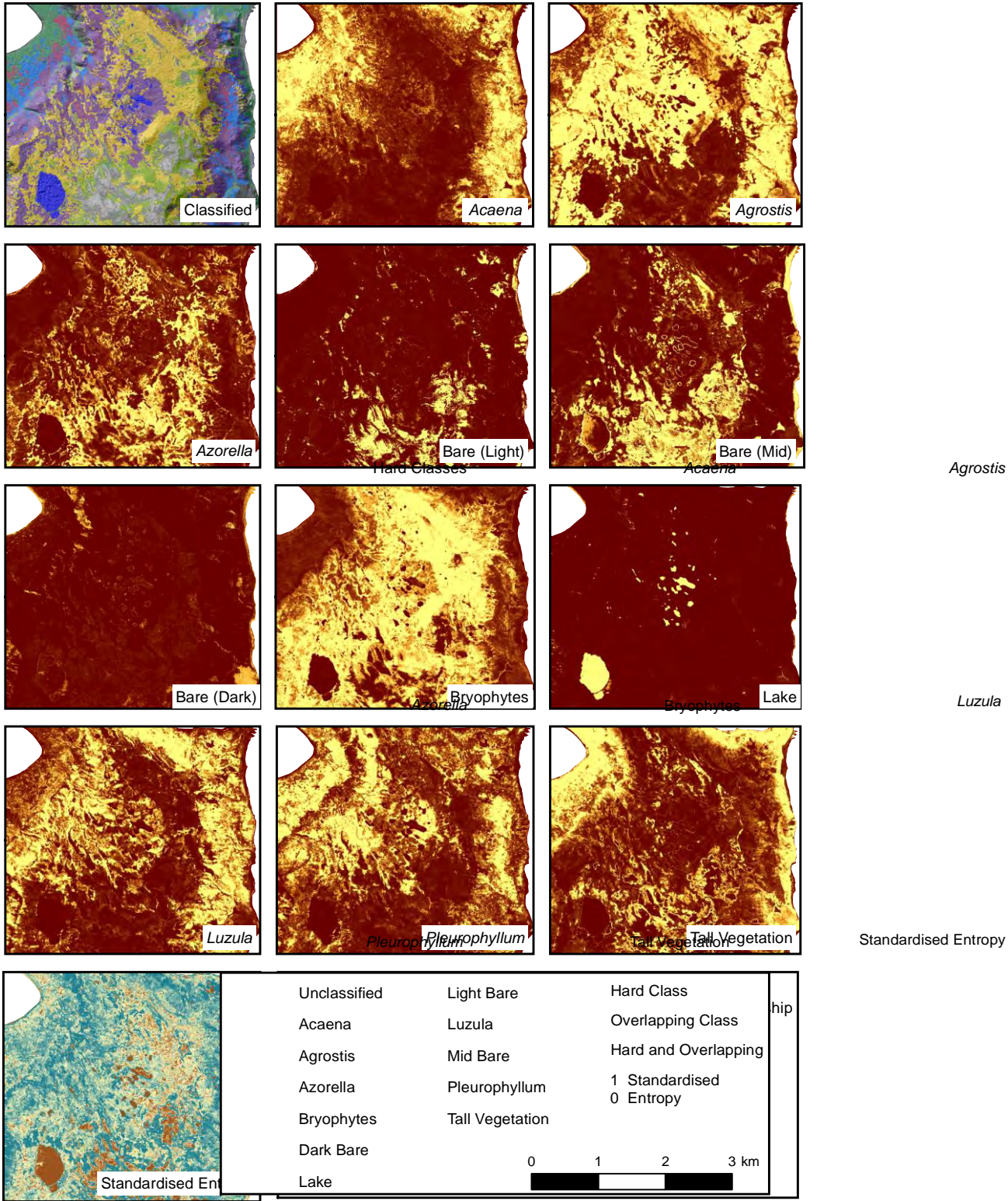
Fig. 2: Probabilities for membership of each land cover class and the standardised entropy measure of the confusion of the classification. Acaena, Agrostis, and Pleurophyllum are associated with the coastal slopes. The tall vegetation has the highest probabilities on the coastal flats and with mire areas. The bare ground and Azorella classes are most commonly associated with the highest parts of the island, though the dark bare ground class is most associated with the steep shadow-affected slopes on the west coast.

Table 2: Spearman rank correlations between the probability of class membership and the percent-cover of individual cover class at the validation field sites. The relationships between percent cover and the probability of class membership were typically weak to moderate.

		Percent Cover in Field Sites							
		Acaena	*Agrostis*	*Azorella*	**Bare**	**Bryo**	*Luzula*	*Pleuro*	**Tall**
Probability of Class Membership	*Acaena*	**0.03**	0.31	-0.61	-0.28	-0.15	0.19	-0.17	0.26
	Agrostis	0.05	**0.47**	-0.50	-0.30	0.03	0.31	-0.02	0.18
	Azorella	0.04	-0.21	**0.63**	0.40	-0.06	-0.24	0.05	-0.26
	Dark Bare	-0.27	-0.41	0.04	**0.56**	-0.38	-0.37	0.08	0.03
	Light Bare	-0.10	-0.08	-0.01	**0.43**	-0.26	-0.05	0.02	-0.02
	Mid Bare	-0.24	-0.42	0.32	**0.69**	-0.32	-0.38	0.20	-0.12
	Bryophytes	0.14	0.13	0.30	0.12	**0.44**	0.09	0.25	-0.34
	Luzula	0.04	0.43	-0.51	-0.21	0.00	**0.29**	-0.10	0.20
	Pleuro	-0.09	0.31	-0.48	-0.21	0.00	0.17	**0.09**	0.18
	Tall	-0.24	-0.17	-0.45	0.11	-0.45	-0.12	-0.13	**0.34**

6.3.2. **Overlapping Classes**

The use of hard classes implies that the classes are mutually exclusive. In the case of mapping individual species , this assumption is often unreasonable. For such cases, it is possible to produce overlapping classes by adjusting the probability threshold for class membership. Plotting the probability of class membership against the presence or absence of the species of interest in the validation plots showed that this was feasible for all vegetated classes (Fig. 3). This is similar to the approach of Grossmann et al. (2010) to tuning cut-offs for the presence of small classes.

For the tall vegetation and *Pleurophyllum* classes, the box plots showed clear breaks in the probability between validation plots where the species was present or absent, using a 10% cover threshold for determining presence for all species other than *Azorella* (Fig. 3). This break was reflected in the high accuracies for the hard class map. The box plots indicated that reducing the probability threshold for the *Acaena*, *Agrostis*, *Azorella*, bryophyte and *Luzula* classes would likely increase the accuracy of these classes, though the division is not so clear-cut as for the first two classes. These are the classes which under-estimated the distribution of widely-dispersed species.

The box plots indicated that changing the threshold for classification would have little effect on the accuracy of the bare ground classes.

Using these plots as guidance, the probability thresholds for each class were adjusted to produce overlapping classes. For the species whose distribution had been underestimated by the hard classes, the overlapping classes had lower producer's accuracy, but higher user's accuracy (Table 3). For the tall class, which had over-estimated the distribution of *P. foliosa* and S. *polaris*, the overlapping classes improved the producer's accuracy without changing the user's accuracy. The map showed that the overlapping class excluded several small patches on the plateau that the hard class map had falsely predicted would contain tall vegetation (Fig. 4).

 For the *Pleurophyllum* class, the overlapping classes performed similarly to the hard classes. The 11.1% increase in producer's accuracy meant little, as it represented a single plot due to the small number of validation plots that contained *Pleurophyllum*. Maps of the overlapping class for *Pleurophyllum* showed a slight reduction in predicted area, compared with the hard class.

*Table 3: Comparison of the accuracies of the hard-class and overlapping-class classifications. In general, the overlapping classesincreased the user's accuracy at the expense of the producer's accuracy. This was most obvious in species with broad distributions. For the classes with restricted ranges (*Pleurophyllum *and tall vegetation), overlapping classes increased the producer's accuracy, with no effect on the the user's accuracy.*

	Hard Classes		Overlapping Classes		Change	
Class	**Producer's Accuracy**	**User's Accuracy**	**Producer's Accuracy**	**User's Accuracy**	**Producer's Accuracy**	**User's Accuracy**
Acaena	57.9	50	33	72.7	-24.9	22.7
Agrostis	88	38.0	82.1	59.9	-5.9	21.9
Azorella	89.1	35.3	77	66.4	-12.1	31.1
Bryophytes	95.2	24.9	88.9	43.9	-6.3	19.0
Luzula	88.9	12.7	34.8	36.5	-54.1	23.8
Pleurophyllum	88.9	88.9	100	88.9	11.1	0.0
Tall	44.7	91.3	58.3	91.3	13.6	0.0
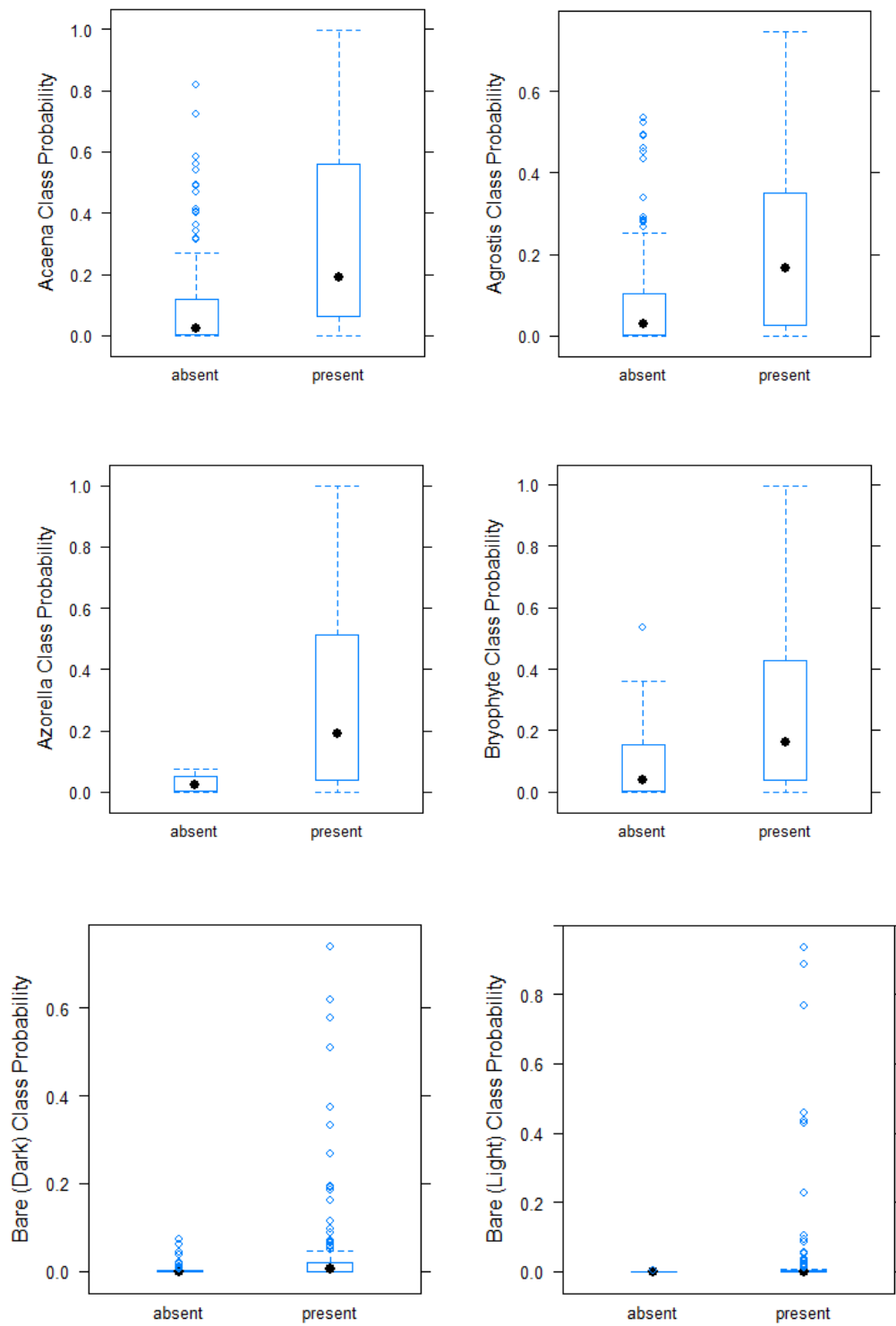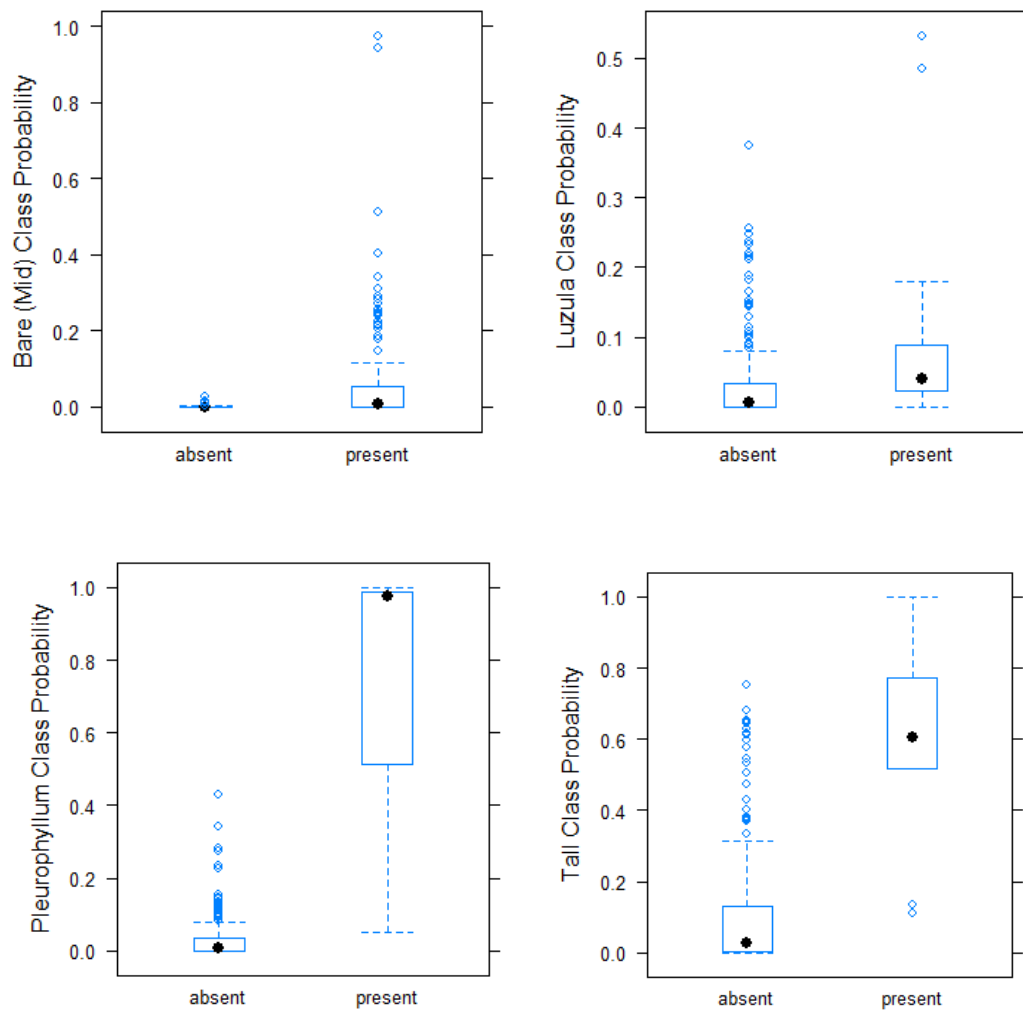
Fig. 3: Box plots showing the probability of class membership as a function of whether the species is present within validation plots. For all species apart from Azorella, *"present" means >= 10% cover. For* Azorella, *"present" means > 0% cover. There were clear differences in the class membership probabilities between the sites at which* Pleurophyllum *and tall vegetation, the most restricted species, were present and those where they were absent. For the more widespread species, there were no obvious differences in the class membership probabilities that could be used to easily choose appropriate thresholds for membership of the overlapping classes.*

The overlapping classes increased the user's accuracy for the widely dispersed species by increasing the sensitivity of the classes to areas with lower percent-cover of the target species (Fig. 4). This was done at the expense of decreasing the producer's accuracy. Thus, these maps incorporated areas of mixed vegetation into the distribution of these species, becoming more sensitive to the presence of the species and simultaneously less specific. This pattern was especially notable for *Luzula*, as this

species most commonly occurs in mixed stands. The overlap class covered 2795.3 ha, while the hard class covered just 126.3 ha. For the hard classes that performed well or over-estimated the species distributions, the overlapping classes improved the producer's accuracy without reducing the user's accuracy. Thus, these classes became more specific to the target species.
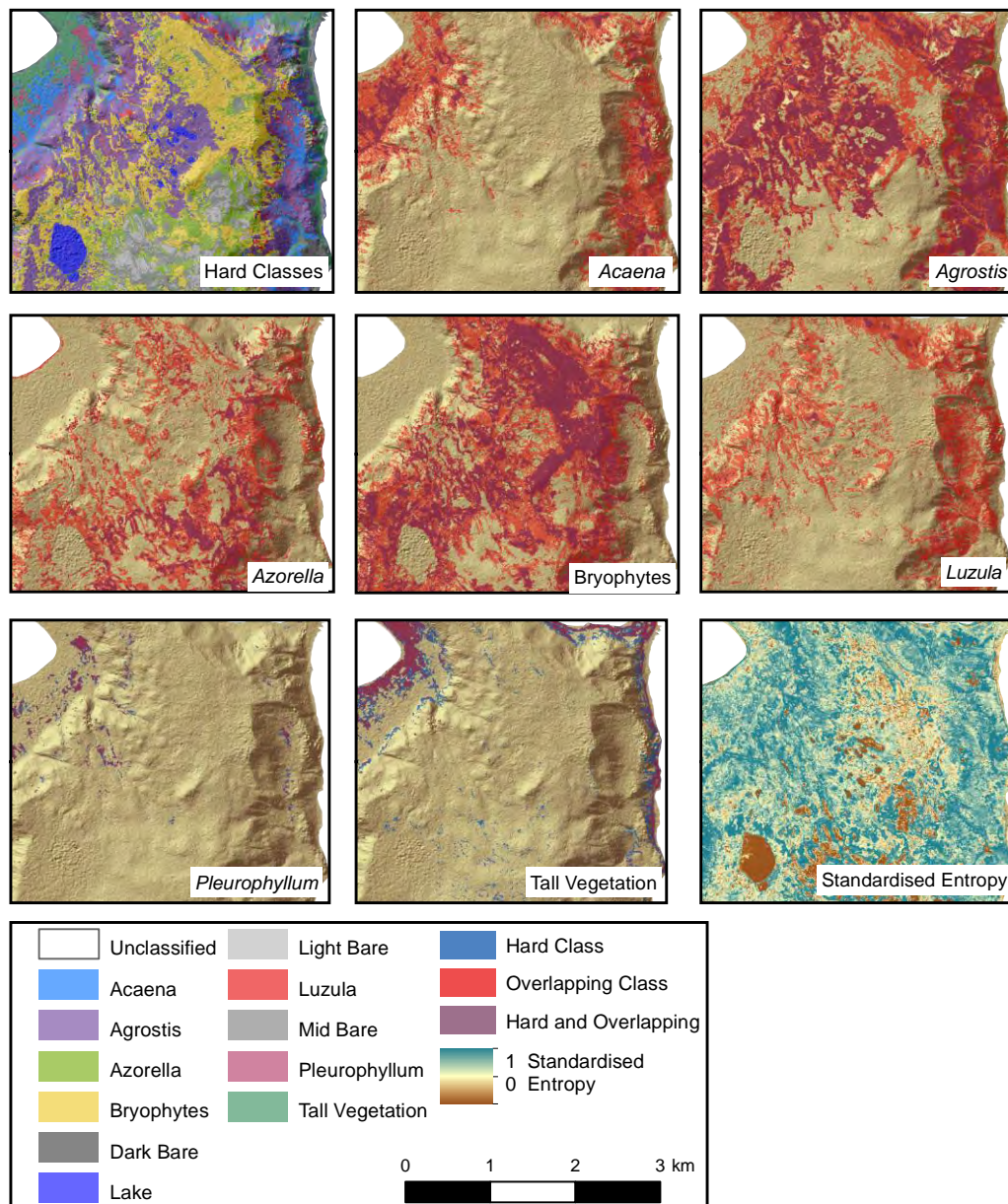
Fig. 4: Comparison of hard and overlapping classes for all vegetated classes. For the widespread species, the overlapping classes covered much larger areas than the hard classes. This relationship was reversed for those classes with restricted ranges.

6.4.   Conclusions

Multi-class mapping of species distributions on Macquarie Island identified core habitat areas for all classes except tall vegetation, which could not be distinguished from the featherbed mire. For almost all species, both the hard- and overlapping-classes had much lower accuracies than the binary classifier for *Azorella* presented in Chapter 5. On the basis of the validation, we consider that only the *Pleurophyllum* class was sufficiently accurate to be used as a baseline for change detection mapping.

There are two key potential reasons for this lower accuracy, and they indicate directions for future research opportunities. Firstly, the classifier presented here was trained on sites with the highest cover of the target species. As many of the target species grow in a range of environments across the island with varying densities, the classifier was not tuned for areas with lower densities. Thus, the effect of training samples on the resulting classification needs to be examined. Secondly, binary classifiers have been shown to be more computationally efficient (Boyd et al. 2006) in other environments, and a future study could apply a series of binary classifiers to test whether these improve the accuracy for the target classes.

The species for which this classifier was most accurate, *Pleurophyllum hookeri*, occurs as isolated individuals in a wide range of environments, but only forms a major component of the canopy in a restricted area. The species for which this classifier performed most poorly were those that grow mostly in intermixed stands across a wide range of environments.

Although the accuracies for most classes in this study were not high enough to provide reliable baselines for change detection, the maps presented here improve on the existing maps of Macquarie Island vegetation. To date, these maps have showed structural vegetation classes digitised from 30 m resolution satellite imagery (Selkirk et al. 2000) or changing patterns in NDVI (Bergstrom et al. 2009a).

This study has shown that multi-class RF classification of species distribution on Macquarie Island is appropriate for depicting the general distribution of core habitat of those species, although the accuracy is too low for the classes to act as baseline maps for change detection. Further experimentation with training samples and binary classifiers may further improve the accuracy of these maps.

7. Conclusions

In this study, approaches to mapping the vegetation of sub-Antarctic Macquarie Island have been proposed and tested for their robustness and accuracy. All stages of the mapping process have been examined, from sampling design and field techniques to the methods used for classifying satellite imagery. On this isolated and heavily disturbed island, there is particular need for efficiency in field sampling methods and for a sampling design that is adaptable to the practicalities of fieldwork in a harsh environment. Recent rapid vegetation change on the island in response to management activities and environmental change have prompted interest in accurately monitoring those changes at a range of spatial scales. This study provides tools for accurate monitoring at the plot and landscape scales.

7.1.   Field Methods

Objective 1: to establish a reliable and efficient field sampling method to capture species abundance

data for vascular plants at the required scale of image analysis

Chapter 2 demonstrated that photo-sampling can act as a valid replacement for quadrat-based field sampling of field species cover in several vegetation types on Macquarie Island. Two methods of photo-interpretation were compared to field quadrat cover estimates for their capacity to identify the species present in a plot and to estimate the canopy cover of each species. Pseudoturnover rates among the three observation methods (24.6% ± 10.5) were similar to those found in double-sampling tests of field quadrat methods elsewhere. There were no significant differences among the total cover estimates for each morphological group. Finally, multivariate clustering showed that the three observations from each plot were more similar to each other than those from other plots, so long as the clustering incorporated species-abundance rather than simply presence-absence data.

There are two key features of sub-Antarctic tundra vegetation that likely contributed to the success of this approach. Firstly, the vegetation has a simple vertical structure, with rarely more than two

overlapping strata. Secondly, the island also has a depauperate flora, which aids in the identification

of specimens in the photographs. This approach has been used in marine studies, but has rarely

been tested in terrestrial environments. This chapter therefore demonstrates the practicality and

reliability of photo-sampling for low-statured terrestrial vegetation. Photo-sampling takes less time

in the field, but the need for increased data processing means that for most terrestrial

environments, field quadrat methods are likely more time efficient overall. For environments where

field access is limited, such as sub-Antarctic islands, the increase in field efficiency may compensate

for the additional data processing requirement. The point-intercept method of photo interpretation

was the most time-efficient, but must be supported by the collection of a species list in the field, as

rare species are frequently missed by point-intercept sampling. Photo-sampling has an additional

benefit in providing a permanent digital record that can be used for plot-scale change detection

Objective 2: to establish a geographically stratified random sampling method to capture the full

variation of both the terrain and vegetation patterns

Chapter 3 described GeoStrat, a geographically stratified proportional random sampling design that

uses existing knowledge on terrain and vegetation to locate field sites. The first task in developing

GeoStrat was to classify the island into homogeneous strata on the basis of terrain and vegetation

data. Fuzzy *c*-means classification was applied to six terrain variables believed likely to affect

vegetation patterns (elevation, slope, surface curvature, wetness index, solar radiation and

topographically deflected wind speed) and a normalised difference vegetation index (NDVI) to find

seven strata. The proportion of potential sites per strata was set by ranking the strata according to

three criteria: area, variance of the NDVI within each stratum, and a subjective assessment of the

ecological variability and likelihood of access difficulty. These three rankings were combined by

weighted pairwise voting to calculate the proportions of sites for each stratum. A random sampling

tool was used to assign the locations for 380 potential field sites. Over two summers, 288 of these

sites were visited, along with 58 sites that were purposively located in perceived core areas of plant

communities. The excluded sites were either inaccessible (43 sites) or were excluded due to time constraints (53 sites). Most commonly, sites were excluded from the three coastal strata.

To test the performance of the GeoStrat and purposive samples, three other randomised samples were generated: an equal stratified random sample, an area-proportional stratified random sample, and an unstratified random sample. Each of these was resampled to simulate the access limitations experienced in the field. As a control, a very large unrestricted random sample was generated.

All the randomised sampling designs came close to matching the control in representing the variation in the terrain variables used to drive the stratification, while the purposive sample was strongly biased towards the vegetationally complex coastal regions. Compared with the other randomised sampling designs, GeoStrat exhibited slight bias towards the coastal regions, with small but significant deviations from the true distributions of elevation, solar radiation, NDVI, and surface curvature. GeoStrat therefore acts as a compromise between true random sampling and purposive sampling. In this study, GeoStrat performance was similar to equal stratification, but this was likely an artefact of the most ecologically complex and difficult to access areas occurring in the smallest strata. GeoStrat provides a tool to ensure that the sampling can be targeted at the most complex vegetation, regardless of stratum size. Finally, this chapter demonstrated that (almost) random sampling could be practically implemented despite the predicted challenges in accessing large parts of this remote and steep island.

7.2. Mapping methods

Objective 3: to establish a robust statistical method for identifying plant communities that can then be translated into mapping units

Classifications of Macquarie Island's vegetation into plant communities have produced little agreement among researchers on the definitions of those communities in the past. In a bid to remedy this issue, unsupervised classification was used to identify stable groupings in floristic data

collected using GeoStrat and purposive sampling. The following choices involved in numerical classification were tested: the sampling design, taxa included in the analysis, strength of the data transformation, number of clusters, and clustering algorithm. The stability of the resulting clusters was primarily tested by measuring their mean silhouette width. Regardless of the sampling design, included taxa, transformation, number of clusters, or algorithm, no clustering attempt produced a mean silhouette width greater than 0.32. A mean silhouette width less than 0.3 indicates that there is no evidence of cluster structure. It was therefore concluded that stable groupings, suitable for defining robust mapping units, could not be found in the floristic data.

Objective 4: to produce accurate high resolution maps of the plant communities, and to use these maps to describe the spatial ecology of the vegetation of Macquarie Island. These maps should be sufficiently accurate to be used as baselines for future change detection.

The instability of the groupings in the floristic data meant there was little basis for defining robust, stable mapping units using the community paradigm. It was therefore considered more appropriate to map individual species or small groups of spectrally and ecologically similar species. Chapter 5 examined the effectiveness of a suite of mapping tools to accurately map the distribution of the endangered endemic cushion plant, *Azorella macquariensis*. This species was chosen because it has an ecologically distinct distribution, an independent validation dataset exists, and because accurate mapping is vital to understand the recently discovered widespread die-off.

There are two main approaches to vegetation mapping: species distribution modelling on the basis of environmental parameters and satellite image interpretation. Recently, attention has focussed on hybrid approaches that use both spectral and environmental data to predict the distribution of a species or community. Other developments in the field of image interpretation, most notably object-based image analysis and sophisticated machine learning tools, such as random forest classification, have also increased accuracies in vegetation mapping.

Chapter 5 examined the effectiveness of spectral data from very high resolution WorldView-2 satellite imagery, high resolution terrain data, pixel-based and object-based image analysis, and random forest classifications to maximise the accuracy of *Azorella* distribution maps. The accuracy of the maps was assessed using independent validation, the out-of-bag (OOB) accuracy estimate in the random forest tool, and visual inspection of the maps.

Of the three image training methods examined here, extracting multiple pixels per field site most improved the validation accuracy scores and the realistic appearance of the maps, but it made the OOB accuracy estimates unreliable. The highest accuracy was found when using a hybrid of terrain and spectral variables to train the classifier. Accuracies for the binary classification (presence/absence) were very high regardless of whether the input variables for the reduced classifications were chosen statistically (90.9%) or on the basis of *a priori* judgements of utility of those variables (93.2%). Using a three-class classification (absent, sparse, and moderate cover classes), the highest validation accuracy was found using a hypothesis-driven subset of input variables (89%).

Chapter 6 explored the application of the image interpretation techniques developed in chapter 5 to a multi-class classification of several dominant plant species on Macquarie Island. Two main approaches were examined: a hard classification that produced a map of the most likely species present in each pixel, and a map of overlapping classes to maximise the accuracy of the distribution of individual species. The preliminary results presented here showed that both approaches produced generally lower accuracies than the binary classification of *Azorella* presented in the last chapter. An exception to this general pattern was *Pleurophyllum hookeri*, a megaherb with a limited ecological range, with map accuracies of 88.9%. The overlapping classification improved user's accuracies for widespread species by including areas with mixed vegetation into the distribution of each target species. These maps also excluded some areas that the hard classification spuriously allocated to that class. Of the target species mapped in this chapter, only the *Pleurophyllum* class was considered to be sufficiently accurate to be used as a baseline for change detection, though these maps were

still an improvement on existing vegetation mapping for Macquarie Island. This chapter ended by describing directions for future research to improve map accuracies.

In summary, this study has developed and assessed methods for the major stages of mapping vegetation on sub-Antarctic Macquarie Island. These methods exploit recent developments in satellite image interpretation, terrain analysis, and classification methods. I demonstrated that those tools can produce extremely high accuracies in mapping Macquarie Island plants and is hence suitable for producing baseline maps for change detection. Furthermore I have produced the first distribution map for the critically endangered endemic species, *Azorella macquariensis* on Macquarie Island.

Appendix: Supplementary materials for Chapter 5

This appendix presents the results of binary classifications of *Azorella* distribution excluded from chapter 5.

Kappa statistics of the OOB errors for the binary classifications were affected by the spatial autocorrelation in the multiple pixel training method. The single pixel and object-based classifications performed similarly for all subsets of input variables. (Fig. S1)

The statistical hybrid and hypothesis-driven subsets of variables had the highest accuracies, performing significantly better than terrain or spectral classifications in isolation, with the spectral classification having the lowest accuracies for all training methods.
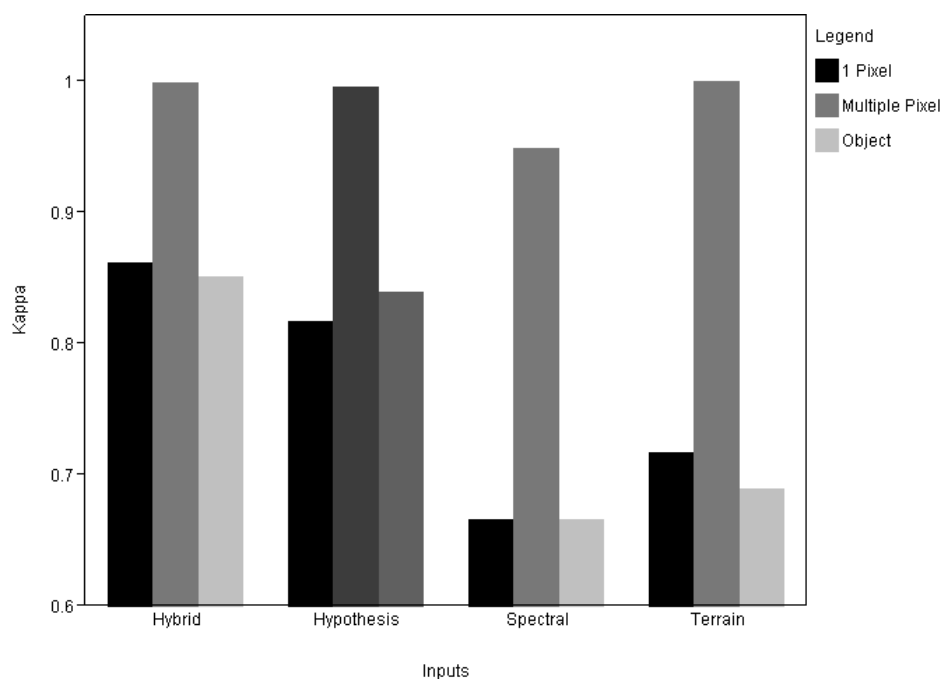
Fig. S1: Kappa statistics of the OOB accuracy measures for the two-class classifications of the presence and absence of Azorella*. For the multiple pixel-trained classifications, the kappa statistic is much higher than the other training methods for all sets of inputs. OOB, as implemented in the randomForest package, is unsuitable for use with multiple pixel training methods common in remote sensing applications.*

Terrain-based classifications

This section presents the confusion matrices (Table S1) and comparative maps (Fig. S2) for the terrain-based classifications for the three image training methods. All three terrain-based classifications produced solid patches of potential *Azorella* habitat across most of the plateau. The differences among the three maps are marginal, with the multiple pixel-trained map covering a slightly greater extent than the other maps.

Table S1: Confusion matrices for OOB accuracy estimates for single pixel-based (A); object-based (B); and multiple pixel-based (C) RF classifications based on terrain data only. These accuracy estimates are for classifications based on the smallest possible subset of input variables without reducing the accuracy of the classification.

(A)

		Predicted			
		Absent	**Present**	**Accuracy**	**Kappa**
Observed	**Absent**	123	13	90.4%	
	Present	12	53	81.5%	
	Accuracy	91.1%	80.3%	87.6%	0.717

(B)

		Predicted			
		Absent	**Present**	**Accuracy**	**Kappa**
Observed	**Absent**	123	12	91.1%	
	Present	15	50	76.9%	
	Accuracy	89.1%	80.6%	86.5%	0.689

(C)

		Predicted			
		Absent	**Present**	**Accuracy**	**Kappa**
Observed	**Absent**	3435	1	100.0%	
	Present	0	1660	100.0%	
	Accuracy	100.0%	99.9%	100.0%	0.9996

Fig. S2: Predicted Azorella *presence on northern Macquarie Island on the basis of terrain variables. The main panel shows the hard classes for* Azorella *presence based on single pixel, multiple pixel and object-based classifications. The predicted cover layers are partially transparent to demonstrate the overlaps in the predictions. Each of the image training methods resulted in solid blocks of predicted* Azorella *habitat on the plateau and mountain peaks, which is likely to show potential, rather than actual,* Azorella *distribution. The inset maps show the probability of* Azorella *presence based on single pixel (A); object (B); and multiple pixel -based (C) random forest classification of terrain.*

Spectral-based classifications

This section presents the confusion matrices (Table S2) and comparative maps (Fig. S3) for the spectral-based classifications for the three image training methods. The three spectral-based classifications produced much more fragmented patches of potential Azorella habitat than the terrain-based classifications. These maps all spuriously included beaches and rocky coastal headlands in the area of predicted *Azorella* presence. The differences among the three maps are subtle, with the multiple pixel-trained map covering a slightly greater extent than the other maps.

Table S2: Confusion matrices for the pixel-based (A) and object-based (B) classifications of the field sites on the basis of spectral data only.

(A)

		Predicted			
		Absence	Presence	Percentage	Kappa
Observed	**Absence**	123	13	90.4%	
	Presence	16	49	75.4%	
	Percentage	88.5%	79.0%	85.6%	0.666

(B)

		Predicted			
		Absence	Presence	Percentage	Kappa
Observed	**Absence**	122	13	90.4%	
	Presence	16	49	75.4%	
	Percentage	88.4%	79.0%	85.5%	0.666

(C)

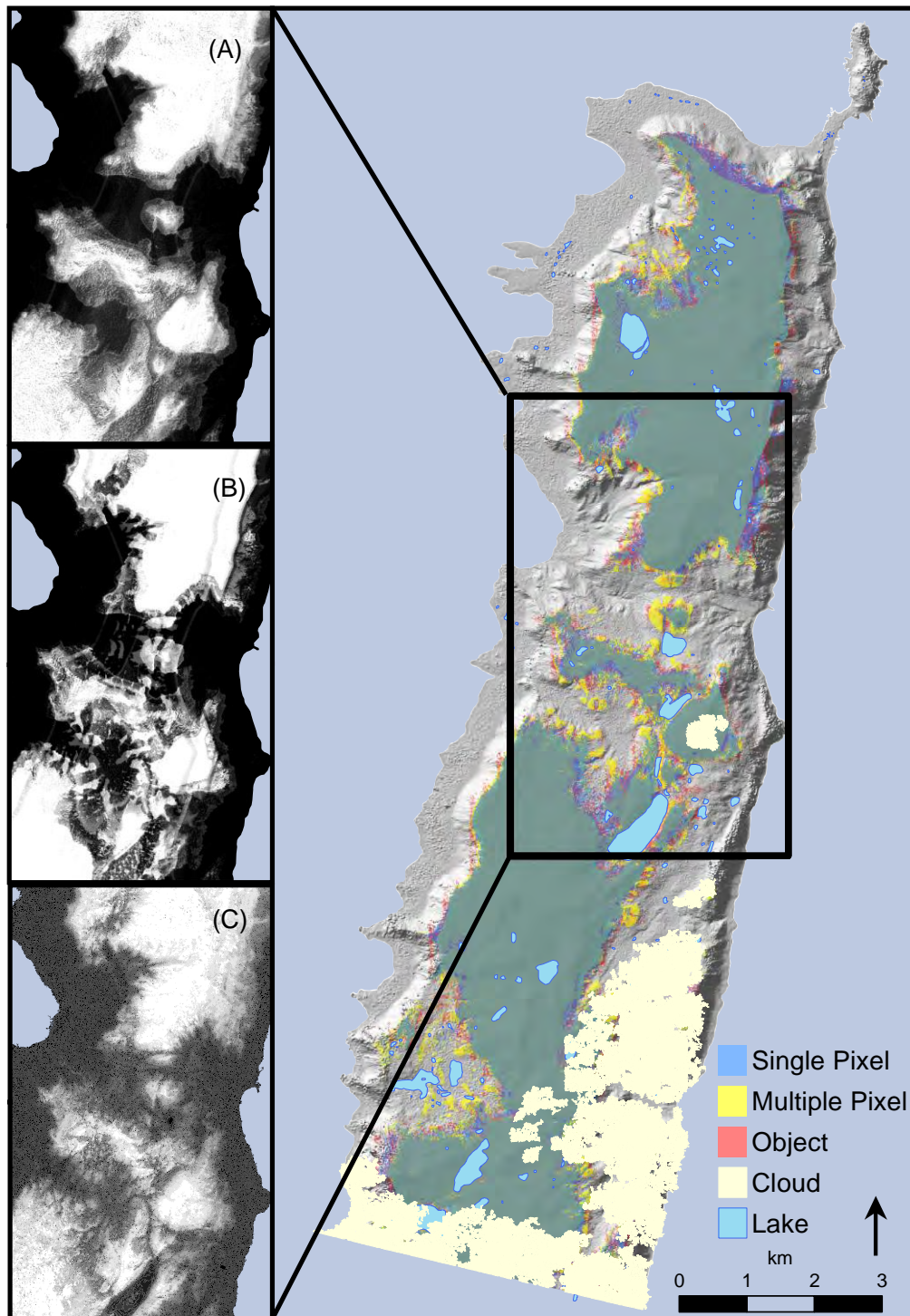|		Predicted			
		Absent	Present	Accuracy	Kappa
Observed	**Absent**	3381	55	98.4%	
	Present	60	1600	96.4%	
	Accuracy	98.3%	96.7%	97.7%	0.949

Fig. S3: Predicted Azorella *presence on northern Macquarie Island on the basis of spectral variables. The main panel shows the hard classes for* Azorella *presence based on single pixel, multiple pixel and object-based classifications. The predicted layers are partly transparent to demonstrate the overlaps among the predictions. The differences among the three image training methods are subtle. The spectral classifications resulted in much more fragmented predicted distributions than the terrain-based classifications, and predicted* Azorella *in coastal areas where the species is known not grow. The inset maps show the probability of* Azorella *presence based on single pixel (A); multiple pixel (B); and object-based (C) random forest classification of terrain.*
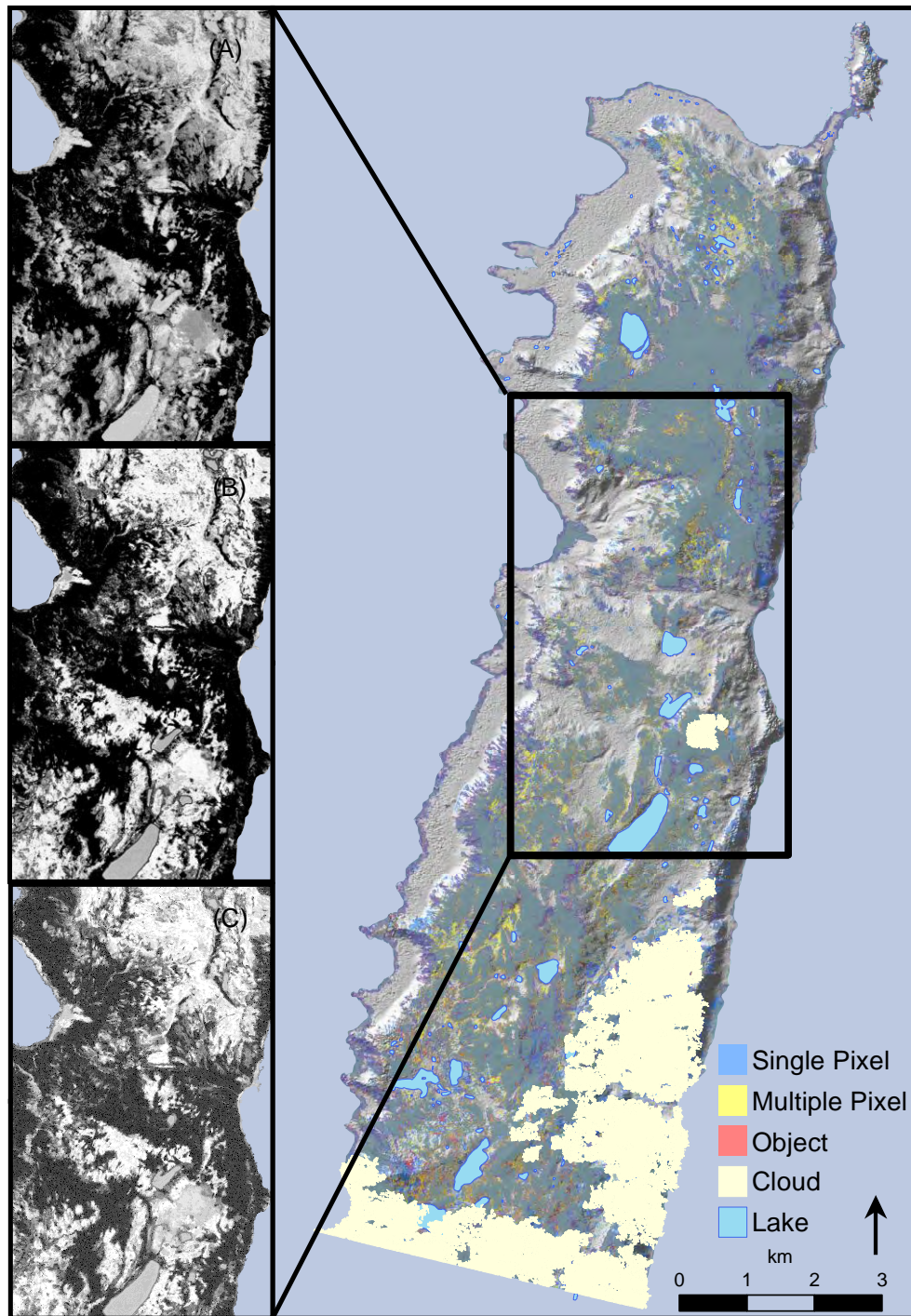
Hybrid Classifications

This section presents the confusion matrices (Table S3) and comparative maps (Fig. S4) for the hybrid classifications for the three image training methods, with the input variables chosen on statistical grounds. The three hybrid classifications all performed as compromises between the hybrid and spectral classifications. The areas of predicted presence were all confined to the higher parts of the plateau, and were more fragmented than those in the terrain-based classification. The differences among the classifications based on the three training methods were subtle, with the single pixel-trained classification having the most restricted distribution, and the multiple pixel and object-based classifications extending around the fringes of the single pixel-trained presence class.

Table S3: Confusion matrices for the single pixel-based (A); object-based (B); and multiple pixel-based (C) RF classifications, using both terrain derivatives and spectral data.

(A)

Observed		Predicted			
		Absent	Present	Accuracy	Kappa
	Absent	131	5	96.3%	
	Present	7	58	89.2%	
	Accuracy	94.9%	92.1%	94.0%	0.862

(B)

Observed		Predicted			
		Absent	Present	Accuracy	Kappa
	Absent	129	6	95.6%	
	Present	7	58	89.2%	
	Accuracy	94.9%	90.6%	93.5%	0.851

(C)

Observed		Predicted			
		Absent	Present	Accuracy	Kappa
	Absent	3435	1	100.0	
	Present	1	1659	99.9	
	Accuracy	100.0	99.9	100.0	0.999
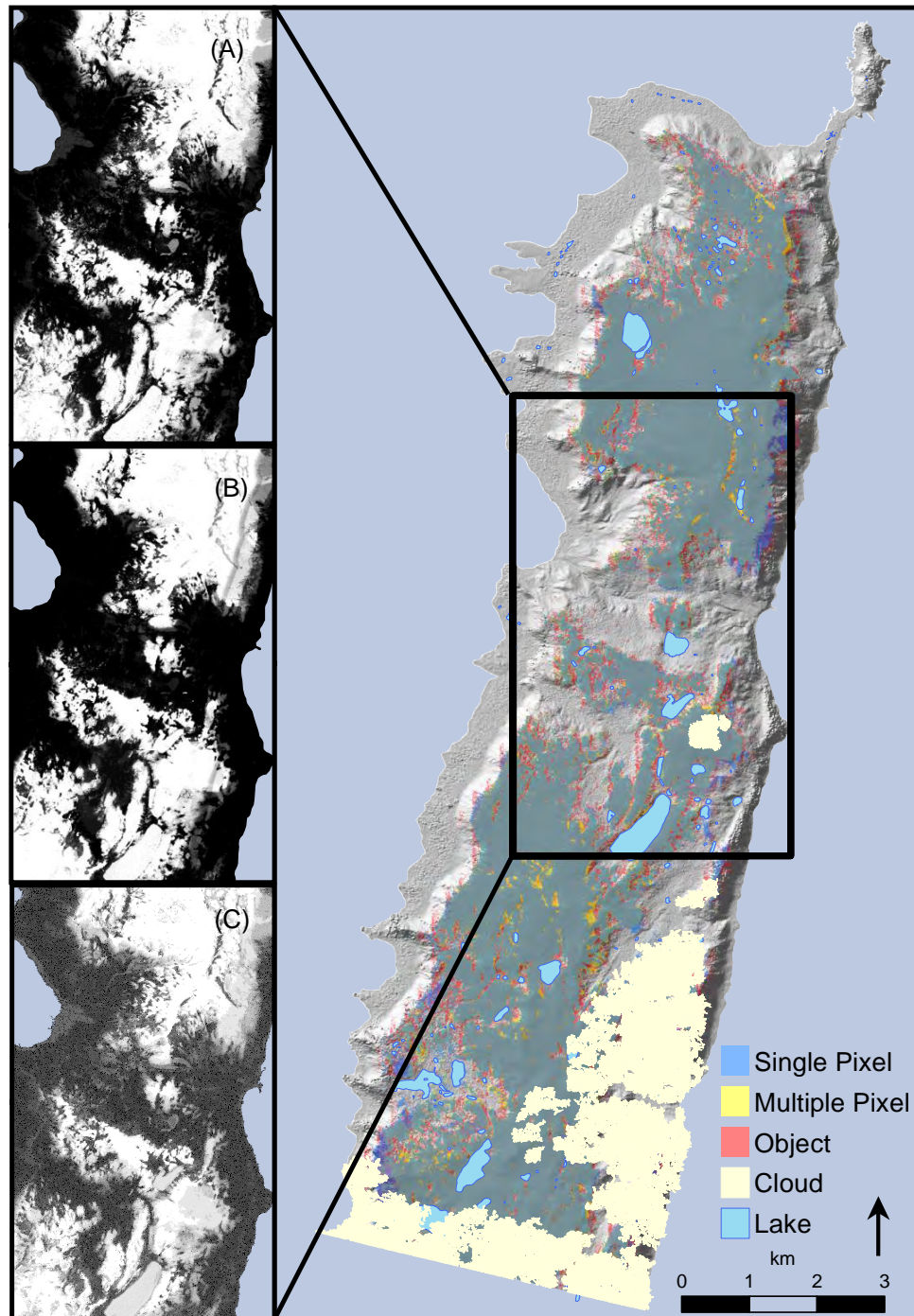
Fig S4: Predicted Azorella *presence on northern Macquarie Island, from single pixel, object and multiple pixel based classifications of the hybrid variables. The predicted distribution layers are partly transparent, to demonstrate the overlaps among the predicted distributions. The hybrid RF classifications were more fragmented than the terrain-based classifications and lacked the spurious predictions of Azorella presence in coastal areas that the spectral-based classifications produced.The main map shows the hard classes for* Azorella *presence while the inset maps show the probability of* Azorella *presence based on single-pixel (A); multiple pixel (B); and object (C) trained classifications.*

Hypothesis-driven Classifications

This section presents the confusion matrices (Table S4), partial dependence plots (Fig. S5) and comparative maps (Fig. S6) for the classifications based on a hypothesis-driven subset of the available input variables for the three image training methods. These three maps appeared very similar to those based on statistically-derived hybrid sets of variables, though the accuracies were slightly higher.

Classifications based on a hypothesis-driven subset of variables showed slightly improved accuracies compared to the hybrid models, when measured by the independent validation and slightly lower when measured by the OOB accuracy estimate. The single and multiple pixel classifications both had accuracies greater than 90% on both measures, while the object-based classification had lower accuracy, with 80.7% by independent validation, and 87.7% by OOB for the present class.

Table S4: Confusion matrices for the single pixel-based (A); object-based (B); and multiple pixel-based (C) RF classifications, using a hypothesis-drive subset of the terrain derivatives and spectral data.

(A)

		Predicted		
		Absent	**Present**	**Accuracy**
Observed	**Absent**	129	7	94.9%
	Present	9	56	86.2%
	Accuracy	93.5%	88.9%	92.0%

(B)

		Predicted		
		Absent	**Present**	**Accuracy**
Observed	**Absent**	129	6	95.6%
	Present	8	57	87.7%
	Accuracy	94.2%	90.5%	93.0%

(C)

		Predicted		
		Absent	**Present**	**Accuracy**
Observed	**Absent**	3432	4	99.9
	Present	4	1656	99.8
	Accuracy	99.9	99.8	99.8

The partial dependence plots showed that many of the input variables had non-linear relationships with the probability of *Azorella* presence. These plots showed that *Azorella* presence was linked to high values for elevation, solar radiation, slope, solar radiation, and blue and yellow reflectance; with lower values for red edge and NIR 2 reflectance, GLCM Mean; and with extreme values for green reflectance, GLCM homogeneity and entropy, curvature, valleyness and wetness index.
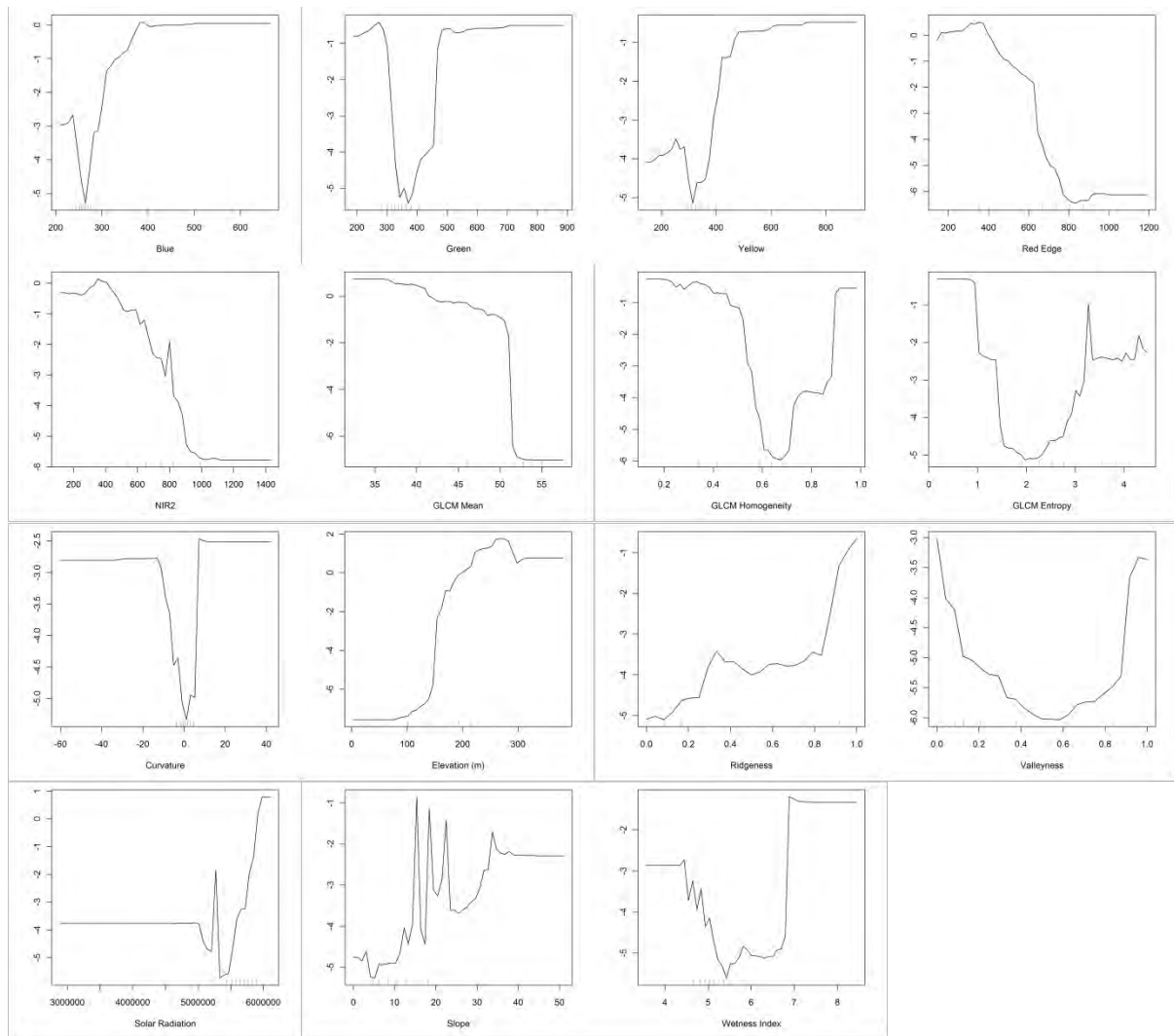
Fig. S5: Partial dependence plots for the present class for the multiple pixel-based classification of hypothesis-driven subset of input variables. Azorella *presence was associated with high values for blue and yellow reflectance, elevation, ridgeness, slope, and solar radiation; with low levels of red edge and NIR2 reflectance, and the GLCM mean texture measure; and with mixed values for green reflectance, GLCM homogeneity and entropy, curvature, valleyness, and wetness index.*
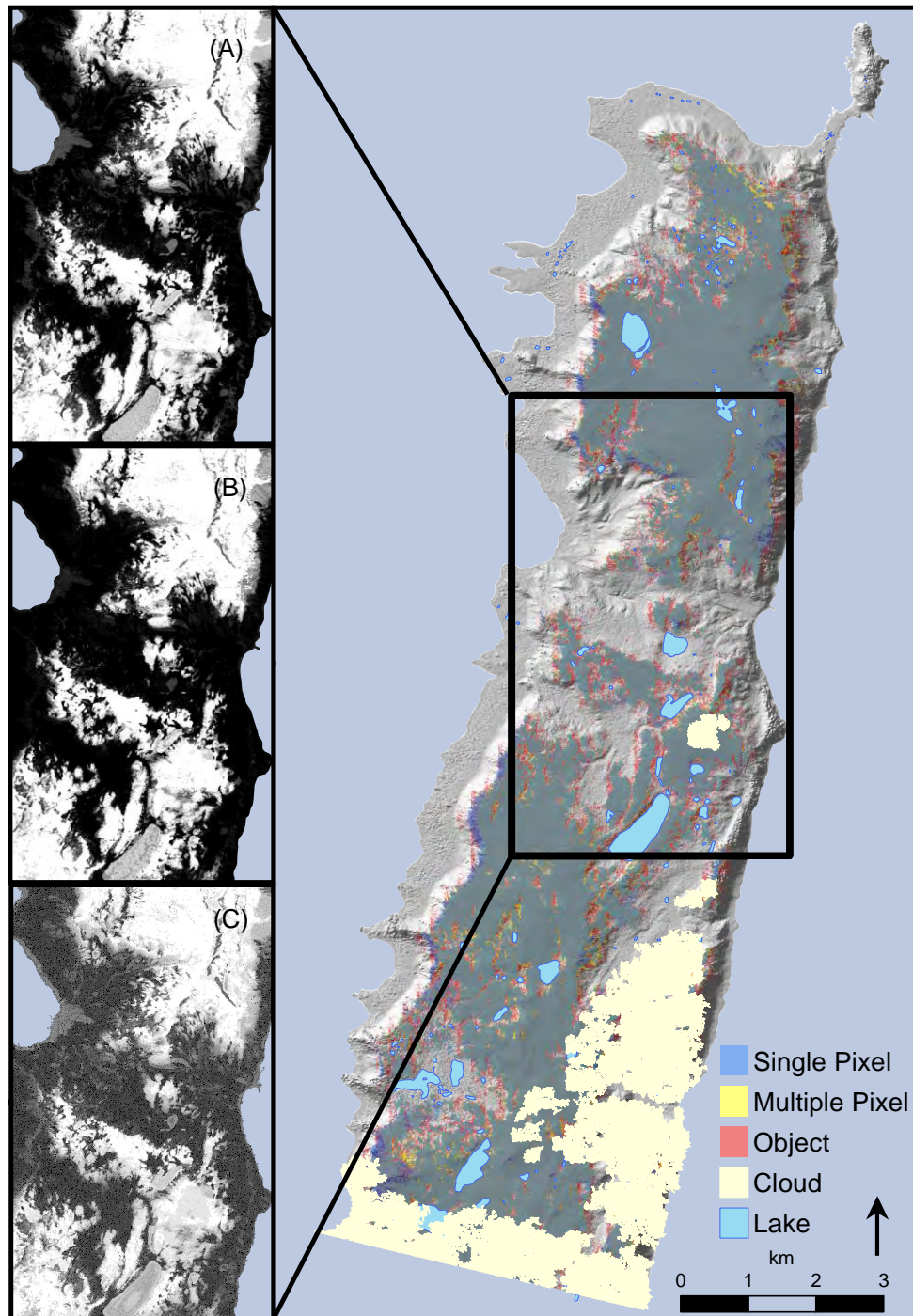
Fig. S6: Predicted Azorella *presence on northern Macquarie Island, from single pixel, object and multiple pixel based classifications of the hypothesis-driven subset of variables. The main map shows the hard classes for* Azorella *presence, with the three predictive layers made partially transparent to demonstrate the overlaps among the predictions. Each of these predictions was similar to the classifications based on the statistically-chosen subsets of hybrid variables. The inset maps show the probability of* Azorella *presence based on single-pixel (A); multiple pixel (B); and object (C) trained classifications.*

References

ABRS (1993) Flora of Australia: Volume 50, Oceanic Islands 2. AGPS and CSIRO Publishing

Addink EA, Jong SMD, Pebesma EJ (2007) The Importance of Scale in Object-based Mapping of Vegetation Parameters with Hyperspectral Imagery. Photogrammetric Engineering & Remote Sensing 73:905-912

Alquezar R, Boyd W (2007) Development of rapid, cost effective coral survey techniques: tools for management and conservation planning. Journal of Coastal Conservation 11:105-119

Archaux F, Bergès L, Chevalier R (2007) Are Plant Censuses Carried Out on Small Quadrats More Reliable than on Larger Ones? Plant Ecology 188:179-190

Archaux F, Camaret S, Dupouey J-L, Ulrich E, Corcket E, Bourjot L, Brêthes A, Chevalier R, Dobremez J-F, Dumas Y, Dumé G, Forêt M, Forgeard F, Lebret Gallet M, Picard J-F, Richard F, Savoie J-M, Seytre L, Timbal J, Touffet J (2009) Can we reliably estimate species richness with large plots? An assessment through calibration training. Plant Ecology 203:303-315

Archaux F, Gosselin F, Bergès L, Chevalier R (2006) Effects of sampling time, species richness and observer on the exhaustiveness of plant censuses. Journal of Vegetation Science 17:299-306

Ashby E (1935) The Quantitative Analysis of Vegetation. Annals of Botany XLIX:779-801

Aspinall RJ, Lees BG (1994) Sampling and Analysis of Spatial Environmental Data. In: Advances in GIS Research. Taylor & Francis, Southampton, pp 1086-1098

Aspinall R, 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 59:537-543

Austin MP (1985) Continuum concept, ordination methods, and niche theory. Annual Review of Ecology and Systematics 16:39-61

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

Austin MP, Heyligers PC (1989) Vegetation Survey Design for Conservation: Gradsect Sampling of Forests in North-eastern New South Wales. Biological Conservation 50:13-32

Benedetti-Cecchi L, Airoldi L, Abbiati M, Cinelli F (1996) Estimating the abundance of benthic invertebrates: a comparison of procedures and variability between observers. Marine Ecology Progress Series 138:93-101

Bennett LT, Judd TS, Adams MA (2000) Close-Range Vertical Photography for Measuring Cover Changes in Perennial Grasslands. Journal of Range Management 53:634-641

Bergstrom DM, Chown SL (1999) Life at the front: history, ecology and change on southern ocean islands. Trends in Ecology and Evolution 14:472-477

Bergstrom DM, Lucieer A, Kiefer K, Wasley J, Belbin L, Pedersen TK, Chown SL (2009a) Indirect effects of invasive species removal devastate World Heritage Island. Journal of Applied Ecology 46:73- 81

Bergstrom DM, Lucieer A, Kiefer K, Wasley J, Belbin L, Pedersen TK, Chown SL (2009b) Management implications of the Macquarie Island trophic cascade revisited: a reply to Dowding et al. (2009). Journal of Applied Ecology 46:1133-1136

Bergstrom DM, Whinam J, Belbin L (2002) A Classification of Subantarctic Heard Island Vegetation. Arctic, Antarctic, and Alpine Research 34:169-177

Beven KJ, Kirkby MJ (1979) A physically based, variable contributing area model of basin hydrology. Hydrological Sciences Journal 24:43-69

Beyer HL (2004) Hawth's Analysis Tools for ArcGIS.

Beyer HL (2009) Geospatial Modelling Environment (0.6.2.0).

Bezdek JC, Erhlich R, Full W (1984) FCM: The fuzzy c-means clustering algorithm. Computers & Geosciences 10:191-203

Blaschke T (2010) Object based image analysis for remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing 65:2-16

Booth DT, Cox SE, Meikle T, Zuuring HR (2008) Ground-cover measurements: assessing correlation among aerial and ground-based methods. Environmental management 42:1091-1100

Boyd DS, Sanchez-Hernandez C, Foody GM (2006) Mapping a specific class for priority habitats monitoring from satellite sensor data. International Journal of Remote Sensing 27:2631-2644

Bray JR, Curtis JT (1957) An Ordination of the Upland Forest Communities of Southern Wisconsin. Ecological Monographs 27:326-349

Breiman L (2001) Random Forests. Machine Learning 45:5-32

Brolsma H (2008) Macquarie Island AIRSAR DEM (Digital Elevation Model), Australian Antarctic Data Centre, http://data.aad.gov.au.

Bråkenhielm S, Qinghong L (1995) Comparison of field methods in vegetation monitoring. Water, Air and Soil Pollution 79:75-87

Bureau of Meterorology B (2011) Climate statistics for Australian locations: Summary statistics Macquarie Island., http://www.bom.gov.au/climate/data/

Burrough PA, Frank AU (1996) Geographic Objects with Indeterminate Boundaries. Taylor and Francis, London, UK

Burrough PA, Gaans PFM van, MacMillan RA (2000) High-resolution Landform Classification Using Fuzzy *k*-means. Fuzzy Sets and Systems 113:37-52

References

Burrough PA, Wilson JP, Gaans PFMV, Hansen AJ (2001) Fuzzy *k*-means classification of topo-climatic data as an aid to forest mapping in the Greater Yellowstone Area , USA. Landscape Ecology 16:523-546

Cairns DM, Moen JON (2004) Herbivory influences tree lines. Journal of Ecology 92:1019-1024

Carlsson ALM, Bergfur J, Milberg P (2005) Comparison of data from two vegetation monitoring methods in semi-natural grasslands. Environmental Monitoring and Assessment 100:235-48

Chapuis J-L, Frenot Y, Lebouvier M (2004) Recovery of native plant communities after eradication of rabbits from the subantarctic Kerguelen Islands , and influence of climate change. Biological Conservation 117:167-179

Cherrill A, McClean C (1999) The reliability of `Phase 1' habitat mapping in the UK : the extent and types of observer bias. Landscape and Urban Planning 45:131-143

Chiarucci A (2007) To sample or not to sample? That is the question ... for the vegetation scientist. Folia Geobotanica 42:209-216

Chown S, Rodrigues ASL, Gremmen NJM, Gaston KJ (2001) World Heritage Status and Conservation of Southern Ocean Islands. Conservation Biology 15:550-557

Chown SL, Smith VR (1993) Climate change and the short-term impact of feral house mice at the sub-Antarctic Prince Edward Islands. Oecologia 96:508-516

Chown SL, Spear D, Lee JE, Shaw JD (2009) Animal Introductions to Southern Systems: Lessons for Ecology and for Policy. African Zoology 44:248-262

Chytrý M (2001) Phytosociological data give biased estimates of species richness. Journal of Vegetation Science 12:439-444

Clarke KR (1993) Non-parametric multivariate analyses of changes in community structure. Australian Journal of Ecology 18:117-143

Clarke K, Somerfield P, Chapman M (2006) On resemblance measures for ecological studies, including taxonomic dissimilarities and a zero-adjusted Bray–Curtis coefficient for denuded assemblages. Journal of Experimental Marine Biology and Ecology 330:55-80

Clarke K, Somerfield P, Gorley R (2008) Testing of null hypotheses in exploratory community analyses: similarity profiles and biota-environment linkage. Journal of Experimental Marine Biology and Ecology 366:56-69

Coker PD (2000) Vegetation analysis, mapping and environmental relationships at a local scale, Jotunheimen, Southern Norway. In: Alexander R, Millington AC (eds) Vegetation Mapping. John Wiley & Sons, Chichester, England, pp 136-158

Convey P (2007) Influences on and origins of terrestrial biodiversity of the sub-Antarctic islands. Papers and Proceedings of the Royal Society of Tasmania 141:83-93

Cooley T, Anderson GP, Felde GW, Hoke ML, Ratkowski AJ, Chetwynd JH, Gardner JA, Adler-Golden SM, Matthew MW, Berk A, Bernstein LS, Acharya PK, Miller D, Lewis P (2002) FLAASH , a

MODTRAN4-based Atmospheric Correction Algorithm , Its Application and Validation. In: Geoscience and Remote Sensing Symposium, IGARSS '02. IEEE International, pp 1414-1418

Copson GR (1984) An Annotated Atlas of the Vascular Flora of Macquarie Island.

Copson G, Whinam J (1998) Response of Vegetation on Subantarctic Macquarie Island to Reduced Rabbit Grazing. Australian Journal of Botany 46:15-24

Crimmins MA, Crimmins TM (2008) Monitoring plant phenology using digital repeat photography. Environmental management 41:949-58

Cushman SA, Evans JS, McGarigal K, Kiesecker JM (2010) Toward Gleasonian Landscape Ecology: From Communities to Species, From Patches to Pixels.

Cutler DR, Edwards TC, Beard KH, Cutler A, Hess KT, Gibson J, Lawler JJ (2007) Random forests for classification in ecology. Ecology 88:2783-2792

Daemane ME, Cilliers SS, Bezuidenhout H (2010) An ecological study of the plant communities in the proposed Highveld National Park, in the peri-urban area of Potchefstroom, South Africa. Koedoe 52:Art. #708, 8 pages

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

Department, of Primary Industries P and W (2009) Tasmanian Threatened Species Listing Statement: Azorella macquariensis (macquarie cushions). Tasmania, Australia

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

Dirnböck T, Dullinger S, Gottfried M, Ginzler C, Grabherr G (2003) Mapping alpine vegetation based on image analysis, topographic variables and Canonical Correspondence Analysis. Applied Vegetation Science 6:85-96

Dobrowski SZ, Safford HD, Cheng YB, Ustin SL (2008) Mapping mountain vegetation using species distribution modeling, image-based texture analysis, and object-based classification. Applied Vegetation Science 11:499-508

Dowding JE, Murphy EC, Springer K, Peacock AJ, Krebs CJ (2009) Cats, rabbits, Myxoma virus, and vegetation on Macquarie Island: a comment on Bergstrom et al. (2009). Journal of Applied Ecology 46:1129-1132

Drummond SP, Connell SD (2005) Quantifying percentage cover of subtidal organisms on rocky coasts: a comparison of the costs and benefits of standard methods. Marine and Freshwater Research 56:865-876

Evans JS, Cushman S a. (2009) Gradient modeling of conifer species using random forests. Landscape Ecology 24:673-683

References

Faber-Langendoen D, Aaseng N, Hop K, Lew-Smith M, Drake J (2007) Vegetation classification, mapping, and monitoring at Voyageurs National Park, Minnesota: An application of the U.S. National Vegetation Classification. Applied Vegetation Science 10:361-374

Feick RD, Hall GB (2002) Balancing consensus and conflict with a GIS-based multi-participant , multi-criteria decision support tool. GeoJournal 53:391-406

Field JG, Clarke KR, Warwick R (1982) A Practical Strategy for Analysing Multispecies Distribution Patterns. Marine Ecology Progress Series 8:37-52

Field SA, O'Connor PJ, Tyre AJ, Possingham HP (2007) Making monitoring meaningful. Austral Ecology 32:485-491

Fielding AH (2007) Cluster and Classification Techniques for the Biosciences. Cambridge University Press, Cambridge, UK

Fisher P, Arnot C, Wadsworth R, Wellens J (2006) Detecting change in vague interpretations of landscapes. Ecological Informatics 1:163-178

Fisher P, Wood J, Cheng T (2004) Where is Helvellyn? Fuzziness of multi-scale landscape morphometry. Transactions of the Institute of British Geographers 29:106-128

Foody GM (2010) Assessing the accuracy of land cover change with imperfect ground reference data. Remote Sensing of Environment 114:2271-2285

Fortin M-J, Drapeau P, Legendre P (1989) Spatial autocorrelation and sampling design in plant ecology. Vegetatio 83:209-222

Fortin M, Olson RJ, Ferson S, Iverson L, Hunsaker C, Edwards G, Levine D, Butera K, Klemas V (2000) Issues related to the detection of boundaries. Landscape Ecology 15:453-466

Foster MS, Harrold C, Hardin DD (1991) Point vs. photo quadrat estimates of the cover of sessile marine organisms. Journal of Experimental Marine Biology and Ecology 146:193-203

Franklin J (1995) Predictive vegetation mapping: geographic modelling of biospatial patterns in relation to environmental gradients. Progress in Physical Geography 19:474-499

Franklin J, Keeler-Wolf T, Thomaas KA, Shaari DA, Stine PA, Michaelsen J, Miller J (2001) Stratified sampling for field survey of environmental gradients in the Mojave Desert ecoregion. In: Millington AC, Walsh SJ, Osborne PE (eds) GIS and Remote Sensing Applications in Biogeography and Ecology. Kluwer Academic Publishers, Boston, Dordrecht, London, pp 133-139

Franklin SE, Maudie AJ, Lavigne MB (2001) Using Spatial Co-Occurrence Texture to Increase Forest Structure and Species Composition Classification Accuracy. Photogrammetric Engineering & Remote Sensing 67:849-855

Franklin J, Miller JA (2009) Mapping Species Distributions: Spatial Inference and Prediction. Cambridge University Press, Cambridge, UK

Frenot Y, Chown SL, Whinam J, Selkirk PM, Convey P, Skotnicki M, Bergstrom DM (2005) Biological invasions in the Antarctic: extent, impacts and implications. Biological Review 80:45-72

Fretwell PT, Convey P, Fleming AH, Peat HJ, Hughes KA (2011) Detecting and mapping vegetation distribution on the Antarctic Peninsula from remote sensing data. Polar Biology 34:273-281

Fu P, Rich PM (2000) A geometric solar radiation model and its applications in agriculture and forestry. In: Proceedings of the Second International Conference on Geospatial Information in Agriculture and Forestry.pp 357-364

Fu P, Rich PM (2002) A geometric solar radiation model with applications in agriculture and forestry. Computers and Electronics in Agriculture 37:25-35

Gallant JC, Wilson JP (2000) Primary topographic attributes. In: Terrain Analysis—Principles and Applications. Wiley, New York, pp 51-85

Gilbert JA, Butt KR (2009) Evaluation of digital photography as a tool for field monitoring in potentially inhospitable environments. Mires and Peat 5:1-6

Gislason PO, Benediktsson JA, Sveinsson JR (2006) Random Forests for land cover classification. Pattern Recognition Letters 27:294-300

Goedickemeier I, Wildi O, Kienast F (1997) Sampling for vegetation survey: Some properties of a GIS-based stratification compared to other statistical sampling methods. Coenoses 12:43-50

Grabherr G, Reiter K, Willner W (2003) Towards objectivity in vegetation classification : the example of the Austrian forests. Plant Ecology 169:21-34

Gray AN, Azuma DL (2005) Repeatability and implementation of a forest vegetation indicator. Ecological Indicators 5:57-71

Green DR, Hartley S (2000) Integrating photointerpretation and GIS for vegetation mapping: some issues of error. In: Alexander R, Millington AC (eds) Vegetation Mapping. John Wiley and Sons, Chichester, England, pp 103 - 134

Grossmann E, Ohmann J, Kagan J, May H, Gregory M (2010) Mapping Ecological Systems with a Random Forest Model: Tradeoffs between Errors and Bias. Gap Analysis Bulletin 17:16-22

de Gruijter J (1999) Spatial sampling schemes for remote sensing. In: Stein A, van Der Meer F, Gorte B (eds) Spatial Statistics for Remote Sensing. Kluwer Academic Publishers, Dordrecht, pp 211-242

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

Haralick RM, Shanmugan K, Dinstein I (1973) Textural Features for Image Classification. IEEE Transactions on Systems, Man and Cybernetics SMC-3:610-621

Harris G (2003) Ecological paradigms, change detection and prediction. In: Zerger A, Argent RM (eds) MODSIM 2005 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, Townsville, Australia

References

Harris GP, Heathwaite AL (2012) Why is achieving good ecological outcomes in rivers so difficult? Freshwater Biology 57:91-107

Hearn SM, Healey JR, McDonald MA, Turner AJ, Wong JLG, Stewart GB (2011) The repeatability of vegetation classification and mapping. Journal of Environmental Management 92:1174-84

Hill RA, Thomson AG (2005) Mapping woodland species composition and structure using airborne spectral and LiDAR data. International Journal of Remote Sensing 26:3763-3779

Hirzel A, Guisan A (2002) Which is the optimal sampling strategy for habitat suitability modelling. Ecological Modelling 157:331-341

Horn BKP (1981) Hill shading and the reflectance map. Proceedings of the IEEE 69:14-47

Huang S, Rich PM, Crabtree RL, Potter CS, Fu P (2008) Modeling Monthly Near-Surface Air Temperature from Solar Radiation and Lapse Rate: Application over Complex Terrain in Yellowstone National Park. Physical Geography 29:158-178

Hüttich C, Herold M, Strohbach BJ, Dech S (2011) Integrating in-situ, Landsat, and MODIS data for mapping in Southern African savannas: experiences of LCCS-based land-cover mapping in the Kalahari in Namibia. Environmental Monitoring and Assessment 176:531-47

Jano AP, Jefferies RL, Rockwell RF (1998) The detection of vegetational change by multitemporal analysis of LANDSAT data: the effects of goose foraging. Journal of Ecology 86:93-99

Januchowski SR, Pressey RL, VanDerWal J, Edwards A (2010) Characterizing errors in digital elevation models and estimating the financial costs of accuracy. International Journal of Geographical Information Science 24:1327-1347

Kaufman L, Rousseeuw PJ (2005) Finding Groups in Data. John Wiley & Sons, Hoboken, New Jersey, U.S.A.

Kendeigh SC (1954) History and Evaluation of Various Concepts of Plant and Animal Communities in North America. Ecology 35:152-171

Kennedy AD (1995) Antarctic terrestrial ecosystem response to global environmental change. Ecology, Evolution, and Systematics 26:683-704

Kirkpatrick JB, Scott J (2002) Change in Undisturbed Vegetation on the Coastal Slopes of Subantarctic Macquarie Island, 1980-1995. Arctic, Antarctic, and Alpine Research 34:300-307

Klasner FK, Fagre DB (2002) A Half Century of Change in Alpine Treeline Patterns at Glacier National Park , Montana. Arctic, Antarctic, and Alpine Research 34:49-56

Klimeš L, Dančak M, Hájek M, Jongepierová I, Kučera T (2001) Scale-dependent biases in species counts in a grassland. Journal of Vegetation Science 12:699-704

Kohler KE, Gill SM (2006) Coral Point Count with Excel extensions (CPCe): A Visual Basic program for the determination of coral and substrate coverage using random point count methodology. Computers & Geosciences 32:1259-1269

Kriwoken LK, Holmes N (2007) Emerging Issues of Australia's Sub-Antarctic Islands: Macquarie Island and Heard Island and McDonald Islands. In: Looking South: Australia's Antarctic Agenda. Federation Press, Annandale, NSW, Australia, pp 149-161

Laliberte AS, Rango A, Havstad KM, Paris JF, Beck RF, McNeely R, Gonzalez AL (2004) Object-oriented image analysis for mapping shrub encroachment from 1937 to 2003 in southern New Mexico. Remote Sensing of Environment 93:198-210

Laliberte AS, Rango A, Herrick JE, Fredrickson EL, Burkett L (2007) An object-based image analysis approach for determining fractional cover of senescent and green vegetation with digital plot photography. Journal of Arid Environments 69:1-14

Lawrence R, Wood S, Sheley R (2006) Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (RandomForest). Remote Sensing of Environment 100:356-362

Lepš J, Hadincová V (1992) How reliable are our vegetation analyses? Journal of Vegetation Science 3:119-124

Lepš J, Šmilauer P (2007) Subjectively sampled vegetation data: Don't throw out the baby with the bath water. Folia Geobotanica 42:169-178

Leujak W, Ormond R (2007) Comparative accuracy and efficiency of six coral community survey methods. Journal of Experimental Marine Biology and Ecology 351:168-187

Liang L, Schwartz MD, Fei S (2011) Examining Spring Phenology of Forest Understory using Digital Photography. In: Proceedings of the 17th Central Hardwood Forest Conference.pp 51-57

Liaw A, Wiener M (2002) Classification and Regression by randomForest. R News 2/3:18-22

Liston GE, Sturm M (1998) A snow-transport model for complex terrain. Journal of Glaciology 44:498-516

Lu D, Weng Q (2007) A survey of image classification methods and techniques for improving classification performance. International Journal of Remote Sensing 28:823-870

Lucieer A (2008) Mapping grazed vegetation communities on Macquarie Island using a binary ensemble classifier. In: Australasian Remote Sensing and Photogrammetry Conference (ARSPC). Darwin, Australia

Luscier JD, Thompson WL, Wilson JM, Gorham BE, Dragut LD (2006) Using digital photographs and object-based image analysis to estimate percent ground cover in vegetation plots. Frontiers in Ecology and the Environment 4:408-413

Lájer K (2007) Statistical tests as inappropriate tools for data analysis performed on non-random samples of plant communities. Folia Geobotanica 42:115-122

Maechler M, Rousseeuw P, Struyf A, Hubert M (2002) Cluster Analysis Basics and Extensions. R package version 1.14.2.

References

Mallinis G (2008) Object-based classification using Quickbird imagery for delineating forest vegetation polygons in a Mediterranean test site. ISPRS Journal of Photogrammetry and Remote Sensing 63:237-250

Mannel S, Price M, Hua D (2011) Impact of reference datasets and autocorrelation on classification accuracy. International Journal of Remote Sensing 32:5321-5330

Michel P, Mathieu R, Mark AF (2010) Spatial analysis of oblique photo-point images for quantifying spatio-temporal changes in plant communities. Applied Vegetation Science 13:173-182

Mitchell AD, Meurk CD, Wagstaff SJ (1999) Evolution of Stilbocarpa, a megaherb from New Zealand's sub-antarctic islands. New Zealand Journal of Botany 37:205-211

Moore ID, Grayson RB, Ladson AR (1991) Digital Terrain Modelling: A review of hydrological, geomorphological, and biological applications. Hydrological Processes 5:3-30

Murray H, Lucieer A, Williams R (2010) Texture-based classification of sub-Antarctic vegetation communities on Heard Island. International Journal of Applied Earth Observation and Geoinformation 12:138-149

Nagendra H (2001) Using remote sensing to assess biodiversity. International Journal of Remote Sensing 22:2377-2400

van Niel KP, Laffan SW, Lees BG (2004) Effect of error in the DEM on environmental variables for predictive vegetation modelling. Journal of Vegetation Science 15:747-756

Nilsen L, Brossard T, Joly D (1999) Mapping plant communities in a local Arctic landscape applying a scanned infrared aerial photograph in a geographical information system. International Journal of Remote Sensing 20:463-480

Nilsson IN, Nilsson SG (1985) Experimental Estimates of Census Efficiency and Pseudoturnover on Islands: Error Trend and Between-Observer Variation when Recording Vascular Plants. Journal of Ecology 73:65-70

Northup BK, Brown JR, Dias CD, Skelly WC, Radford B (1999) A technique for near-ground remote sensing of herbaceous vegetation in tropical woodlands. Rangelands Journal 21:229-243

Noy-Meir I, van der Maarel E (1987) Relations between community theory and community analysis in vegetation science: some historical perspectives. Vegetatio 69:5-15

Oksanen J, Sarjakoski T (2005) Error propagation of DEM-based surface derivatives. Computers & Geosciences 31:1015-1027

Pal M (2005) Random forest classifier for remote sensing classification. International Journal of Remote Sensing 26:217-222

Palmer MW, White PS (1994) On the existence of ecological communities. Journal of Vegetation Science 5:279-282

Parks and Wildlife Service (2007) Plan for the Eradication of Rabbits and Rodents on Subantarctic Macquarie Island.

Paruelo JM, Lauenroth WK, Roset PA (2000) Technical note: Estimating aboveground plant biomass using a photographic technique. Journal of Range Management 53:190-193

Pendlebury SF, Barnes-Keoghan IP (2007) Climate and climate change in the sub-Antarctic. Papers and Proceedings of the Royal Society of Tasmania 141:67-82

Pike RJ (2000) Geomorphometry -diversity in quantitative surface analysis. Progress in Physical Geography 24:1-20

Pillar VD (1998) Sampling sufficiency in ecological surveys. Abstracta Botanica 22:37-48

Pillar VD (1999) How Sharp Are Classifications? Ecology 80:2508-2516

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

R Core Development Team (2010) R: A Language and Environment for Statistical Computing, http://www.r-project.org.

Ricotta C (2007) Random sampling does not exclude spatial dependence: The importance of neutral models for ecological hypothesis testing. Folia Geobotanica 42:153-160

Roleček J, Chytrý M, Hájek M, Lvončík S, Tichý L (2007) Sampling design in large-scale vegetation studies: Do not sacrifice ecological thinking to statistical purism! Folia Geobotanica 42:199-208

Roush W, Munroe JS, Fagre DB (2007) Development of a Spatial Analysis Method Using Ground-Based Repeat Photography to Detect Changes in the Alpine Treeline Ecotone, Glacier National Park, Montana, U.S.A. Arctic, Antarctic, and Alpine Research 39:297-308

Rousseeuw PJ (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics 20:53-65

Le Roux PC, McGeoch M, Nyakatya MJ, Chown SL (2005) Effects of a short-term climate change experiment on a sub-Antarctic keystone plant species. Global Change Biology 11:1628-1639

Scott WA, Hallam CJ (2002) Assessing species misidentification rates through quality assurance of vegetation monitoring. Plant Ecology 165:101-115

Scott JJ, Kirkpatrick JB (2008) Rabbits, landslips and vegetation change on the coastal slopes of subantarctic Macquarie Island, 1980 – 2007: implications for management. Polar Biology 31:409-419

Selkirk PM (2007) The nature and importance of the sub-Antarctic. Papers and Proceedings of the Royal Society of Tasmania 141:1-6

Selkirk PM, Adamson DA, Harris U (2000) Map of structural vegetation types and drainage on subantarctic Macquarie Island, Australian Antarctic Data Centre, http://data.aad.gov.au

Selkirk PM, Seppelt RD, Selkirk DR (1990) Subantarctic Macquarie Island: Environment and Biology. Cambridge University Press

References

Sellers PJ (1985) Canopy Reflectance, Photosynthesis and Transpiration. International Journal of Remote Sensing 6:1335-1372

Serreze MC, Walsh JE, Osterkamp T, Dyurgerov M, Romanovsky V, Oechel WC, Morison J, Zhang T, Barry RG (2000) Observational evidence of recent change in the northern high-latitude environment. Climatic Change 46:159-207

Shaw JD, Hovenden MJ, Bergstrom DM (2005) The impact of introduced ship rats (Rattus rattus) on seedling recruitment and distribution of a subantarctic megaherb (Pleurophyllum hookeri). Austral Ecology 30:118-125

Shaw J, Terauds A, Bergstrom DM (2011) Rapid commencement of ecosystem recovery following aerial baiting on sub-Antarctic Macquarie Island. Ecological Management & Restoration 12:1-5

Sluiter R, Pebesma EJ (2010) Comparing techniques for vegetation classification using multi- and hyperspectral images and ancillary environmental data. International Journal of Remote Sensing 31:6143-6161

Smith VR (2002) Climate change in the sub-Antarctic: an illustration from Marion Island. Climatic Change 52:345-357

Smith VR, Steenkamp M (1990) Climatic change and its ecological implications at a subantarctic island. Oecologica 85:14-24

Sparks TH, Huber K, Croxton PJ (2006) Plant development scores from fixed-date photographs: the influence of weather variables and recorder experience. International Journal of Biometeorology 50:275-9

Stevens JP, Blackstock TH, Howe EA, Stevens DP (2004) Repeatability of Phase 1 habitat survey. Journal of Environmental Management 73:53-9

Stohlgren TJ, Ma P, Kumar S, Rocca M, Morisette JT, Jarnevich CS, Benson N (2010) Ensemble habitat mapping of invasive plant species. Risk Analysis 30:224-35

Stow D, Hope A, McGuire D, Verbyla D, Gamon J, Huemmrich F, Houston S, Racine C, Sturm M, Tape K, Hinzman L, Yoshikawa K, Tweedie C, Noyle B, Silapaswan C, Douglas D, Griffith B, Jia G, Epstein H, Walker D, Daeschner S, Petersen A, Zhou L, Myneni R (2004) Remote sensing of vegetation and land-cover change in Arctic Tundra Ecosystems. Remote Sensing of Environment 89:281-308

Sturm M, Racine C, Tape K (2001) Increasing shrub abundance in the Arctic. Nature 411:546-547

Sykes JM, Horrill AD, Mountford MD (1983) Use of Visual Cover Assessments as Quantitative Estimators of Some British Woodland Taxa. Journal of Ecology 71:437-450

Sørensen R, Zinko U, Seibert J (2006) On the calculation of the topographic wetness index: evaluation of different methods based on field observations. Hydrology and Earth System Sciences 10:101-112

Tapley I, Dijkstra AH, Brolsma H (2004) The Geology, Landforms and Topography of Sub- Antarctic Macquarie Island, Australia, as revealed by AIRSAR. In: IEEE International Geoscience and Remote Sensing Symposium. IEEE Computer Society Press, Anchorage, Alaska

Taylor BW (1955) The flora, vegetation and soils of Macquarie Island, A.N.A.R.E. Reports Series B. Volume II Botany, Antarctic Division, Department of External Affairs, Melbourne, Australia.

Tkachenko K (2005) An evaluation of the analysis system of video transects used to sample subtidal epibiota. Journal of Experimental Marine Biology and Ecology 318:1-9

Townsend PA, Walsh SJ (2001) Remote sensing of forested wetlands : application of multitemporal and multispectral satellite imagery to determine plant community composition and structure in southeastern USA. Plant Ecology 157:129-149

Tweedie CE, Bergstrom DM (2000) A climate change scenario for surface air temperature at Subantarctic Macquarie Island. In: Proceedings of the VIII SCAR International Biology Symposium, Antarctic Ecosystems: Models for wider ecological understanding.pp 272-281

Vanha-Majamaa I, Salemaa M, Tuominen S, Mikkola K (2000) Digitized photographs in vegetation analysis - a comparison of cover estimates. Applied Vegetation Science 3:89-94

Vegetation Subcommittee Federal Geographic Data Committee (2002) The National Vegetation Classification Standard, Version 2.

Verrelst J, Geerling G, Sykora K, Clevers J (2009) Mapping of aggregated floodplain plant communities using image fusion of CASI and LiDAR data. International Journal of Applied Earth Observation and Geoinformation 11:83-94

Vittoz P, Bayfield N, Brooker R, Elston DA, Duff EI, Theurillat J-P, Guisan A (2010) Reproducibility of species lists, visual cover estimates and frequency methods for recording high-mountain vegetation. Journal of Vegetation Science 21:1035-1047

Vittoz P, Guisan A (2007) How reliable is the monitoring of permanent vegetation plots? A test with multiple observers. Journal of Vegetation Science 18:413-422

Wallace L (2006) GIS based snow accumulation model for the Windmill Islands. School of Geography and Environmental Studies, University of Tasmania, Studio Project

Wallentin G, Tappeiner U, Strobl J, Tasser E (2008) Understanding alpine tree line dynamics: An individual-based model. Ecological Modelling 218:235-246

Walther G-R, Beißner S, Burga CA (2005) Trends in the upward shift of alpine plants. Journal of Vegetation Science 16:541-548

Wang L, Sousa WP, Gong P (2004) Integration of object-based and pixel- based classification for mapping mangroves with IKONOS imagery. International Journal of Remote Sensing 25:5655-5668

Watts JD, Lawrence RL, Miller PR, Montagne C (2009) Monitoring of cropland practices for carbon sequestration purposes in north central Montana by Landsat remote sensing. Remote Sensing of Environment 113:1843-1852

References

Whittaker RH (1951) A Criticism of the Plant Association and Climatic Climax Concepts. Northwest Science XXV:7-31

Wilson JP, Gallant JC (2000) Digital Terrain Analysis. In: Terrain Analysis: Principles and Applications. John Wiley & Sons, New York, U.S.A., pp 225-237

Wilson AM, Silander JA, Gelfand A, Glenn JH (2011) Scaling up: linking field data and remote sensing with a hierarchical model. International Journal of Geographical Information Science 25:509-521

Wood J (1996) The geomorphological characterisation of digital elevation models. PhD Thesis, University of Leicester, UK

Xie Y, Sha Z, Yu M (2008) Remote sensing imagery in vegetation mapping: a review. Plant Ecology 1:9-23

Zevenbergen LW, Thorne CR (1987) Quantitative analysis of land surface topography. Earth Surface Processes and Landforms 12:47-56

Zhou Q, Robson M (1998) Automated cover estimation using ground digital images. In: Li D, Gong J, Chen X (eds) Spatial Information Science, Technology and Its Applications. Wuhan Technical University of Surveying and Mapping Press, Wuhan, pp 745-751

Zhou Q, Robson M, Pilsejo P (1998) On the ground estimation of vegetation cover in Australian rangelands. International Journal of Remote Sensing 19:1815-1820

Zimmermann NE, Kienast F (1999) Predictive Mapping of Alpine Grasslands in Switzerland: Species versus Community Approach. Journal of Vegetation Science 10:469-482

Økland RH (2007) Wise use of statistical tools in ecological field studies. Folia Geobotanica 42:123-140

Černá L, Chytrý M (2005) Supervised classification of plant communities with artificial neural networks. Journal of Vegetation Science 16:407-414