ARE DISTRIBUTION PATTERNS CORRELATED WITH PLANT TRAITS?

BENJAMIN MAGANA-RODRIGUEZ

Victoria
UNIVERSITY OF WELLINGTON
Te Whare Wānanga
o te Ūpoko o te Ika a Māui

A thesis submitted to the Victoria University of Wellington in partial fulfilment of the requirements for the Degree of Doctor of Philosophy in Ecology and Biodiversity

School of Biological Sciences
Faculty of Sciences
Victoria University of Wellington
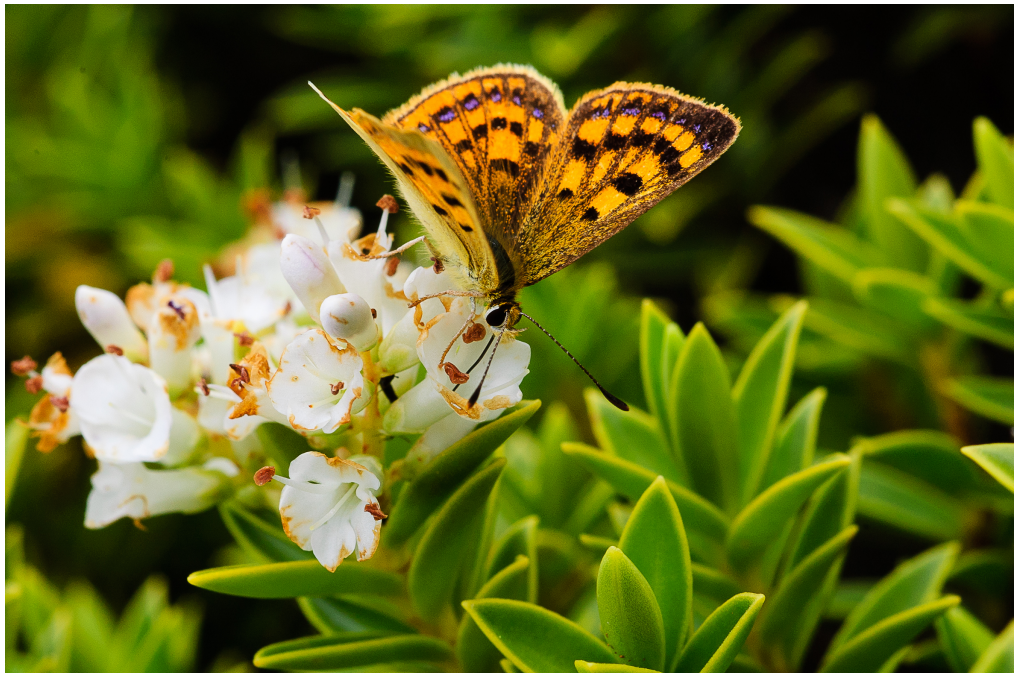April 2012

"Great intellects are skeptical"

— FRIEDRICH NIETZSCHE

Dedicated to my family, specially to my mum, dad, Miranda and Ale,
who share the pleasure of exploring this world.

ABSTRACT

The current crisis in loss of biodiversity requires rapid action. Knowledge of species' distribution patterns across scales is of high importance in determining their current status. However, species display many different distribution patterns on multiple scales. A positive relationship between regional (broad-scale) distribution and local abundance (fine-scale) of species is almost a constant pattern in macroecology. Nevertheless interspecific relationships typically contain much scatter. For example, species that possess high local abundance and narrow ranges, or species that are widespread, but locally rare. One way to describe these spatial features of distribution patterns is by analysing the scaling properties of occupancy (e.g., aggregation) in combination with knowledge of the processes that are generating the specific spatial pattern (e.g., reproduction, dispersal, and colonisation). The main goal of my research was to investigate if distribution patterns correlate with plant life-history traits across multiple scales. First, I compared the performance of five empirical models for their ability to describe the scaling relationship of occupancy in two datasets from Molesworth Station, New Zealand. Secondly, I analysed the association between spatial patterns and life history traits at two spatial scales in an assemblage of 46 grassland species in Molesworth Station. The spatial arrangement was quantified using the parameter k from the Negative Binomial Distribution (NBD). Finally, I investigated the same association between spatial patterns and life-history traits across local, regional and national scales, focusing in one of the most diverse families of plant species in New Zealand, the *Veronica* sect. *Hebe* (Plantaginaceae). The spatial arrangement was investigated using the mass fractal dimension. Cross-species correlations and phylogenetically independent contrasts were used to investigate the relationships between plant life-history traits and spatial patterns on both data bases. There was no superior occupancy-area model overall for describing the scaling relationship, however the results showed that a variety of occupancy-area models can be fit to different data sets at diverse spatial scales using nonlinear regression. Additionally, here I showed that it is possible to deduce and extrapolate information on occupancy at fine scales from coarse-scale data. For the 46 plant-assemblage in Molesworth Station, Specific leaf area (SLA) exhibits a

positive association with aggregation in cross-species analysis, while leaf area showed a negative association, and dispersule mass a positive correlation with degree of aggregation in phylogenetic contrast analysis at a local-scale (20 × 20 m resolution). Plant height was the only life-history trait that was associated with degree of aggregation at a regional-scale (100 × 60 m resolution). For the *Veronica* sect. *Hebe* dataset, leaf area showed a positive correlation with aggregation while specific leaf area showed a negative correlation with aggregation at a fine local-scale (2.5-60 m resolution). Inflorescence length, breeding system and leaf area showed a negative correlation with degree of aggregation at a regional-scale (2.5-20 km resolution). Height was positively associated with aggregation at national-scale (20-100 km resolution). Although life-history traits showed low predictive ability in explaining aggregation throughout this thesis, there was a general pattern about which processes and traits were important at different scales. At local scales traits related to dispersal and completion such as SLA , leaf area, dispersule mass and the presence of structures in seeds for dispersal, were important; while at regional scales traits related to reproduction such as breeding system, inflorescence length and traits related to dispersal (seed mass) were significant. At national scales only plant height was important in predicting aggregation. Here, it was illustrated how the parameters of these scaling models capture an important aspect of spatial pattern that can be related to other macroecological relationships and the life-history traits of species. This study shows that when several scales of analysis are considered, we can improve our understanding about the factors that are related to species' distribution patterns.

RESUMEN

La actual pérdida de biodiversidad exige actuar de manera rápida para frenarla. El conocimiento de los patrones de distribución de las especies es de gran importancia en la determinación de su estado actual. Sin embargo, las especies muestran muchos patrones de distribución en diferentes escalas. La relación positiva entre la distribución regional (a escalas gruesas) y la abundancia local (a escalas finas) de las especies, es casi un patrón constante en macroecología. Sin embargo, las relaciones intraespecíficas contienen mucha variación entre escalas. Por ejemplo, hay especies que tienen mayor abundancia local y distribución restringida, o especies que tienen amplia distribución pero

que son poco abundantes a escala local. Una forma de describir estas propiedades espaciales de los patrones de distribución es mediante el análisis de las propiedades de escalamiento de ocurrencia de las especies (por ejemplo, la agregación de especies), en combinación con el conocimiento de los procesos que están generando los patrones espaciales (por ejemplo, reproducción, dispersión y colonización). El objetivo principal de mi investigación fue analizar si los patrones de distribución tienen correlación con atributos funcionales de las plantas en diferentes tamaños de escala. En primer lugar, se comparó el desempeño de cinco modelos para determinar su capacidad para describir la relación de escalamiento de la ocurrencia de especies en dos bases de datos de Molesworth Station, Nueva Zelanda. En segundo lugar, se analizó la asociación entre los patrones espaciales a escala local y regional, así como los atributos funcionales en un ensamble de 46 especies de pastizal en Molesworth Station. La distribución espacial se cuantificó usando el parámetro *k* del modelo de distribución negativa binomial. Finalmente, se analizó la misma asociación entre patrones espaciales y atributos funcionales a escalas locales, regionales y nacionales, centrándose en una de las familias de plantas más diversas en Nueva Zelanda, la familia Plantaginaceae (*Veronica* sect. *Hebe*). La distribución espacial se analizó utilizando el índice de la geometría fractal. Se utilizaron dos tipos de correlaciones para investigar las relaciones entre atributos funcionales y patrones espaciales: correlaciones tomando en cuenta la filogenia y correlaciones sin tomar en cuenta la filogenia, en ambas bases de datos. Los resultados muestran que no hubo un modelo superior de ocurrencia-área para describir la relación de escalamiento. Sin embargo, se demostró que los modelos de ocurrencia-área pueden ajustarse a diferentes bases de datos utilizando regresión no lineal. Además, en esta tesis se demostró que es posible deducir y extrapolar información de la ocurrencia de especies a escalas finas a partir de escalas más gruesas. Para el ensamble de 46 especies de plantas en Molesworth Station, el área foliar específica (AFE) muestra una asociación positiva con la agregación de especies (sin tomar en cuenta la filogenia), mientras que el área foliar mostró una asociación negativa, y el tamaño de los propágulos mostró una correlación positiva con los patrones de distribución tomando en cuenta la filogenia a escala local (resolución de 20×20 m). La altura de las plantas fue el único atributo functional que estuvo asociado con patrones dedistribución a escala regional (resolución de 100×60 m). En las especies pertenecientes al grupo

Veronica sect. *Hebe*, el tamaño de hoja mostró una correlación positiva con la agregación de especies, mientras que el área foliar específica mostró una correlación negativa con la agregación de especies a escala local (resolución de 2.5-60 m). El tamaño de la inflorescencia, el tipo de sexualidad de las plantas y el tamaño de la hoja mostraron una correlación negativa con los patrones de distribución a escala regional (resolución de 2.5-20 km). La altura estuvo asociada positivamente con los patrones de distribución a escala nacional (resolución de 20-100 km). Aunque los atributos funcionales mostraron baja capacidad predictiva para explicar los patrones de distribución, en esta tesis se demostró cómo los parámetros de estos modelos de escalamiento pueden capturar un aspecto importante de los patrones espaciales que puede ser afines con otras relaciones macroecológicas y con los atributos funcionales de las especies. Este estudio muestra que cuando se consideran escalas de diferentes tamaños en los análisis, se puede mejorar nuestra comprensión de los factores relacionados con los patrones de distribución de las especies.

"With regard to general problems of biogeography, the biota of New Zealand has been, perhaps, the most important of any in the world. It has figured prominently in all discussions of austral biogeography, and all notable authorities have felt obliged to explain its history: explain New Zealand and the world falls into place around it"

— Nelson 1975

ACKNOWLEDGMENTS

such an isolated area. Phil Garnock-Jones was of great help in finding the raw data of *Veronica* sect. *Hebe* published in Bayly M.J. & Kellow A.V. (2006, An illustrated guide to New Zealand hebes, Te Papa press, Wellington NZ), for me to use in Chapter 4. Marc Hasenbank was of great help in sharing his knowledge in ArcGIS® and R software. Alejandra Perea and Andrea Rojas worked hard as field assistants; it was gratifying to had you both in such a demanding task, I could not have made it without your help, thank you for the laughs and good meals after finishing the day's work. A special recognition goes to Sandra Taylor and Mary Murray who know the entire Faculty from head to toes, making hassle free all paperwork at Uni. In addition, their friendship is immeasurable. I also express my gratitude to Shona De Sain for her help and support. Patricia Stain has been a keystone in providing everything a student needs for their well-being, including advices, emotional support and wake up calls from time to time. Also I want to thank all the administrative staff of the Faculty. The PhD experience was really enjoyable in the companion of my colleagues and good friends Alejandro Pérez, Erasmo Macaya, Pelayo Salinas de León, Marisol Castro, Alba Costales, Sergio Carrasco, Andrea Varela, Tyler Eddy, Daniela Díaz, Jade Berman, Philipp Neubauer, Enrique Mundaca, Mariana Lazzaro, Cesar Cárdenas, Anita Ojeda and Úrsula Rojas. Un enorme agradecimiento goes for Keko and Andrea who have been very supportive and without their help this thesis would not have been physically existed. Israel Reyes and Nataliya Soboleva constantly encouraged me during this PhD journey. Thanks for that and for the memorable chats and dinners. Alejandra Perea, my delightful motivation, with no doubt has been the bedrock of this journey and adventure that we took together. I have no words to express my gratitude for all her help, support and love given. Finally, no set of acknowledgements would be complete without mention of my immediate family, particularly my mum Socorro and my dad José, who have supported me wholeheartedly through all my life. When travelling and camping with them in remote areas to collect plants at my early age, they didn't think they were defining my vocation and interest for biology. I love you both!

<div align="right">Benjamín Magaña-Rodríguez.</div>

Morelia, Michoacán, México

CONTENTS

xi

LIST OF TABLES

ACRONYMS

AIC   Akaike's Information Criterion

ANOVA Analysis of variance

AOO   Area of Occupancy

APG   Angiosperm Phylogeny Group

CSR   Complete spatial randomness

DoC   Department of Conservation

DTM   droopy tail model

EOO   extent of occupancy

GPS   Global Positioning System

IUCN The World Conservation Union

MA   major axis

MAM   Minimum Adequate Model

MCP   minimum convex polygon

mybp million years before present

NBD   Negative Binomial Distribution

NLL   logarithm of the likelihood

NZMG New Zealand map grid

PIC   Phylogenetic independent contrast

PUE   proportion of underestimation

SAR   species area relationship

SLA   Specific leaf area

WGS84 World Geodetic System

Part I

INTRODUCTION

GENERAL INTRODUCTION

1.1 SPATIAL PATTERN

Since the time of Alfred Russel Wallace (b.1823- d.1913) and before, ecologists and biogeographers have sought explanations for the way in which plants and animals are distributed on Earth (Wallace 1876; Guisan & Thuiller 2005). Species' distribution is the empirical way in which taxa are spatially arranged. When this arrangement has a certain amount of intrinsic or extrinsic predictability, and can be described quantitatively, we use the term spatial pattern (Dale 1999). When ecologists collate information on species distribution in nature, they find it is the rule rather than the exception that most systems are not spatially homogeneous, rather they exhibit some kind of spatial pattern (Dale et al. 2002; Perry et al. 2002). The study of spatial patterns in nature seeks to understand the underlying biological processes such as establishment, competition, growth and reproduction (Brown 1984; Dale 1999; Perry et al. 2002). Spatial patterns can be thought as a snapshot of a process or a combination of processes at one given time (Fortin et al. 2003). One of the main questions of this thesis was: are processes such as competition, growth and reproduction related to spatial patterns?

1.1.1 *Types of spatial point patterns*

There are three broad classes of spatial patterns. First, the term *random pattern* refers to individuals that are distributed completely independent of each other (Figure 1). Secondly, the terms *regular, negatively autocorrelated, inhibited, uniform* and *even* refer to the type of non-random pattern that might develop as a result of mutual repulsive interactions between individuals (Perry et al. 2002; Figure 1). This inhibited pattern could arise because of competition that eliminates individuals that are too close together (Crawley 2007). Thirdly, the terms *aggregated, patchy, positively autocorrelated, clustered* and *clumped* all refer to the type of non-random pattern that might develop as a result of positive or attractive associations between individuals (Perry et

| **Random** | **Regular** | **Aggregated** |
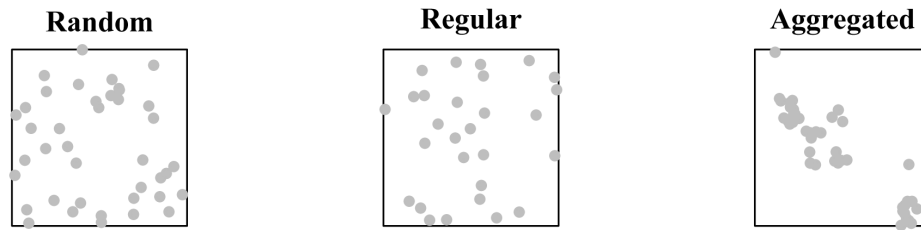
Figure 1: Three broad classes of spatial point patterns.

al. 2002; Figure 1). This may be caused by growth, limited dispersal or environmental conditions where favourable and unfavourable patches influence aggregated patterns (Dale et al. 2002; Crawley 2007). In this study I used several indices of aggregation as a measure of spatial patterns (see below).

1.1.2 *Patterns of counts across sample units*

Sample units may be naturally defined by the environment (e.g., plants with counts of insects on them) or more arbitrarily defined by the researcher by the use of quadrats or grid squares. If the data are in the form of mapped point locations (as Figure 1), they can be converted into the equivalent of grids of contiguous quadrats by setting up a matrix in which the elements of the matrix are the number of points in a corresponding square of the map (Figure 2). The use of smaller and thus more numerous quadrats will lose less information in this conversion. With data on the densities of points (e.g., species) in a grid of contiguous quadrats, the data are used directly (see Dale 1999 for a discussion on spatial analysis methods). One feature of the concepts of random, regular and aggregated patterns is that the phenomena are scale dependent. Therefore, we should not ask whether two different point patterns are scattered or aggregated; we should ask at what scale or scales they are scattered and aggregated. In the same way, throughout this thesis several scales of analysis were considered. I specifically asked, at which scale the relationships between plant life-history traits and distribution patterns are relevant?

1.2 SCALE

The term scale has been used by ecologists as the size of the sample unit (e.g., quadrat) or cell size at which a distribution is gridded (He & Condit 2007). Many species' distributions exhibit different spatial

Random

3	4	5	2
4	3	1	2
4	2	5	5

Regular

3	2	2	3
2	4	2	2
2	2	3	1

Aggregated

6	0	0	0
8	16	2	1
0	0	0	11
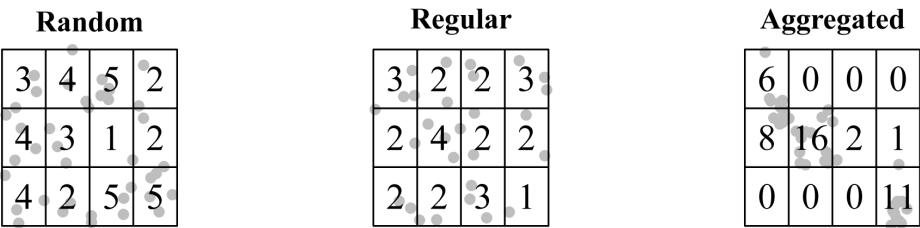
Figure 2: Three same patterns above, but as counts of individuals in gridcells.

patterns when viewed at one spatial scale compared to another; this means that they are scale-dependent (Perry et al. 2002; Hartley & Kunin 2003). For example, the same spatial pattern could appear random when viewed with small quadrats, regular when investigated with medium-size quadrats and aggregated when analysed with large quadrats (Figure 3). Scale plays a major role in this thesis, I compared species' distribution patterns from local-scale (few metres) to national-scale (hundred of kilometres).
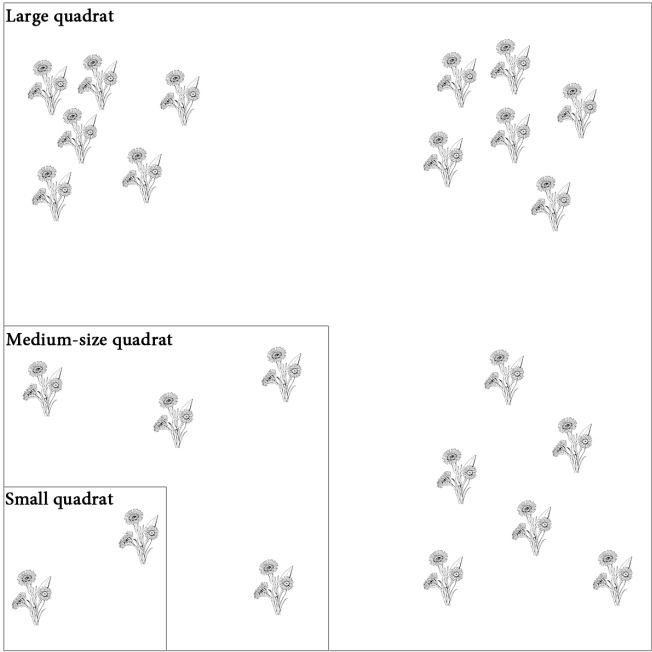
Figure 3: Spatial patterns at different quadrat sizes

1.3 MEASURING AGGREGATION

Quantifying aggregation levels of spatial patterns is essential in ecological studies to infer the underlying mechanisms (e.g., competition,

dispersal, reproductive behaviours, habitat heterogeneity, and disturb-
ances) of species coexistence (He et al. 2000; Zillio & He 2010). One
of the most widely used models for quantifying spatial aggregation
across sample units is reference to the negative binomial distribution.
(He & Gaston 2000; Green & Plotkin 2007). The negative binomial
distribution (NBD) can characterise patterns at one (arbitrary) scale,
whereby the grain of the pattern is set by the size of the sample unit.
More recently, fractals (Mandelbrot 1979) have been used for character-
ising spatial distributions (e.g., Cousens et al. 2004b; Fortin et al. 2005)
across a range of scales. Different forms of fractal dimension can be
calculated for both point patterns and distribution counts. Throughout
this thesis, spatial distribution, measured as the degree of aggregation,
was quantified using these two models.

1.3.1 *The negative binomial*

The negative binomial distribution is the statistical model distribu-
tion most frequently used to model aggregated patterns of spatial
occurrence (Holt et al. 2002; He & Gaston 2003). The NBD takes the
form

$$p = 1 - \left(1 + \frac{\mu}{k}\right)^{-k} \tag{1.1}$$

where p is occupancy, k is a "clumping" parameter and μ is the mean
abundance across sampling unit. The parameter k assumes lower
values as the degree of aggregation increases. Poisson or random
distribution is equivalent to $k \to \infty$ (Hartley 1998). The k parameter
takes the form

$$k = \frac{\mu^2}{\sigma^2 - \mu} \tag{1.2}$$

where σ^2 is the variance in the abundance of a species. The negative
binomial distribution predicts a relationship between μ and σ^2, of the
form

$$\sigma^2 = \mu + \frac{\mu^2}{k} \tag{1.3}$$

One of the drawbacks of using the NBD is that the values of k often exhibit an empirical dependency and are dependent on the mean density (Holt et al. 2002).

1.3.2   *The fractal dimension*

There are many ways of defining or describing the concept of fractal objects and fractal dimensions. The key properties of a fractal are: 1) a degree of self-similarity across a range of spatial scales (or resolutions) of observation and; 2) they do not completely occupy or fill the Euclidean dimensional space in which they are embedded e.g., a wiggly line appears to occupy more "space" on a page than a straight line, even though both lines may be of equal length and infinitely thin. For this reason fractals are usually described as possessing a non-integer dimension, as opposed to Euclidean objects which can only have integer dimensions (Halley et al. 2004). A scale-invariant pattern can be detected if the geometric measure (e.g., "space" occupied) against scale of observation (e.g., resolution of pixels) follows a power-law (or equivalently a linear relationship on log-log axes). The exponent of the power-law (or slope of the log-log plot) has the interpretation of a dimension D (Cousens et al. 2004), which measures the object's ability to fill the Euclidean space E in which it is embedded (Mandelbrot 1983). There is a wide range of methods within the fractal family differing in the information they provide, and an array of methods for their calculations. The most appropriate method for measuring a given type of fractal dimension will depend on the nature of the object and the purpose of the analysis. Generally, distributions that are point-like and highly fragmented tend to have $D \approx 0$; and distributions with continuous and even patterns tend to have $D \approx 2$ (Hartley & Kunin 2003; Figure 31). In this thesis, the box-counting dimension(D_b) and the mass fractal dimension (D_m) were used. The methods for measuring these types of fractals are explained below following Halley et al. (2004).

1.3.2.1   *Box-counting dimension*

This method is well-suited to analyse presence-absence datasets. It measures the number of "boxes" (e.g., grid cells) required to cover an object (e.g., the species distribution pattern) as a function of scale

(the size of the box). The box-counting dimension is usually found by estimating the slope (b) of log $N(s)$ plotted against log(s), where $N(s)$ is the proportional occupancy observed at a particular scale and s is the scale of analysis. Once the log-log slope (b) is calculated, we can convert the slope (b) into a box-fractal dimension (D_b) by rearranging Equation 1.4 (Hartley & Kunin 2003) and obtaining Equation 1.4a.

$$b = 1 - \left(\frac{D_b}{2} \right) \tag{1.4}$$

$$D_b = 2 - 2(b) \tag{1.4a}$$

1.3.2.2 *The mass fractal dimension*

This method is more appropriate for analysing point pattern data (Voss 1988). The method counts the number of points occupied by an object in squares or circles of radius δ of increasing size from the centre of the domain of interest. One counts the number of occupied points and estimates the mass $m(\delta)$, Equation 1.5

$$m(\delta) = \frac{N_o(\delta)}{N_t(\delta)} \tag{1.5}$$

where $N_o(\delta)$ is the number of occupied points and $N_t(\delta)$ is the total number of points within an observation window of size δ. The slope of the linear behaviour of $m(\delta)$ against δ in a log-log plot provides the estimate of the mass fractal dimension D_m. These computations are repeated for various values of δ. The mass fractal dimension D_m is defined as:

$$m(\delta) = k\delta^{D_m} \tag{1.6}$$

where k is a constant (not to be confused with k of the negative binomial distribution).

More details on the calculation of the mass fractal dimension are presented in Chapter 4 and a detailed computation of the mass fractal dimension is given in Appendix E.

1.4 OCCUPANCY AS A MEASURE OF ABUNDANCE

Due to the high cost and the logistical difficulty of doing surveys of individual counts (abundance) for studying distribution patterns, most

species data sets are available in the presence-absence form, without associated abundance count (Hui & McGeoch 2007). Occupancy, which is the number of grid cells (or sample units) in a map where a species is found, has been frequently used as a surrogate of abundance since it is easy to document, interpret and it generally correlates with species abundance (Gaston et al. 2000; He & Condit 2007). The World Conservation Union (IUCN) uses measures of occupancy to classify species as vulnerable, endangered or critically endangered if detailed information on numbers of individuals and extinction probabilities are not available (World Conservation Union 2001). Nevertheless, occupancy frequently fails to capture spatial features of distribution; two species or populations that have the same occupancy can display very different patterns (He & Hubbell 2003; Figure 4). Thus, the question is how to describe the spatial features of species on multiple scales using occupancy? Occupancy-area models can describe important features of distribution patterns across scales (e.g., aggregation). Occupancy-area models were used in this thesis and some parameters of these relationships were used as a measure of aggregation pattern.
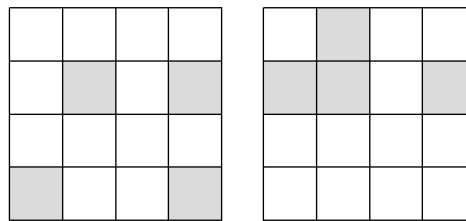
Figure 4: Two hypothetical species having the same occupancy but different distribution patterns with different levels of aggregation.

1.5 OCCUPANCY AS A MEASURE OF RANGE SIZE

Range size, sometimes defined as the extent of occupancy (EOO), is another important concept utilised by the IUCN and others as a measure of species status rank (Hartley & Kunin 2003). EOO is a coarse-scale measure of geographic range, frequently defined by the area of the minimum convex polygon (MCP) that includes all known records of the species. The value of extent of occupancy can be highly influenced by the presence of a single outlying population or individual (Hartley & Kunin 2003; Figure 5). Ranges of EOO can include regions in which the species are not actually present. This may be because those areas are entirely unsuitable, or simply uncolonised (Gaston 1991), thus EOO has the propensity to overestimate the area occupied (Figure 5).

Occupancy can also be used as a surrogate of range of size since Area of Occupancy (AOO) can be adjusted across a wide range of spatial scales. Area of occupancy is more accurate in measuring range size given that is measured by dividing the study area into a number of contiguous sample units and then summing the area of occupied units (Hartley & Kunin 2003). Figure 5 compares the concepts of Area of occupancy and extent of occupancy. The problem of using AOO is again how to capture the spatial features of distribution of two species having the same occupancy but different patterns (Figure 4)?

EOO *extent of occupancy*

MCP *minimum convex polygon*

AOO *area of occupancy*
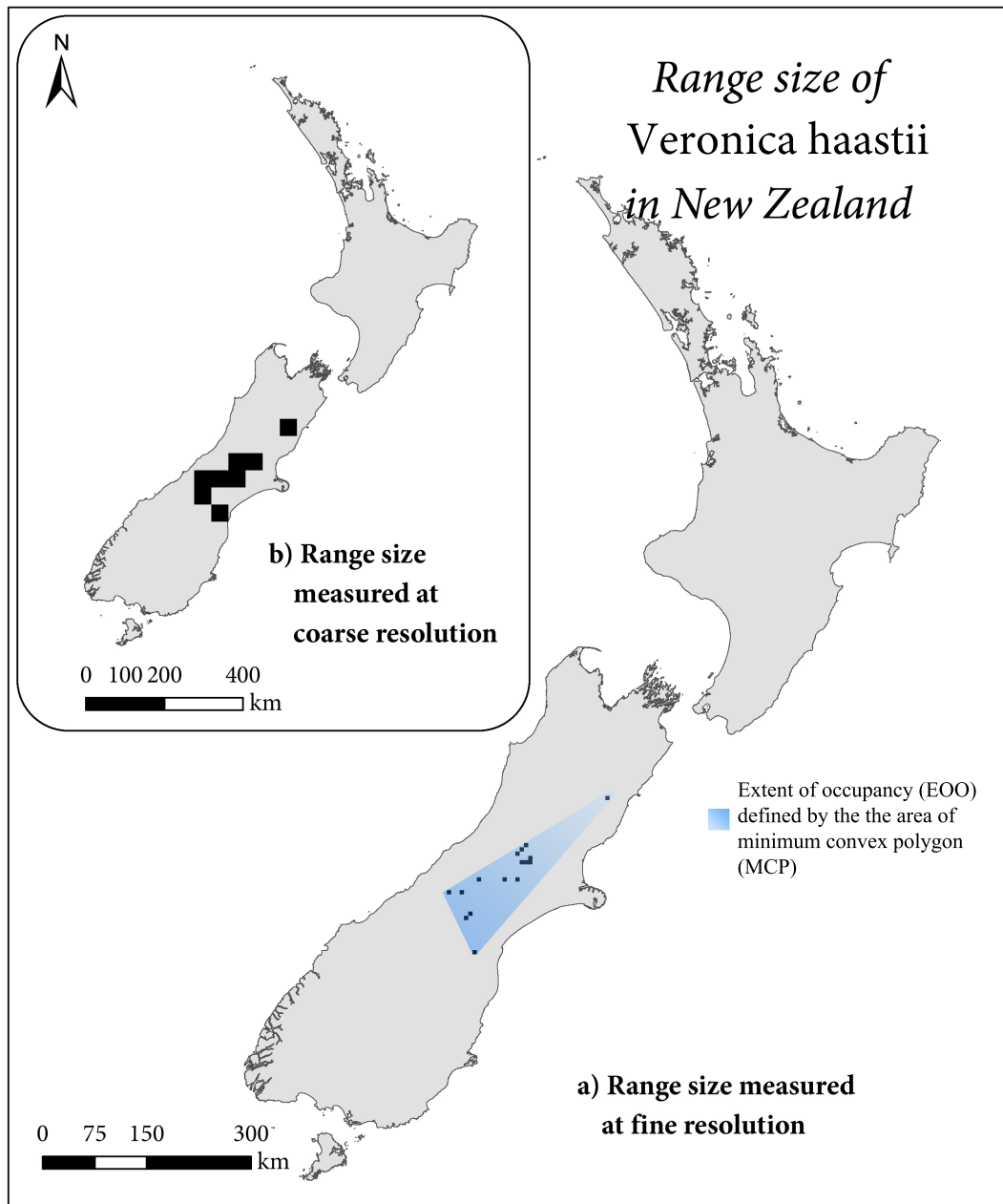
Figure 5: Area of occupancy (AOO) and the extent of occupancy (EOO) of *Veronica haastii* in New Zealand. The blue polygon represents the minimum convex polygon (MCP). The area of each cell for fine-scale is 36 km², thus the total area occupied of *Veronica haastii* is $36 \times 16 = 576$ km², for coarse-scale each cell is 1600 km², thus the total is $1600 \times 8 = 12,800$ km² and for the extent of occupancy defined by the minimum convex polygon the total area is 16,875 km². The MCP clearly overestimates range size. Area of occupancy is presented at fine and coarse-scale being a better measure of range of size. Data from Bayly & Kellow (2006).

1.6 OCCUPANCY-AREA MODELS: AGGREGATION OF OCCUPANCY PATTERNS

As illustrated above, area of occupancy (AOO) or similarly proportional occupancy, may be calculated at different scales using different size grid cells. The occupancy-area relationship is obtained by plotting species occupancy (the proportion of grid cells in which a species is found) against spatial scale or area (defined as the cell size at which a distribution is gridded; Figure 6), which yields a monotonically increasing curve whose slope is steep at first but gradually becomes flat as scale increases (Figure 6 bottom left). Several models have been proposed to examine the scaling properties of the occupancy-area relationship. The purpose of these models is to obtain a straight slope of the occupancy-area relationship leading to an invariant property of the relationship (independent of scale, e.g., Nachman model, Figure 6 bottom right). Hypothetically, once we have described the slope across scales of the occupancy-area relationship we could estimate abundance from occupancy or predict distribution at fine scales from coarse scales (Kunin 1998; He & Condit 2007; He & Gaston 2007). Examples of the occupancy-area models that have been developed to describe the properties of the spatial patterns across scales are: the power-law model, the Poisson model, the NBD model, the Nachman model and the logistic model (He & Condit 2007). Generally the relationships are plotted in their linear form (log-log), yielding some problems for empirical data of rare species where occupancy at small scales is often zero if the study area has been subsampled at the finer scales. The performance of five of these occupancy-area models was tested in Chapter 2 in their non-linear form and with occupancies of zero at fine-scales. The application of nonlinear regression to the occupancy-area models can improve occupancy prediction and can be used in surveys that have been carried out in nested quadrats or when sample units are of a mixture of sizes, not necessarily in gridded maps.
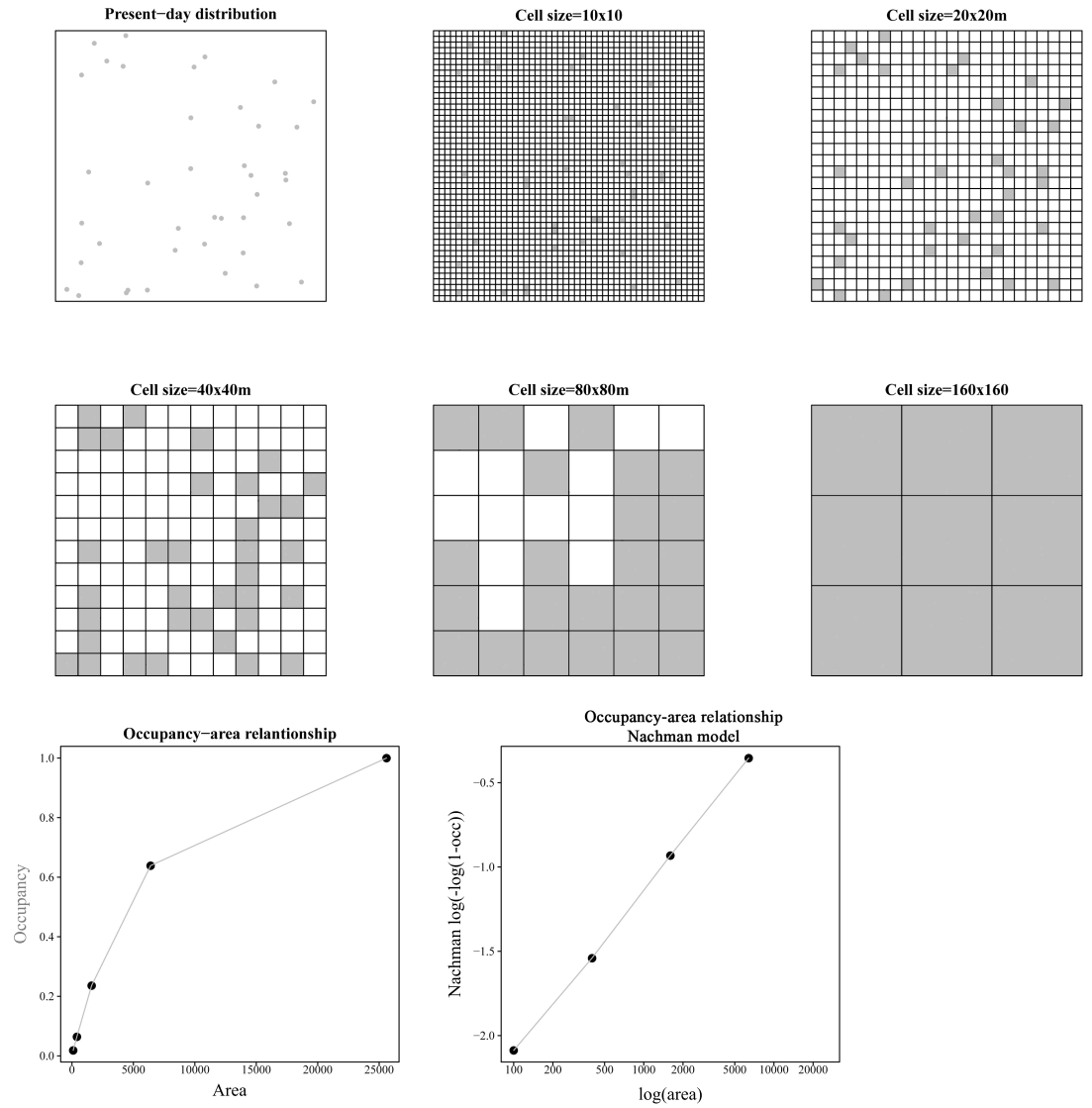
Figure 6: Hypothetical distribution of 40 points (upper left). Panels show the occurrence at five scales: 10×10, 20×20, 40×40, 80×80 and 160×160 (grid cell size or area). Bottom left plot is the relationship between the proportion of occupied area of each map and cell size (the occupancy-area curve). Bottom right plot corresponds to the same relationship but with the Nachman model, occupancy-area curve fitted in linear form.

1.7 LIFE-HISTORY TRAITS AS PREDICTORS OF DISTRIBUTION PATTERNS

The current crisis in biodiversity requires rapid action. Knowledge of species' distribution patterns across scales is important in determining their current status; whether they are rare of common species, declining or expanding. It has been suggested that several plant life-history traits might be causally related to distribution patterns (Lloyd et al. 2003). Some plant life-history traits have the advantage of being easy to measure for a large sample of species. Life-history and ecological traits (e.g., in plants; seed size, diaspore dispersal mechanism, mode of reproduction) appear to be important predictors that link organismal traits to abundance and distribution patterns (Tilman 1994; Suding et al. 2003). Recent studies have demonstrated that life-history traits are important predictors of species that become invasive or threatened (Phillips et al. 2010, Bradshaw et al. 2008, Hao et al. 2009, Fynn et al. 2009). However, the results of the different studies are far from consistent (Moles et al. 2008). Other studies have found associations between plant life-history traits and distribution patterns (Zhang et al. in press, Tremlova and Munzbergova, 2007, Van der Veken et al. 2007) and have also found contradictory results. One possible reason of these contradictory results is that these studies frequently centre their analyses at one single scale, giving different results according to the scale used. In this thesis, a multi-scale analysis approach was taken in order to address this gap; I tested the relationships between some plant life-history traits and distribution patterns at several scales.

In New Zealand, several studies have investigated the association between plant life-history traits and distribution patterns. Lloyd et al. (2003) found positive relationships between geographic range size and plant life-history traits such as growth rate, dispersal capacity and environmental tolerance; as well as negative relationships between range size and species with more variable flowering between years in species of *Chinochloa* (Poaceae). Another study that looked at attributes in fruits of species of *Acaena* (Roasaceae) in relation to distribution, found that species with barbed spines generally have the broadest geographical range and habitat distribution within New Zealand (Lee et al. 2001). In the same genus (*Acaena*) Lloyd et al. (2002) investigated growth and reproductive traits in relation to rarity and abundance. They found that common species or the species with larger range size had higher relative growth rates, faster lateral expansion, produced

inflorescences for a shorter period and held their capitula higher above the canopy. Although these studies have looked at the associations between plant life-history traits and distribution patterns, they mostly focused again on AOO at one particular scale. In this thesis, I asked: are plant life-history traits correlated with aggregation patterns? And at which scale these relationships were relevant? Multi scale-studies have the advantage of being better able to detect scale-specific associations between pattern and process that may not always be apparent at a single scale of study. They can help increase the robustness of cross-scale predictions of occupancy and to understand the relationship between abundance and occupancy.

1.8   IMPLICATIONS

Regarding occupancy-area models, the potential use of predicting properties of fine-scale distribution pattern from coarse-scale ("scaling down") is particularly useful for conservation purposes (Kunin 1998; Moody-Weis et al. 2008). In this thesis, I compared and looked at the scaling properties of several occupancy-area models. Frequently, information on invasive or rare species is needed at a fine-scale for identifying dangerously invasive or small populations; nevertheless, in most cases the information is limited to coarse scale distributional maps (Tosh et al. 2004). In conjunction, knowledge of pattern (e.g., range, abundance, and aggregation) in combination with knowledge of process (e.g., reproduction, dispersal, colonisation), is essential to predict distribution patterns across scales that may provide some guidance in the assessment or design of effective nature preserves, conducting efficient censuses, mitigating the effects of habitat loss, and predicting the effects of natural and anthropogenic disturbances (Brown 1984; Hanski & Gyllenberg 1997; Conlisk et al. 2007). Here, I correlated distribution patterns (e.g., aggregation) with processes such as reproduction, dispersal, colonisation. The results of this thesis may provide a guidance of the factors that are shaping aggregation patterns at different scales. Additionally, understanding and comparing life-history and ecological traits between rare and common species has the potential to provide useful biological information for the conservation and management of rare species (Murray et al. 2002). It has been shown that rare species tend to differ in many life-history and ecological traits from common species (Kunin & Gaston 1993; Pocock et al. 2006).

1.9 THESIS STRUCTURE

The main focus of my thesis was to identify plant life-history traits that correlate with distribution patterns across multiple scales, and to determine which plant life-history traits are important at which scales. First, in Chapter 2, I examined five occupancy-area models that may provide accurate description of distribution patterns across multiple scales because this remains a largely unsolved problem (He & Condit 2007). Additionally, I tested the performance of the same models in predicting occupancy at fine-scales from coarse-scales. In Chapter 3, I looked at the relationships between distribution patterns and plant life-history traits in a high country or sub-alpine plant assemblage in Molesworth Station, New Zealand. The NBD model was used to describe distribution pattern as it allows a straight-forward link between occupancy and local abundance (Hartley 1998). Seven morphological and life-history traits were measured and compiled from field and from specialised literature; they were: leaf area, SLA, woodiness, plant height, seed weight and dispersal mode. In Chapter 4, the national distribution of *Veronica* sect. *Hebe* in New Zealand (84 spp. from the hebe and semi-whipcord hebe clades), was quantified at two different scales and then correlated with life-history traits. Fractal dimensions were used to describe these distribution patterns, since it is a useful descriptor of the spatial aggregation of point-pattern data across scales (Storch et al. 2008). Seven plant traits were used as correlates of important processes like colonisation, dispersal and competition. I hypothesised the following relationships:

1. *Seed size* (mass). It is expected that species with larger seeds will tend to have more aggregated distributions

2. *Presence or absence of seed wings*. Species with presence of wings should have more diffuse distributions than those lacking this capacity.

3. *Breeding system*. Species that are cosexual and self-compatible could form and persist in more fragmented ranges leading to more cohesive distributions

4. *Growth habit*. Trees are expected to have relatively large and more scattered ranges.

5. *Leaf area and specific leaf area* (SLA). I predict that species with small and thicker leaves (small SLA values) will tend to have more aggregated patterns.

6. *Height*. It is hypothesised that taller species would lead to point-like distributions, since taller species could increase dispersal capacity into previously unoccupied habitats.

7. *Inflorescence length*. The hypothesis was that the species with longer inflorescences would lead to point-like distributions.

Finally in Chapter 5, the summary of the thesis is presented. Future directions and implications of the thesis in conservation, management, surveying and monitoring are discussed. The data Chapters of this thesis have been intentionally written as independent manuscripts. Some repetition of general information may occur between them.

The data Chapters will intend to be submitted to following journals:

– **Chapter 2.** Occupancy-scale relationships: a new approach for nested quadrats. *Methods in Ecology and Evolution*

– **Chapter 3.** Are plant abundance and distribution patterns correlated with ecological plant traits: a multi-scale analysis of a grassland community. *Journal of Ecology*

– **Chapter 4.** Life-history traits related to national, regional and local distribution patterns in the *Veronica* sect. *Hebe* (Plantaginaceae). *Diversity and Distributions*

1.10 REFERENCES

– Bradshaw C.J.A., Giam X.L., Tan H.T.W., Brook B.W. & Sodhi N.S. (2008). Threat or invasive status in legumes is related to opposite extremes of the same ecological and life-history attributes. Journal of Ecology, 96, 869-883.

– Brown J.H. (1984). On the relationship between abundance and distribution of species. The American Naturalist, 124, 255-279.

– Conlisk E., Bloxham M., Conlisk J., Enquist B. & Harte J. (2007). A new class of models of spatial distribution. Ecol. Monogr., 77, 269-284.

– Cousens R., Wallinga J. & Shaw M. (2004). Are the spatial patterns of weeds scale-invariant? Oikos, 107, 251-264.

– Crawley M.J. (2007). The R book. John Wiley & Sons Inc.

— Dale M.R.T. (1999). Spatial pattern analysis in plant ecology. Cambridge University Press, New York, USA.

— Dale M.R.T., Dixon P., Fortin M.J., Legendre P., Myers D.E. & Rosenberg M.S. (2002). Conceptual and mathematical relationships among methods for spatial analysis. Ecography, 558-577.

— Fortin M.J., Boots B., Csillag F. & Remmel T. (2003). On the role of spatial stochastic models in understanding landscape indices in ecology. Oikos, 203-212.

— Fortin M.J., Keitt T., Maurer B., Taper M., Kaufman D.M. & Blackburn T. (2005). Species' geographic ranges and distributional limits: pattern analysis and statistical issues. Oikos, 108, 7-17.

— Fynn, R., Wragg, P., Morris, C., Kirkman, K. & Naiken, J. (2009) Vegetative traits predict grass species' invasiveness and the invasibility of restored grassland. African Journal of Range and Forage Science, 26, 59-68.

— Gaston K.J. (1991). How Large Is a Species Geographic Range. Oikos, 61, 434-438.

— Gaston K.J., Blackburn T.M., Greenwood J.J.D., Gregory R.D., Quinn R.M. & Lawton J.H. (2000). Abundance-occupancy relationships. J. Appl. Ecol., 37, 39-59.

— Green J.L. & Plotkin J.B. (2007). A statistical theory for sampling species abundances. Ecology Letters, 10, 1037-1045.

— Guisan A. & Thuiller W. (2005). Predicting species distribution: offering more than simple habitat models. Ecology Letters, 8, 993-1009.

— Halley J.M., Hartley S., Kallimanis A.S., Kunin W.E., Lennon J.J. & Sgardelis S.P. (2004). Uses and abuses of fractal methodology in ecology. Ecology Letters, 7, 254-271.

— Hanski I. & Gyllenberg M. (1997). Uniting two general patterns in the distribution of species. Science, 275, 397-400.

— Hao, J. H., Qiangs, S., Liu, Q. Q. & Cao, F. (2009) Reproductive traits associated with invasiveness in Conyza sumatrensis. Journal of Systematics and Evolution, 47, 245-254.

— Hartley S. (1998). A positive relationship between local abundance and regional occupancy is almost inevitable (but not all positive relationships are the same). Journal of Animal Ecology, 67, 992-994.

— Hartley S. & Kunin W.E. (2003). Scale dependency of rarity, extinction risk, and conservation priority. Conservation Biology, 17, 1559-1570.

— He F. & Condit R. (2007). The distribution of species: occupancy, scale, and rarity. In: Scaling Biodiversity (eds. Storch D, Marquet PA & Brown JH). Cambridge University Press Cambridge.

— He F. & Gaston K.J. (2003). Occupancy, spatial variance, and the abundance of species. American Naturalist, 366-375.

— He F. & Gaston K.J. (2007). Estimating abundance from occurrence: an under-determined problem. The American Naturalist, 170, 655-659.

— He F.L. & Gaston K.J. (2000). Estimating species abundance from occurrence. American Naturalist, 156, 553-559.

— He F.L. & Hubbell S.P. (2003). Percolation theory for the distribution and abundance of species. Phys. Rev. Lett., 91, 4.

— He H.S., DeZonia B.E. & Mladenoff D.J. (2000). An aggregation index (AI) to quantify spatial patterns of landscapes. Landscape Ecology, 15, 591-601.

— Holt A.R., Gaston K.J. & He F.L. (2002). Occupancy-abundance relationships and spatial distribution: A review. Basic and Applied Ecology, 3, 1-13.

— Hui C. & McGeoch M.A. (2007). Capturing the "droopy-tail" in the occupancy-abundance relationship. Ecoscience, 14, 103-108.

— Kattge, J., Diaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Bonisch, G., Garnier, E., Westoby, M., Reich, P. B. & Wright, I. A. N. J. (in press) TRY–a global database of plant traits. Global Change Biology.

— Kunin W.E. (1998). Extrapolating species abundance across spatial scales. Science, 281, 1513-1515.

— Kunin W.E. & Gaston K.J. (1993). The Biology of Rarity - Patterns, Causes and Consequences. Trends in Ecology & Evolution, 8, 298-301.

— Lee W.G., Macmillan B., Partridge T., Lister R. & Lloyd K. (2001). Fruit features in relation to the ecology and distribution of Acaena (Rosaceae) species in New Zealand. New Zealand Journal of Ecology, 25, 17-27.

— Lloyd K.M., Lee W.G. & Wilson J.B. (2002). Growth and reproduction of New Zealand Acaena (Rosaceae) species in relation to rarity and commonness. New Zealand Journal of Ecology, 26, 149-160.

— Lloyd K.M., Wilson J.B. & Lee W.G. (2003). Correlates of geographic range size in New Zealand Chionochloa (Poaceae) species. Journal of Biogeography, 30, 1751-1761.

— Mandelbrot B.B. (1979). Fractals: form, chance and dimension. WH Freeman & Co, San Francisco, USA.

— Mandelbrot B.B. (1983). The fractal geometry of nature. WH Freeman and Co, New York.

— Moles, A. T., Gruber, M. A. M. & Bonser, S. P. (2008) A new framework for predicting invasive plant species. Journal of Ecology, 96, 13-17.

— Moody-Weis J., Antonovics J., Alexander H.M. & Pilson D. (2008). Predicting local colonization and extinction dynamics from coarser-scale surveys. Ecography, 31, 61-72.

— Murray B.R., Thrall P.H., Gill A.M. & Nicotra A.B. (2002). How plant life-history and ecological traits relate to species rarity and commonness at varying spatial scales. Austral Ecology, 27, 291-310.

— Perry J., Liebhold A., Rosenberg M., Dungan J., Miriti M., Jakomulska A. & Citron Pousty S. (2002). Illustrations and guidelines for selecting statistical methods for quantifying spatial pattern in ecological data. Ecography, 25, 578-600.

— Phillips, M. L., Murray, B. R., Leishman, M. R. & Ingram, R. (2010) The naturalization to invasion transition: Are there introduction history correlates of invasiveness in exotic plants of Australia? Austral Ecology, 35, 695-703.

— Pocock M.J.O., Hartley S., Telfer M.G., Preston C.D. & Kunin W.E. (2006). Ecological correlates of range structure in rare and scarce British plants. Journal of Ecology, 94, 581-596.

— R Development Core Team (2011). R: A Language and Environment for Statistical Computing, Vienna, Austria. URL http://www.R-project.org

— Storch D., Sizling A.L., Reif J., Polechova J., Sizlingova E. & Gaston K.J. (2008). The quest for a null model for macroecological patterns: geometry of species distributions at multiple spatial scales. Ecology Letters, 11, 771-784.

— Suding K.N., Goldberg D.E. & Hartman K.M. (2003). Relationships among species traits: Separating levels of response and identifying linkages to abundance. Ecology, 84, 1-16.

— Tilman D. (1994). Competition and Biodiversity in Spatially Structured Habitats. Ecology, 75, 2-16.

— Tosh C.A., Reyers B. & van Jaarsveld A.S. (2004). Estimating the abundances of large herbivores in the Kruger National Park using presence–absence data. Animal Conservation, 7, 55-61.

— Tremlova, K. & Munzbergova, Z. (2007) Importance of species traits for species distribution in fragmented landscapes. Ecology, 88, 965-977.

— Van der Veken S., Bellemare J., Verheyen K. & Hermy M. (2007). Life-history traits are correlated with geographical distribution patterns of western European forest herb species. Journal of Biogeography, 34, 1723-1735.

— Voss R.F. (1988). Fractals in nature: from characterization to simulation. Springer-Verlag New York, USA.

— Wallace A.R. (1876). The geographical distribution of animals: With a study of the relations of living and extinct faunas as elucidating the past changes of the earth's surface. Macmillan and Co.

– World Conservation Union (2001). IUCN Red list Categories and Criteria (version 3.1). World Conservation Union, Gland, Switzerland and Cambridge, United Kindom.

– Zhang, S. B., Slik, J., Zhang, J. L. & Cao, K. F. (in press) Spatial patterns of wood traits in China are controlled by phylogeny and the environment. Global Ecology and Biogeography.

– Zillio T. & He F.L. (2010). Modelling spatial aggregation of finite populations. Ecology, 91, 3698-3706.

Part II

DATA CHAPTERS

OCCUPANCY-AREA RELATIONSHIPS: A NEW APPROACH FOR NESTED QUADRATS

2.1 INTRODUCTION

The distribution of species in broad scale studies is commonly recorded by occupancy because measures of abundance (number of individuals) are too costly and logistically difficult to carry out. Occupancy is an important criterion used for inventorying biodiversity (Gaston et al. 2000), categorising conservation status IUCN and assessing spread of invasive species (Hamilton et al. 2005).

Occupancy is often referred to as the presence of a species in a sample unit, grid cell, site or some other defined locality (Hui and McGeoch, 2007, Hartley, 1998). Here, I use the term sample unit which implies the data were derived from some form of structured survey. The proportion of sample units that is occupied by a species is a function of the overall abundance of the species (e.g., number of individuals), their degree of aggregation and the size (also referred to as the area, scale, or resolution) of the sample units (He and Gaston, 2000a). Occupancy measured at a single scale often fails to capture significant spatial features (He and Condit, 2007), however occupancy measured over multiple scales can reveal spatial properties such as aggregation (He and Condit, 2007). The relationship between area of the sample unit and occupancy has also captured ecologists' attention because it is a promising way to predict abundance and distribution across multiple spatial scales (He and Condit, 2007, Kunin, 1998).

Several mathematical models have been suggested to describe the occupancy-area relationship. Although the majority have been originally proposed in the form of occupancy-abundance models (see Holt et al. 2002 for a review), they can be adapted in the form of occupancy-area relationship.

In a recent paper He & Condit (2007) have tested the fit of three models for abundance-area relationships: the Nachman, the logistic and the power-law models. Although no model was overall superior, they concluded that the most robust model was the Nachman model. The occupancy-area relationship is commonly analysed in linear form (on log-log axes) drawing data from commonly used grid maps (e.g.,

the Barro Colorado Island 50-ha permanent forest plot) in which coarse scale occupancy maps are produced by amalgamating the adjacent grid cells of a spatially comprehensive fine-scale survey. The log-log transformations of the occupancy-area relationships often distort the experimental error since the linear regression assumes that the scatter of points around the line follows a Gaussian distribution an that the standard deviation is the same at every value of X (Motulsky and Christopoulos, 2004). These assumptions are rarely true after transforming the data.

Here, I provide a new tool for analysing the scaling properties of occupancy where data is not transformed and can also contain values of zeroes at fine resolutions. This study is one of the first attempts to fit nonlinear regression to the occupancy-area relationship. Non-linear curve fitting is often considered a better approach than linear regression of transformed values due to a more accurate error distribution of residuals.

Many countries have adopted different plant survey methods. For instance, New Zealand has a series of permanent nested quadrats along the country (Figure 8) to assess alpine grassland conservation status, biodiversity and the impact of invasive species (e.g., *Hieracium pilosella*, *Rosa rubiginosa*) (The National Vegetation Survey Databank, Landcare Research). Such survey methods collect occupancy information at multiple scales, but often the area covered by the fine-scale sample units is only a sub-sample of the area covered by coarser-scale sample-units. In such types of nested quadrats it becomes problematic to fit occupancy-area models in their linear form. It is highly likely that some species will be recorded only in the larger quadrats, and will be completely absent from the sub-sample of small quadrats used to estimate fine-scale occupancy. The problem arises because the logarithm of zero is negative infinite. In this study I demonstrate that it is feasible to fit nonlinear models to data derived from this type of nested quadrat survey, thus overcoming the problem of zeroes in the linear (log-log) regression.

The aims of this Chapter were to test the performance of five of the most common occupancy-area models in their non linear form, and to determine how well they predict occupancy at fine scales from the crude data of coarse-scale occupancy. The application of these models in vegetation surveys with nested quadrats could have the potential of predicting occupancy at scales smaller and larger than the original survey, thus saving time and money. Such information is needed for

accurately categorising species status and for planning conservation priorities.

2.2 METHODS

2.2.1 *Study site*

The models were tested using data on the distribution of sub-alpine species in indigenous grassland at local to regional scales in Molesworth Station, New Zealand. In 2006 and 2007, 80 nested grassland quadrats (Molesworth I) were established by the Department of Conservation in this area with the objective of analysing plant structure, richness and assessing the invasiveness of exotic species (S. Husheer unpub. data). Twenty sampling points were randomly selected. In each point four 20×20 m quadrat were place; each quadrat separated from each other a distance of 150 m in north direction. Another set of 27 larger nested quadrats (Molesworth II) were set up by B. Magana between 2007 and 2010 in the same general area (Figure 7; Table 1). Plots were randomly placed with restrictions (within a buffer area of 3 km on each side of the gravel road due to the inaccessibility of some remote areas).

Molesworth station is situated in the South Island, NW of the Inland Kaikoura Mountains in South Marlborough. It occupies a great part of the headwater catchments of the Awatere River flowing north-east through Marlborough, and of the Acheron River and other streams which drain from the same side into the upper Clarence River (Moore, 1976). The average annual rainfall of Molesworth station ranges from less than 760 mm in the East to more than 2,540 mm in the West (Coulter, 1969). Long dry periods occur at all times of the year, but especially in summer and autumn. The average temperature varies from 14 °C in summer to 10 °C in winter. Air frosts occur in all months and ground frosts are experienced on 218 nights per annum on average (Coulter, 1969). Vegetation reflects this range of rainfall. It is common to find gravelfield and scree communities in dry places and in places where the rainfall is heavier, exotic grassland plants and short native tussocks are common. The range of elevation for the quadrats set up was from 700 to 1460 m.
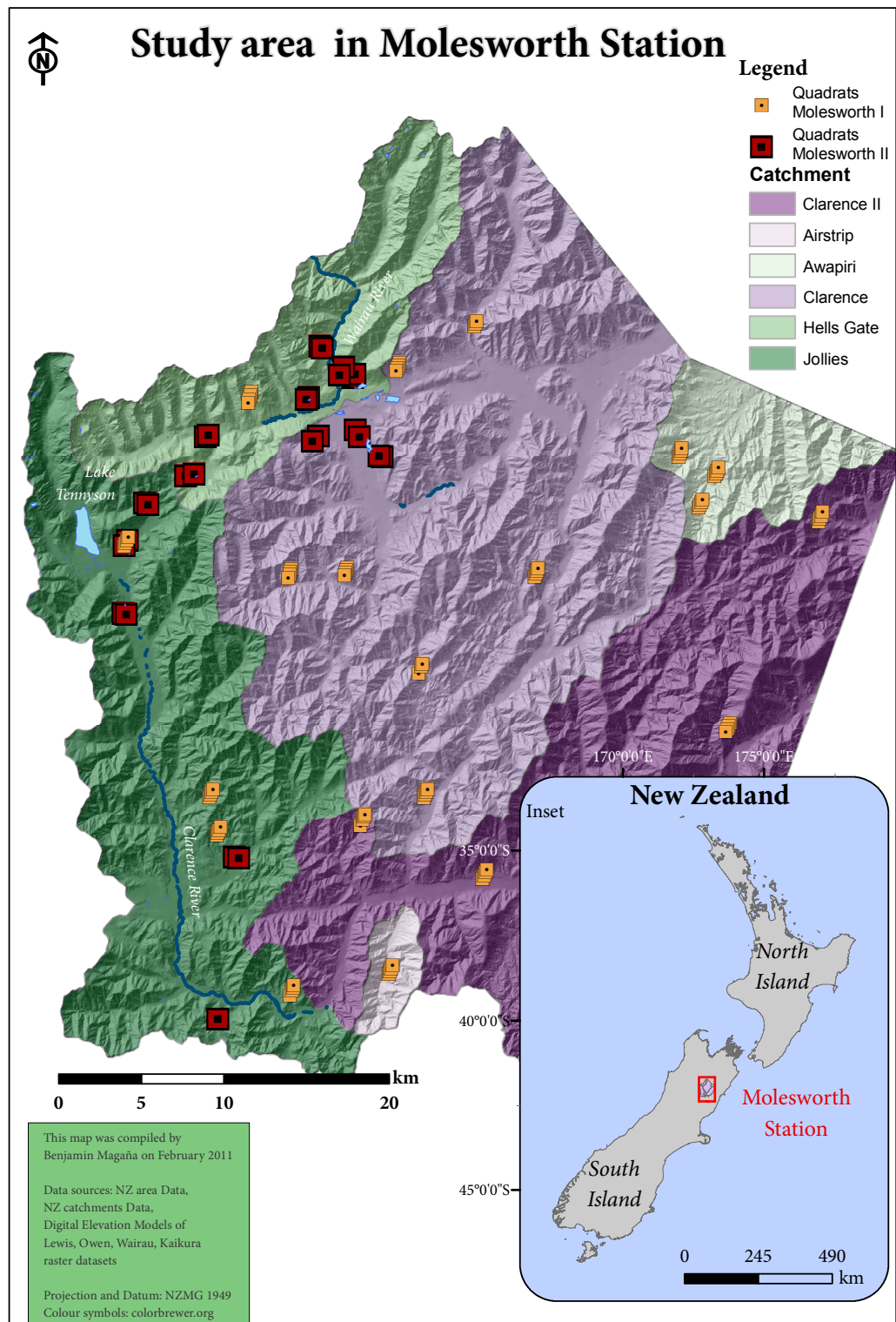
Figure 7: Study area and plot location of the Molesworth I and Molesworth II data sets.

Table 1: Scales and number of quadrats sampled for Molesworth I and Molesworth II data set.

| | MOLESWORTH I (299 spp.) | | | | | | MOLESWORTH II (48 spp.) | | | |
| | Scale (m) | | | | | | Scales (m) | | | |
	20 x 20	10 x 10	5 x 5	2 x 2	1 x 1	0.5 x 0.5	100 x 60	20 x 20	5 x 5	1 x 1
Total area (m^2)	400	100	25	4	1	0.25	6000	400	25	1
Number of quadrats	80	80	80	80	80	80	27	405	432	432

2.2.2 *The negative binomial, power-law, Nachman and logistic models*

There are five models commonly used for describing the occupancy–area
relationship (Table 2). These five models were first developed for differ-
ent purposes, for example: Kunin (1998) proposed a linear (power law),
scale-occupancy curve and presented the slope of the curve as a scale-
independent measure of species abundance. Nachman (1981) provided
an exponential relationship based on the change in the proportion of
empty patches, which does not make any assumptions about the form
of the underlying theoretical spatial distribution. The logistic model
was initially used for modelling the species–area relationship (Hanski
and Gyllenberg, 1997). He and Gaston (2000b) adopted the negative
binomial distribution to obtain an occupancy–abundance relationship
which can be expressed as a two-parameter occupancy-area model as
below (NBD2). This model was further improved and simplified to a
one-parameter occupancy-area model, NBD1, by He and Condit (2007).
The models have been adapted to obtain the occupancy-area curves
(Table 2), where *occ* is occupancy (the proportion of sample units
occupied), *a* is the area of a single sample unit (quadrat size), and *c*, *z*
and *k* are parameters. Occupancy is measured as a proportion (hence
in the range 0-1), the total area (A) of the study site is 1 and then is
removed from the original formulation of He and Gaston (2000a, b).
This causes the parameter N, in the second negative binomial model
to be proportional to the estimated number of individuals across the
total study area rather than a direct estimate, unless the areas of the
quadrats are also expressed as a proportion of the total study area. The
estimated parameters *k* and *z* are generally regarded to reflect degree
of aggregation or positive spatial autocorrelation among occupied
quadrats.

Table 2: Occupancy-area relationship models.

MODEL	OCCUPANCY-AREA RELATIONSHIP	PARAMETER(S) ESTIMATED	REFERENCE
Negative Binomial (NBD1)	$occ = 1 - \left(1 + \frac{a}{k}\right)^{-k}$	1; k	(He and Condit, 2007)
Negative Binomial (NBD2)	$occ = 1 - \left(1 + \frac{N \cdot a}{k}\right)^{-k}$	2; N, k	(He and Gaston, 2000b)
Power-law	$occ = c \cdot a^z$	2; c, z	(Kunin, 1998)
Nachman	$occ = 1 - e^{-c \cdot a^z}$	2; c, z	(Nachman, 1981)
Logistic	$occ = \frac{c \cdot a^z}{1 + c \cdot a^z}$	2; c, z	(Hanski and Gyllenberg, 1997)

2.2.3 *Plot design*

Two data sets were used to evaluate the performance of the five models. The data sets differ in the way the nested quadrats were arranged: The first dataset comprises the local distribution of 229 plant species (Molesworth I) across 80 quadrats 20 × 20 m in size, containing nested quadrats of five smaller sizes ranging from 0.25 to 100 m² (Figure 8a). The second dataset comprises the local to regional distribution of 48 species (Molesworth II), across 27 large quadrats 60 × 100 m in size, containing samples from sub-quadrats of three smaller sizes from 1 to 400 m² (Figure 8b).

2.2.4 *Data analysis*

Occupancy-area curves of the five models were fitted for each species from the two datasets (example of six species in Figure 9), by using R v.2.10 software (R Development Core Team, 2011) with the nlme package and the non-linear least squares function nls (Pinheiro et al. 2010). The R script to fit the different curves was developed by Hartley and Magana-Rodriguez. See Appendix A for details of computation of these five models in R. The 'port' optimisation method was used and parameter values were constrained within sensible limits. For some data sets it was necessary to try alternative start values for the parameter estimates before the model would converge on a solution. Four criteria were used to measure the goodness-of-fit of the models. The first three were the residual variance (σ^2), Akaike's Information Criterion (AIC) and the logarithm of the likelihood (NLL). The residual
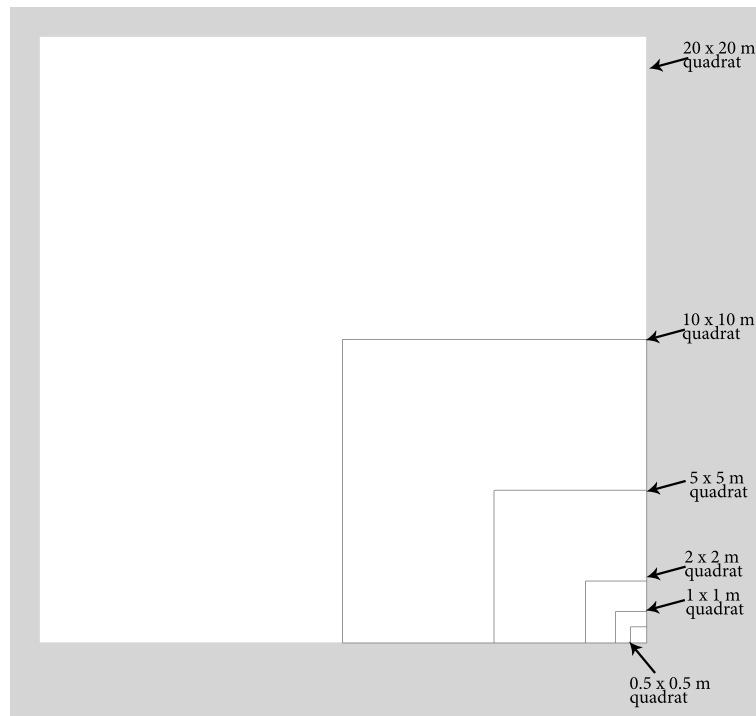
(a) Molesworth I. Each quadrat had nested quadrats of six sizes.
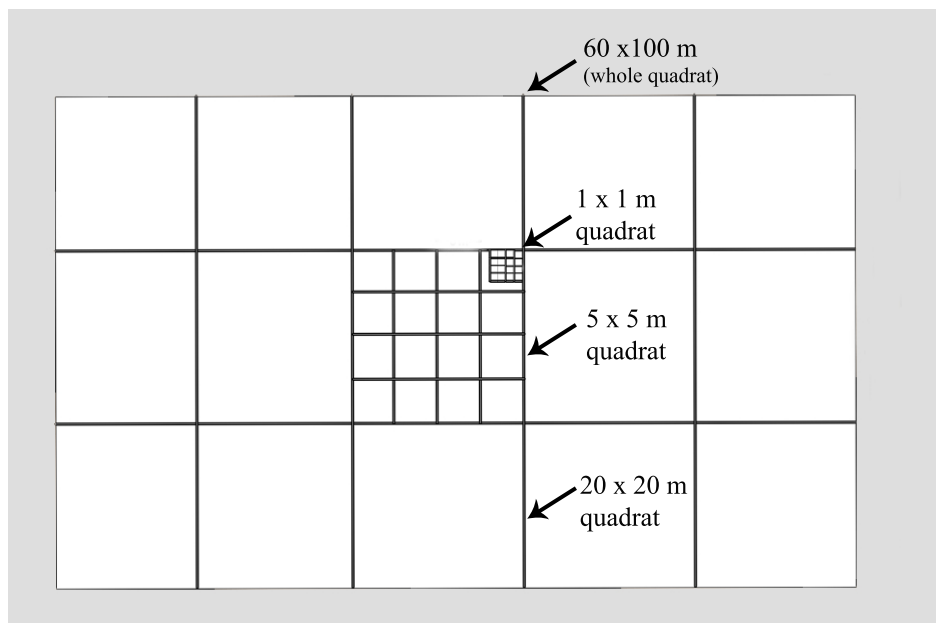
(b) Molesworth II. Each quadrat had nested quadrats of four sizes.

Figure 8: Layout of sample quadrats used to measure subalpine plant species structure and richness in Molesworth Station.

variance σ^2 is obtained as the minimum value of the residual sum of squares (RRS) divided by the degrees of freedom ($n - p$), giving the estimate

$$\sigma^2 = \frac{RRS(\beta)}{n - p} \tag{2.1}$$

where RRS is defined by:

$$RRS(\beta) = \sum_{i=1}^{n} (y_i - f(x_i, \beta))^2 \tag{2.2}$$

where $x_1, ..., x_n$ are the predictor values of the model and $y_1, ..., y_n$ the response values and a given mean function f depending on some unknown parameters. The minimisation of RSS is often referred to as minimising the least-squares criterion or least-squares estimation, and the solution to the minimisation problem leads to the least-squares parameter estimates (Ritz and Streibig, 2008). The fourth criterion was to rank the five models from best to worst (the lowest σ^2) and to report how often a model was judged as the best model, in the top two or in the top three across all species for Molesworth I and Molesworth II data (229 and 48 species, respectively).

For predicting occupancy at small-scales (0.25 m^2) from coarse-scales, five (1, 4, 25, 100 and 400 m^2), four (4, 25, 100 and 400 m^2) and three scales (25, 100 and 400 m^2) were used for Molesworth I. In Molesworth II data, three coarser scales (25, 400, 6000 m^2) were used to predict occupancy at the finer scale of 1 m^2.

To estimate the predictive performance of the models the relationship between the observed and predicted occupancy was compared with three measures; the root-mean-square error ($RMSE$), the proportion of underestimation (PUE) and the correlation coefficient (r). The $RMSE$ measures the differences between values predicted by the model and the values actually observed:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \tag{2.3}$$

where y_i is the observed value for the ith observation and $\hat{y}_i$ is the predicted value.

The PUE is simply calculated from the differences between observed and expected values ($y_i - \hat{y}_i$). Each positive value represents an un-

derestimation, PUE is then the proportion of observations generating positive values. Non-biased predictors should generate a PUE of 0.5.

The correlation coefficient (r) indicates the strength of the linear relationship between predicted and observed occupancy but by itself cannot distinguish between a perfect predictive model and one that under or over-predicts by a constant ratio, hence the correlation coefficient was used in combination with the other two measures of predictive ability.
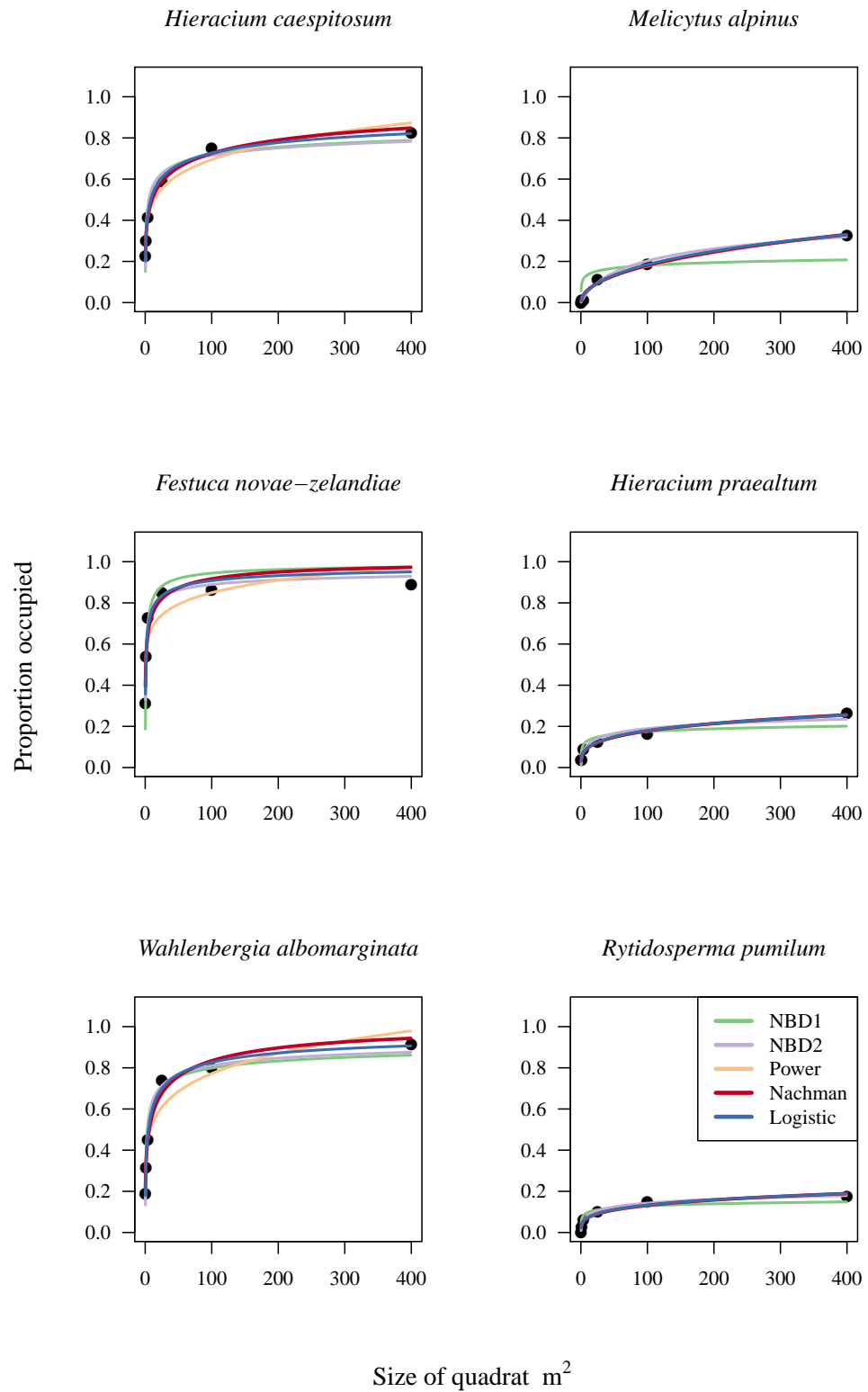
Figure 9: Nonlinear occupancy-area relationship curves for six Molesworth I species. Different colours represent different models. The right plots show the performance of the models with species recorded to have occupancy of zero at fine scale.

2.3 RESULTS

2.3.1 *Empirical evaluation of the models*

For Molesworth I data, the logistic model gave the best average fit (Figure 10). However the second negative binomial model (NBD2) was more often the highest ranking model and the Nachman was most consistently in the set of top 3 models (Table 3). In the Molesworth II dataset the power-law relationship was most often the highest ranking model (Table 3) and the Nachman model gave the best fit on average (Figure 11). By every measure, the negative binomial model 1 (NBD1) produced the worst fits overall.

2.3.2 *Estimating occupancy at fine scales from the occupancy at coarse scales*

The best four models for predicting occupancy at small scales were plotted against the observed occupancy values (Figure 12). The best model for predicting occupancy at small scales, measured as the lowest *RMSE*, PUE (unbiased model = 50%) and large *r* values, varied among the number of scales used for predicting occupancy at small scales and between the Molesworth I and Molesworth II datasets. When analysing Molesworth I data set, the best model for predicting occupancy using five resolutions was the logistic model, although the difference in values of *RMSE* and *r* with the Nachman model was minimal (Figure 12a and 12b; Table 4). When using four resolution scales, the Nachman model was the best and when using three scales for predicting occupancy, the lowest *RMSE* and the largest *r* came from the power-law model (Figure 12c; Table 4). The PUE was constant across the number of scales used for predicting occupancy, and suggested that the best unbiased model is the NBD2. The lowest *RMSE* values were estimated when using five scales, this result indicates that the prediction was more accurate as the number of scales increases (Figure 26a).

PUE *proportion of underestimation*

In the Molesworth II dataset, the Nachman model had the lowest *RMSE* and the largest *r* followed by the power-law. The PUE suggested the logistic model was the least biased estimator (Figure 12d, Table 4).

Table 3: Number of species for which each model was judged as best, in the top 2 or top 3 best models (i.e. cumulative ranks).

MODEL	MOLESWORTH I (229 spp)							MOLESWORTH II (48 spp)						
	Mean σ^2	Median σ^2	Mean AIC	Median AIC	1st	top 2	top 3	Mean σ^2	Median σ^2	Mean AIC	Median AIC	1st	top 2	top 3
NBD1	0.0237	0.0148	-29.9	-30.6	5	19	23	0.0863	0.0814	-7.4	-5.8	0	2	4
NBD2	0.0107	0.0058	-37.8	-39.6	**132**	**138**	139	0.0213	0.0150	-18.8	-16.3	11	16	20
Power-law	0.0125	0.0068	-37.2	-39.1	47	52	89	0.0255	0.0068	-24.2	-19.2	**17**	22	28
Nachman	0.0108	0.0067	-37.8	-39.3	6	91	**211**	**0.0179**	0.0142	**-25.0**	-19.5	8	**33**	**46**
Logistic	**0.0101**	0.0067	**-38.2**	-39.4	7	126	193	0.0191	0.0144	-24.7	-19.2	12	23	**46**
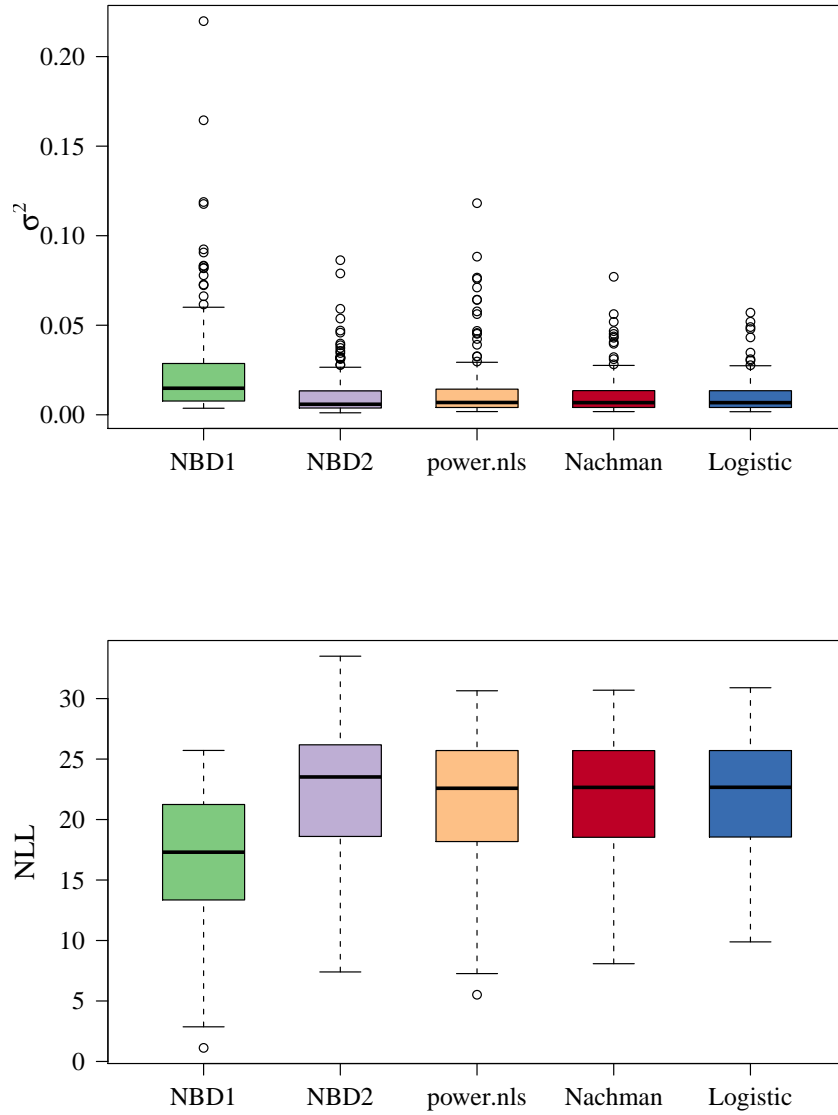
Figure 10: Boxplots of σ^2 and the logarithmic likelihood (NLL) showing the goodness-of-fit of the five models fitted to each of the 229 species distribution of Molesworth I data. Central lines in bold indicate the median value, boxes depict the interquartile range, whiskers depict trimmed ranges and circles are outliers. Low σ^2 and large NLL suggest good fit of the model.
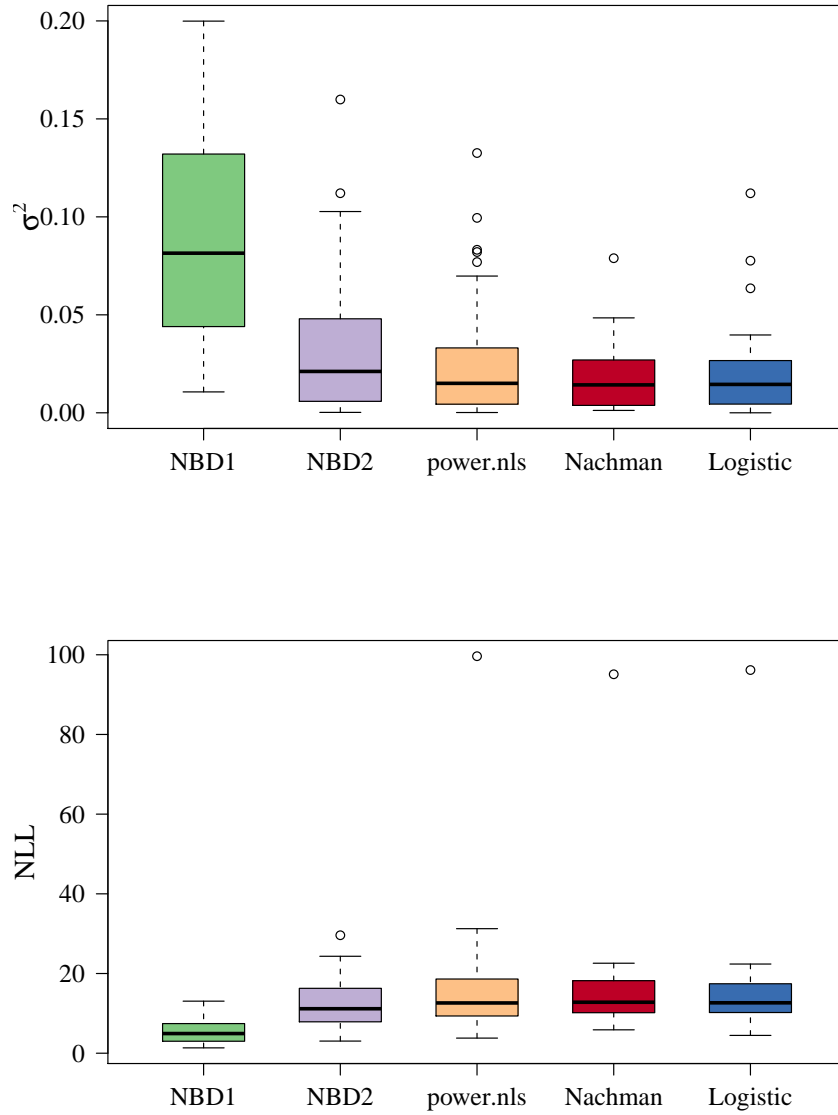
Figure 11: Boxplots of σ^2 and the logarithmic likelihood (NLL) showing the goodness-of-fit of the five models fitted to each of the 48 species distribution of Molesworth II data. Central lines in bold indicate the median value, boxes depict the interquartile range, whiskers depict trimmed ranges and circles are outliers. Low σ^2 and large NLL suggest good fit of the model.
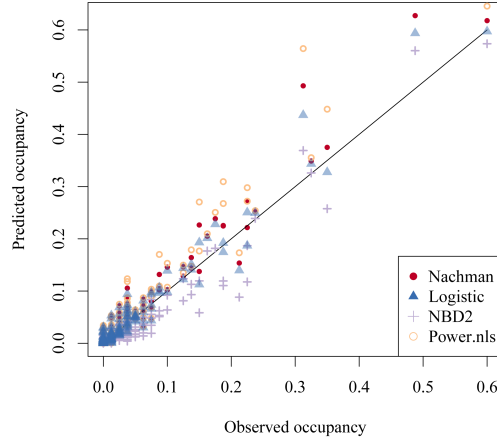
(a) Predicted *occ* at 0.25 m²with 5 resolutions
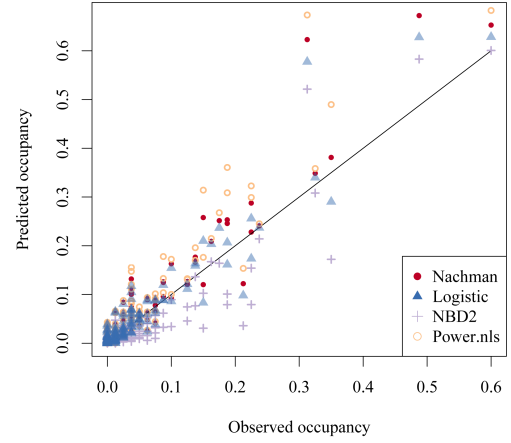
(b) Predicted *occ* at 0.25 m² with 4 resolutions
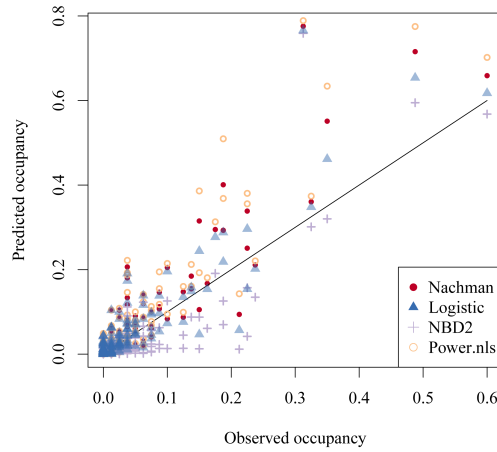
(c) Predicted *occ* at 0.25 m² with 3 resolutions
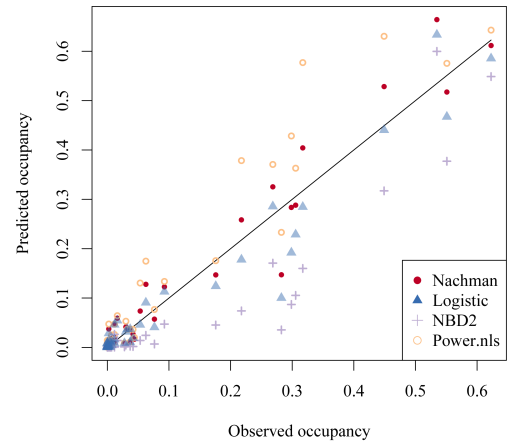
(d) Predicted *occ* at 1 m² with 3 resolutions

Figure 12: Estimated occupancy versus observed occupancy at 0.25 m² scale in plots (a), (b) and (c) in Molesworth I dataset. Plot (d) is the estimated occupancy versus observed occupancy at 1 m² scale in Molesworth II dataset. Red dots represent the Nachman model, blue triangles the logistic model, purple crosses the negative binomial 2 and orange circles the power-law model. The black line is the observed occupancy where the models should lie if accurate.

Table 4: The root-mean-square error (*RMSE*), proportion of underestimation (PUE) and correlation coefficient (*r*) for the Nachman, logistic, negative binomial 2 and the power-law models for estimating occupancy at 0.25 m² and 1 m² in Molesworth I and Molesworth II datasets respectively. Three, four and five resolutions were used to predict occupancy in Molesworth I and only three resolutions were used to predict occupancy in Molesworth II. Low *RMSE* suggest good fit of the model, PUE should attain the nominal level of 50%. Large values of *r* indicate a strong relationship.

MODEL	MOLESWORTH I (229 spp)									MOLESWORTH II (48 spp)		
	3 resolutions			4 resolutions			5 resolutions			3 resolutions		
	RMSE	PUE	*r*	*RMSE*	PUE	*r*	*RMSE*	PUE	*r*	*RMSE*	PUE	*r*
Nachman	0.0298	17.1	0.914	**0.0216**	17.9	**0.957**	0.0158	17.9	0.977	**0.0354**	35.4	**0.977**
Logistic	0.0318	20.0	0.904	0.0231	21.4	0.951	**0.0154**	21.3	**0.978**	0.0416	**42.5**	0.968
NBD2	0.0422	**42.6**	0.842	0.0316	**43.2**	0.910	0.0202	**41.1**	0.963	0.0680	66.2	0.917
Power-law	**0.0288**	15.7	**0.922**	0.0222	14.4	0.954	0.0177	15.7	0.971	0.0413	22.9	0.968

2.4 DISCUSSION

Measures of occupancy-area relationships are important in identifying distribution patterns across scales. The identification of distribution patterns is considered the first steps towards understanding the processes that are generating these distribution patterns. The use of the novel approach suggested here to identify distribution patterns on multiple scales, has the potential of being used in data obtained in a series of nested quadrats or in different quadrat sizes or resolution within a study area. The measurement of the occupancy-area curves in these types of nested quadrats had been remained elusive so far by the conventional linear methods. Additionally, in this Chapter I demonstrate that predicting occupancy at fine scales from coarse scales is achievable with the use of these five models in their nonlinear form.

In this study, it has been shown that a variety of occupancy-area models can be fit to different data sets at diverse spatial scales using nonlinear regression. Although there was no superior occupancy-area model overall for describing the scaling relationship, all models apart from the negative binomial model 1 (NBD1) performed well in their nonlinear form. The logistic and the Nachman model were the most robust with the lowest σ^2 on average in the Molesworth I and Molesworth II, respectively. The Nachman model was the best model for predicting occupancy at small-scales from coarse scales (except when three resolutions were used in Molesworth I). These results obtained using nonlinear regression are consistent with what He and Condit (2007) found when comparing three of the occupancy-area models mentioned above but using transformed data into a linear regression. He et al. (2002) also found similar results when comparing the same five models tested in this study plus the unified model for predicting the occupancy-abundance relationship. They concluded that the best models to use were the Nachman for bird species and the logistic for tree species. In a more recent study Hui and McGeoch (2007) also tested these five models plus a model that improves the estimate of occupancy at coarse scales, the droopy tail model (DTM). Their results showed that the NBD2 and the droopy tail model performed better than the other models. In conclusion, there is no overall superior model. The advantage of fitting these models in nonlinear regression is that they can be fitted even if values of zero occupancy are present at fine scales.

NBD *negative binomial model*

One of the advantages of the NBD2 over the other models in estimating occupancy is that it provides a direct estimate of abundance N (two parameters estimated: k = aggregation, N = total abundance; Table 2). The relationship of the occupancy-area curves can be scaled down using a sufficiently fine scale so that each occupied cell (area or resolution) contains a single individual. Thus, in theory, being able to predict abundance from occupancy, although extra information on the average size of an individual is needed (Hartley and Kunin, 2003, He and Gaston, 2000a). Measuring abundance is important because it helps to determine at-risk populations, guide conservation decisions and further knowledge of species natural histories (Conlisk et al. 2009). It is worth mentioning that estimating number of individuals N doesn't always make sense for plant species with modular growth and vegetative reproduction. Thus, the use of the other models to estimate occupancy at fine scales is well based too. The occupancy-area models can be applied in the opposite direction; there are some cases in which one might want to scale-up from fine-scale data to coarse-scale data. In forestry, for example, census at fine-scale among individual plants could predict forest total occupancy at a particular extent which is extremely desirable (Rastetter et al. 1992). Another example of the application on scaling up is to predict human alteration in natural environments. These global alterations (e.g., reduction in species occupancy) could be predicted from short-term studies made on small plots (less than 100×100 m) (Rastetter et al. 2003).

RMSE *root-mean-square error*

Prediction of occupancy is feasible when using nonlinear fitting to the occupancy-area models. The estimation improved when taking into account more quadrats at different sizes (scales) for predicting occupancy at finer-scales. A minor drawback when fitting models in their nonlinear regression form with the `nls` function, is that at least three resolutions are needed to make them work (to avoid the situation of zero residuals), compared to the linear regression where only two resolutions are needed to make predictions. Contrary to the *RMSE*, the proportion of underestimation (PUE) remained remarkably constant independent of the scales used for predicting occupancy at fine scales. There was a regular underestimation by the second negative binomial (NBD2) and an overestimation by the power-law model. These results are consistent with Kunin et al. (2000). They applied the NBD2 to a data set of scarce British plants and found that it consistently underestimated fine-scale occupancy. On the other hand Kunin (1998) found that the fractal model (i.e. the power-law model)

tended to overestimate the area occupied at fine scale in the same data set. Despite the weaknesses and the strengths of the models' performance in any particular scale, this study has showed that it is possible to deduce and extrapolate information on occupancy at fine scales from coarse-scale data.

A considerable amount of coarse-scale information about species distribution is available in many countries (e.g., Jalas and Suominen, 1988, Mitchell-Jones et al. 1999). Nevertheless, much of this information is not utilised so far for categorising extinction risk following the criteria of the IUCN (World Conservation Union, 2001). An example is the information of plant species distribution in Britain which is summarised as presences and absences in grids of 100 km² squared cells (e.g., Wigginton, 1999). The IUCN considers species as critically endangered if they occupy < 10 km²; under this criterion none of the plant species in Britain would qualify as critically endangered . By generating information on occupancies at fine scales from coarser scale we could overcome this problem. A coarse-scale map, however, contains less information than a fine-scale map of the same area. Nonetheless, occupancy-area curves can be used to estimate some attributes of fine-scale distributions from coarse-scale data (Kunin, 1998).

One of the oldest and best studied scaling phenomenons in ecology is the species area relationship (SAR); in which larger sample units typically contain a greater number of species. Occupancy-area relationships are essentially the single-species equivalent. Larger sample units have a higher probability of containing a given species. The occupancy-area curves can be interpreted as the probabilities of a given area that will contain the species in question. The sum of such probabilities across species determines the expected number of species to be found in a given area, and thus provides the SAR (Tjørve et al. 2008). Tjørve (2009) reviewed the many different ways in which sample units could be arranged to generate species-area relationships (SARs) and the same situations apply to individual species' occupancy scaling relationships. Some of the designs may result in the species remaining undetected in small-size sample units, hence the non-linear methods such as the ones applied in this Chapter are much more flexible in the range of designs they can handle. Historically, researchers have preferred to use data transformations to create linear models because they are much easier to fit. Modern computing power and the wide availability of optimisation software (e.g., Guilhaumon et al.) have removed this

barrier to adoption. Though it must be recognised that non-linear fitting methods sometimes suffer from problems of non-convergence. In most cases these problems can be overcome by supplying sensible and alternative starting values for the parameters under estimation.

Although most studies emphasise the ecological processes that would regulate species abundance, any factors that affect aggregation in space would be equally important in explaining occupancy patterns (Holt et al. 2002). This Chapter illustrated how the parameters of the occupancy-area models capture an important aspect of spatial pattern, e.g. aggregation. The novelty of the work presented here when using nonlinear occupancy-area relationships applied in nested quadrats, could have the potential of predicting occupancy at scales smaller and larger than the original survey, thus saving time and money and facilitating more meaningful comparisons between data collected from disparate survey methods. Such information is needed for accurately categorising species status and for planning conservation priorities, where different species may have been surveyed with different methods. Although this study surveyed plant species, it is possible to apply the same nonlinear method to other taxonomic groups when presence/absence data is available.

2.5   REFERENCES

— Conlisk, E., Conlisk, J., Enquist, B., Thompson, J. & Harte, J. (2009) Improved abundance prediction from presence–absence data. Global Ecology and Biogeography, 18, 1-10.

— Coulter, J. D. (1969) The climate of Molesworth. Meteorological Service Miscellaneous Publication, New Zealand.

— Finlayson, J. C. (1999) Species abundances across spatial scales. Science, 283, 1979.

— Gaston, K. J., Blackburn, T. M., Greenwood, J. J. D., Gregory, R. D., Quinn, R. M. & Lawton, J. H. (2000) Abundance-occupancy relationships. Journal of Applied Ecology, 37, 39-59.

— Guilhaumon, F., Mouillot, D. & Gimenez, O. mmSAR: an R package for multimodel species–area relationship inference. Ecography, 33, 420-424.

— Hamilton, M. A., Murray, B. R., Cadotte, M. W., Hose, G. C., Baker, A. C., Harris, C. J. & Licari, D. (2005) Life-history correlates of plant invasiveness at regional and continental scales. Ecology Letters, 8, 1066-1074.

— Hanski, I. & Gyllenberg, M. (1997) Uniting two general patterns in the distribution of species. Science, 275, 397-400.

— Hartley, S. (1998) A positive relationship between local abundance and regional occupancy is almost inevitable (but not all positive relationships are the same). Journal of Animal Ecology, 67, 992-994.

— Hartley, S. & Kunin, W. E. (2003) Scale dependency of rarity, extinction risk, and conservation priority. Conservation Biology, 17, 1559-1570.

— He, F. & Condit, R. (2007) The distribution of species: occupancy, scale, and rarity. Scaling Biodiversity (eds D. Storch, P. A. Marquet & J. H. Brown). Cambridge University Press, Cambridge.

— He, F. & Gaston, K. J. (2000b) Occupancy-abundance relationships and sampling scales. Ecography, 503-511.

— He, F., Gaston, K. J. & Wu, J. (2002) On species occupancy-abundance models. Ecoscience, 9, 119-126.

— He, F. L. & Gaston, K. J. (2000a) Estimating species abundance from occurrence. American Naturalist, 156, 553-559.

— Holt, A. R., Gaston, K. J. & He, F. L. (2002) Occupancy-abundance relationships and spatial distribution: A review. Basic and Applied Ecology, 3, 1-13.

— Hui, C. & McGeoch, M. A. (2007) Capturing the "droopy-tail" in the occupancy-abundance relationship. Ecoscience, 14, 103-108.

— Jalas, J. & Suominen, J. (1988) Atlas Florae Europaeae: distribution of vascular plants in Europe. Cambridge University Press.

— Kunin, W. E. (1998) Extrapolating species abundance across spatial scales. Science, 281, 1513-1515.

— Kunin, W. E., Hartley, S. & Lennon, J. J. (2000) Scaling down: On the challenge of estimating abundance from occurrence patterns. American Naturalist, 156, 560-566.

— Mitchell-Jones, A. J., Mitchell, J., Amori, G., Bogdanowicz, W., Spitzenberger, F., Krystufek, B., Vohralik, V., Thissen, J., Reijnders, P. & Ziman, J. (1999) The atlas of European mammals. T & AD Poyser London.

— Moore, L. B. (1976) The changing vegetation of Molesworth Station, New Zealand, 1944 to 1971. New Zealand Department of Scientific and Industrial Research, Wellington.

— Motulsky, H. & Christopoulos, A. (2004) Fitting models to biological data using linear and nonlinear regression: a practical guide to curve fitting. Oxford University Press, USA.

— Nachman, G. (1981) A Mathematical-Model of the Functional-Relationship between Density and Spatial-Distribution of a Population. Journal of Animal Ecology, 50, 453-460.

— Pinheiro, J., Bates, D., DebRoy, S. & Sarkar, D. (2010) nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-97.

— R Development Core Team (2011) R: A Language and Environment for Statistical Computing. Vienna, Austria.

— Rastetter, E. B., Aber, J. D., Peters, D. P. C., Ojima, D. S. & Burke, I. C. (2003) Using mechanistic models to scale ecological processes across space and time. Bioscience, 53, 68-76.

— Rastetter, E. B., King, A. W., Cosby, B. J., Hornberger, G. M., O'Neill, R. V. & Hobbie, J. E. (1992) Aggregating fine-scale ecological knowledge to model coarser-scale attributes of ecosystems. Ecological Applications, 2, 55-70.

— Raxworthy, C. J., Martinez-Meyer, E., Horning, N., Nussbaum, R. A., Schneider, G. E., Ortega-Huerta, M. A. & Peterson, A. T. (2003) Predicting distributions of known and unknown reptile species in Madagascar. Nature, 426, 837-841.

— Ritz, C. & Streibig, J. C. (2008) Nonlinear regression with R. Springer Verlag, New York, USA.

— Tjørve, E. (2009) Shapes and functions of species–area curves (II): a review of new models and parameterizations. Journal of Biogeography, 36, 1435-1445.

— Tjørve, E., Kunin, W. E., Polce, C. & Calf Tjørve, K. M. (2008) Species–area relationship: separating the effects of species abundance and spatial distribution. Journal of Ecology, 96, 1141-1151.

— Wigginton, M. J. (1999) Vascular plants. Joint Nature Conservancy Council, Peterborough, United Kingdom.

— World Conservation Union (2001) IUCN Red list Categories and Criteria (version 3.1). World Conservation Union, Gland, Switzerland and Cambridge, United Kindom.

ARE PLANT ABUNDANCE AND AGGREGATION PATTERNS CORRELATED WITH ECOLOGICAL PLANT TRAITS? A MULTI-SCALE ANALYSIS OF A NEW ZEALAND GRASSLAND COMMUNITY

3.1 INTRODUCTION

The distribution patterns of species have been of interest since Linnaeus wrote *Flora Lapponica* (1737) where he described the geographical distribution of some plant species. His ideas have had a profound influence and have inspired many biogeographers and ecologists to seek answers to the question "Why are some species widely distributed while some are narrowly dispersed?" The distribution patterns in natural systems that result from processes like reproduction, growth and dispersal occur in space and time. One way to study the distribution patterns of species is by quantifying the degree of spatial aggregation (Réjou Méchain et al., 2011). Understanding aggregation is crucial to elucidate how species are using resources and how it is used as a resource (Condit et al., 2000). Also, aggregation may help infer which processes are important in shaping a particular pattern (Tilman and Kareiva, 1997).

To understand these functional processes we need to identify the relevant spatial scale at which they occur (Fortin & Dale 2005). Nevertheless, explaining the population patterns of even a single species is time consuming because we require an intimate knowledge of the species' natural history, its abiotic limitations and its full range of biotic interactions (Pocock et al. 2006). The potential effect of these processes on aggregation has also been suggested to be scale-dependent (Wiens, 1989), as the spatial scale investigated changes, processes and thus patterns, are also expected to change (Réjou Méchain et al., 2011). For example, competition and reproduction are thought to operate at local scales (Woodward, 1987, Van der Veken et al., 2007) while at intermediate scales, dispersal could affect distributions patterns (Wiens, 1989). However, at broader scales, physical factors (e.g. barriers) may dominate or dissipate these biological effects (Wiens, 1989, Guisan and Thuiller, 2005).

An alternative approach is to statistically examine the distributions of a large number of species, looking for traits that are correlated with

patterns of abundance or distribution (Gaston et al. 2000; Cadotte 2006). There is growing literature that demonstrates how some plant traits are correlated to important processes like colonisation, reproduction and competition (Cornelissen et al. 2003; Lloyd et al. 2003; Diaz et al. 2004; Westoby & Wright 2006; Wright et al. 2007). For example, species with small geographic range sizes tended to be self-compatible or asexual, had lower reproductive investment, dispersal ability and genetic variation, narrower resource usage and smaller body size. However, most of the studies that relate aggregation patterns with processes focus on one single scale yielding contradictory results. Here, I carried out a multi-scale analysis where aggregation patterns were measured at different scales, from local-scale to regional-scale and for all scales. Therefore my first aim was to correlate aggregation patterns with plant traits that might help us understand distribution patterns at different scales. I expect that life-history traits that correlate with competition and reproduction (e.g. specific leaf area) are more important at local-scales, while traits that correlate to dispersal (e.g. seed mass, plant height) would be more important at regional scales. The following hypotheses were tested and the scale at which these traits are expected to be important is indicated.

The first aim of this Chapter was to identify what processes inherent to the species measured as plant life-history traits, were important at different scales. Processes inherent to species (e.g. dispersal, competition) and processes independent of species (e.g. disturbance, environmental conditions), can generate different types of distribution patterns (Fortin & Dale 2005). Several studies have shown that patterns at coarse scale are likely to be controlled by climatic regulators (Guisan & Thuiller 2005) and patterns at fine scale are controlled by interspecific interactions (Van der Veken et al. 2007). However, most of the studies that relate distribution patterns with processes focus on one single scale, yielding contradictory results. Here, I carried out a multi-scale analysis, where the distribution patterns were measured at different scales, form local-scale to regional-scale and for all scales.

Local-scale

Specific leaf area (SLA) *and leaf area*. SLA tends to correlate with metabolic activity and plant growth. Low values of SLA (higher leaf thickness and/or density) are more tolerant to water and nutrient deficiency due to their thicker cuticule layers and slower leaf turnover rates (Wright et al., 2002). Leaves with high SLA utilise low amounts of diffuse radiation more efficiently (Sefton et al., 2002), but they are

less stress tolerant and sensitive of being dried out (Schulze et al., 2005). Therefore, I expect that leaves with high SLA tend to have more aggregated patterns which could be present only at favourable sites, while leaves with low SLA would be able to persist in more scattered distributions in unfavourable places with lower growth rate. On the other hand, for the leaf area plant trait, large leaves have been suggested to experience greater degrees of herbivory (Moles and Westoby, 2000), which may affect their aggregation patterns. Plants with larger leaves are expected to have more scattered distributions because large aggregations of plants disproportionately attract and support specific herbivores and pathogens of the plant (Janzen, 1970).

Regional-scale

Plant height. It is related to dispersal and taller species have been found to increase dispersal capacity into previously unoccupied habitats (Mathews and Bonser, 2005, Thomson et al., 2011). Thus, I expect that taller plants tend to have more scattered distributions.

Woodiness. It is an indirect measure of longevity. Long-lived species may have a more scattered distribution than short-lived species. This is due to the persistence of isolated individuals outside the range within which recruitment is currently possible (Pocock et al., 2006). Woody plants are expected to have relatively large and more scattered distributions.

Seed and dispersule mass. The size and number of seeds produced are subjected to a trade-off in resource allocation in plants and have complex interactions with the mode of dispersal (Thompson et al., 2001). But overall, it is expected that small seeded species are better dispersers and have relatively larger, more scattered distributions.

Dispersal mode. Plants that have mechanisms to disperse further away might develop distributions that are less aggregated (Pocock et al., 2006). Thus, it is expected that plants with no specialised dispersal mechanisms have more aggregated distributions.

A problem encountered when attempting to identify the processes generating such multi-species patterns, is that the species' characteristics of are not independent. Phylogenetically related species share characteristics through their common evolutionary ancestry (Felsenstein, 1985). Two closely related species will therefore tend to be more similar to each other in a range of characters, than two species chosen at random from the entire assemblage being compared (Blackburn, 2004). Therefore I compared results from both cross-species and phy-

logenetic independent contrast analyses. Analyses were conducted of each plant trait individually. Also multivariate analyses were performed in order to assess the importance of interactions between the variables in affecting plant aggregation patterns.

Many indices have been used to describe spatial distribution (see Dale 2000; Dale et al. 2002). In the present Chapter, two aggregation indices were used. First, I used the *k* parameter from the NBD1 which is one of the most common indices to quantify aggregation (He et al., 2002, He and Gaston, 2000) (it is calculated from the mean and variance of individuals). Nevertheless, the traditional NBD1 has the disadvantage of requiring count data which is hard to obtain and not always applicable to plants that do not have unitary growth form. So, I quantified the same parameter *k* but calculated from the occupancy-area relationship from Chapter 2 (NBD2). The advantage of using NBD2 is because it is a measure of aggregation related to the NBD1 that can be derived from presence/absence data as input (as long as there is a multi-scale, nested component to the sampling design). Furthermore it allows the use of data that includes zeroes (calculated from occupancy-area relationship, see Chapter 2). The second aim was to test the correlation between the aggregation parameter *k* from the occupancy-area model (NBD2) with the aggregation parameter *k* from the negative binomial 1(NBD1; e.g., count data). The occupancy-area method (NBD2) of describing aggregation is more naturally suited to species which do not have discrete individuals, but whose distribution and abundance is measured by presence-absence in quadrats and percent cover. The traditional method of calculating *k* from mean and variance assumes individuals can be counted within sample units. If the occupancy-area method is correlated, the approach could potentially be applied in a range of species distribution information obtained from presence/absence or from the typical atlas dot maps.

The third aim was to identify the types of distribution and rarity displayed by the species surveyed in two scales; local-scale (20×20 m quadrats) and regional-scale (100×60 m quadrats). Different types of distribution patterns in plants have been suggested, including several forms of rarity (Rabinowitz 1981; Rabinowitz et al. 1986). However, there is not a straightforward way to identify the different types of distribution patterns, including rarity. Many researchers have adopted their own definition including the breath of geographic range, local frequency and endemism (e.g., Gaston 1994; Quinn et al. 1996), but there is not a universal definition of the different types of rarity and

NBD *negative binomial model*
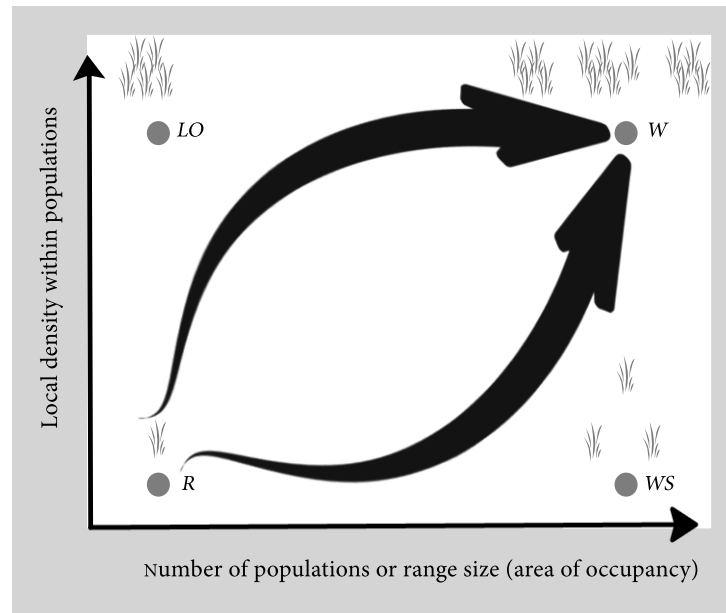
Figure 13: Different forms of rarity (*R, LO, WS*) and different trajectories away from a situation of widespread, high-density (*W*). Fluctuating populations might display consistently different trajectories of decline compared to expansion for the same number of individuals, *LO* species are much more aggregated than *WS* species. The trajectories depicted are moving away from a situation of rarity (as an invasive species might start).

distribution patterns since they are scale-dependant. Here, I identify different types of distribution with two axes of measurement and four types of distribution: species with small populations and in low densities (*R*); species locally abundant but with narrow ranges (*LO*); species that are widespread but with small populations (*WS*); and species that are both regionally widespread and locally abundant (*W*; Figure 13). Species *R, LO* and *WS* constitute different types of rarity. Likewise, there are different spatial scales at which a population can expand or decline; the two extreme trajectories being a uniform decline in density elsewhere, versus loss of discrete populations or a contraction of range margins. In the same manner, there are different ways in which a new invader can expand from *R* towards *W* (arrows in Figure 13). The second aim was to identify the types of distribution and rarity displayed by the species surveyed in two scales; local-scale (20 × 20 m quadrats) and regional-scale (100 × 60 m quadrats).

3.2 METHODS

3.2.1 *Study area and field survey*

The field survey was performed during the summers of 2008, 2009 and 2010 in Molesworth Station which is situated north-west of the Inland Kaikoura Mountains in South Marlborough, New Zealand (Moore 1976, Figure 14).

The average annual rainfall of Molesworth station ranges from less than 760 mm in the East to more than 2,540 mm in the West (Coulter 1969). Long dry periods may occur at all times of the year, but especially in summer and autumn (Coulter 1969). The average temperature varies from 14 °C in summer to 10 °C in winter. Air frosts occur in all months, ground frosts are experienced on 218 nights per annum on average (Coulter 1969). The vegetation reflects this range of rainfall. In dry places it is common to find gravelfield and scree plant communities. In places where the rainfall is heavier, exotic grassland species and short native tussocks are common.

3.2.2 *Distribution data and plot design*

Twenty-seven 100 × 60 m plots were established. Each plot was gridded into fifteen 20 × 20 m quadrats. The central 20 × 20 m quadrat was permanently marked with aluminium pegs, and these central quadrats were subdivided in sixteen 5 × 5 m subquadrats (Figure 15 and 16). Plots were randomly placed with restrictions (within a buffer area of 3 km on each side of the gravel road due to the inaccessibility of some remote areas) along three catchments within the Molesworth Station boundaries (Figure 14).

In order to homogenise the search effort of species within every quadrat and reduce bias, search times were set for each quadrat size. For the 20 × 20 m quadrats the search time was of 30 minutes, and for the each 5 × 5 m subquadrats the search time was of 10 minutes.

I focused on the most common plant species of the Molesworth area. The species sampled represent a range of life-history traits covering a wide range of growth forms. Some common species that occur in the area were not included due to the difficulty of knowing its taxonomic affinities. The total data set comprised 46 species belonging to 19 families (Appendix B). Thirty-two of the species for inclusion on the survey were determined a priori using the species list of Husheer
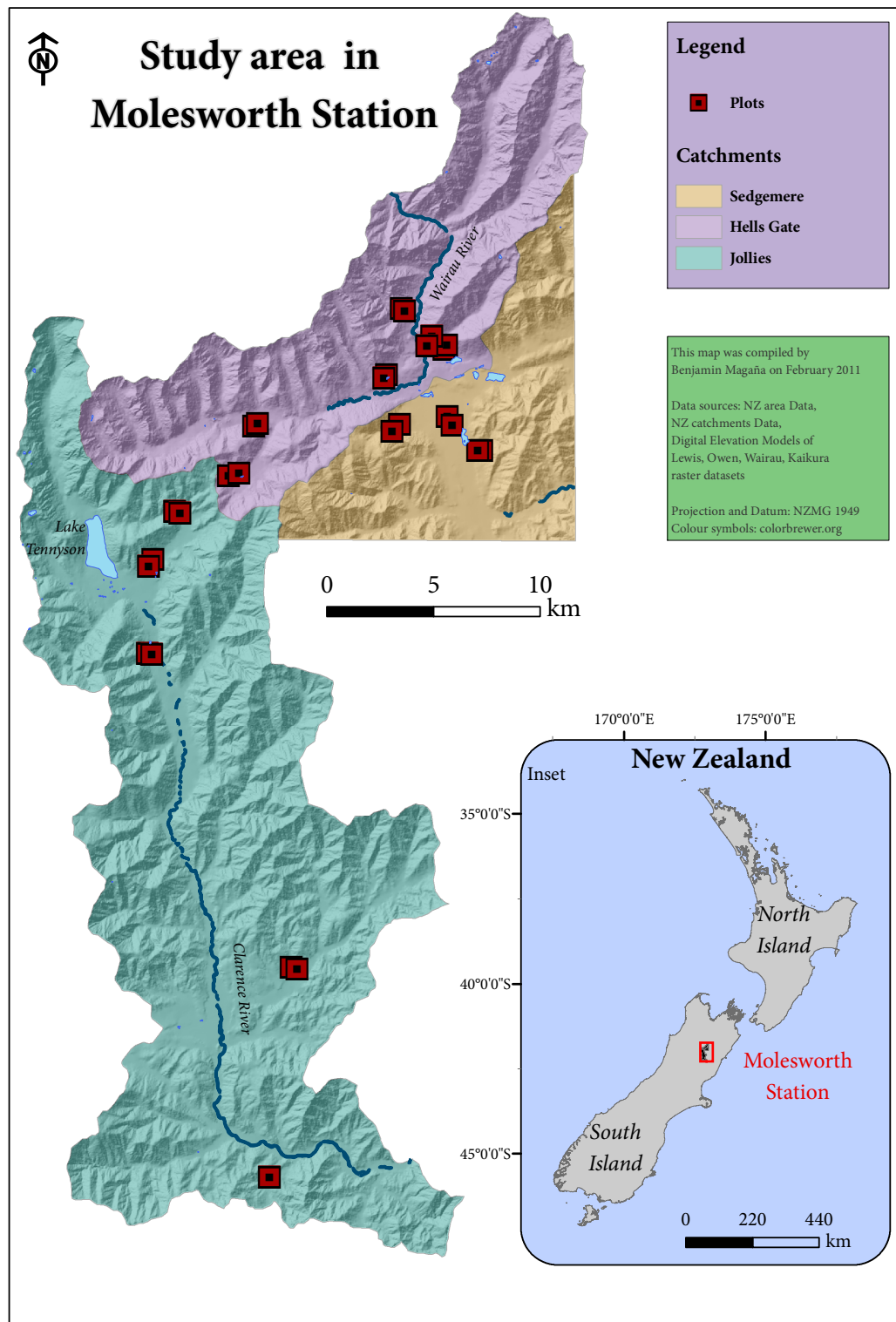
Figure 14: Study area and plot location within the three catchments of the survey.
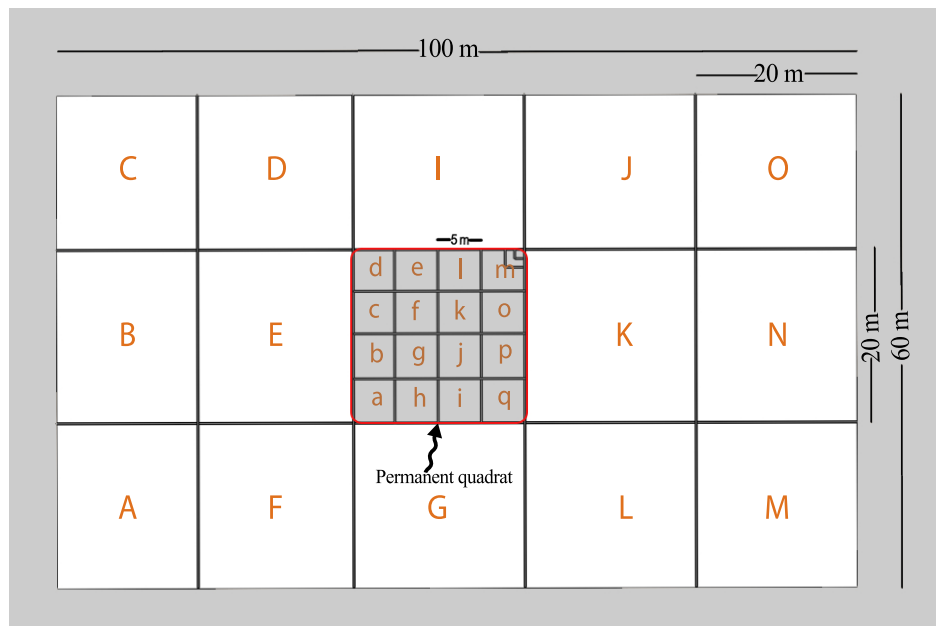
Figure 15: Plot design and arrangement of quadrats and subquadrats. The central quadrat was permanently marked (highlighted).

(unpub data). Fourteen additional species were added to the list upon completion of the first four plots.

3.2.3 *Trait selection and measurement*

Data on seven morphological and life-history traits were compiled from various sources: a) specialised literature (e.g., New Zealand floras), b) the database of ecological traits from Landcare Research (ecotraits) and c) data gathered from own measurements collected in the field (see Appendix S1 and Table 13). Measurement of traits from field-collected specimens followed the methodological guidelines of Cornelissen et al. (2003). Morphological measurements were made on 5-20 leaves and seeds harvested from five individuals from each species (Appendix S2; Table 14). Leaves and seeds were collected in the field from February 2007 to March 2010 following the protocol of Cornelissen et al. (2003). Leaf area, SLA, seed mass, dispersule mass and height were measured in all species surveyed in the field.

SLA *specific leaf area*

Leaves were individually placed on an hp scanjet computer scanner and a high resolution, black and white image was generated. Images were imported into IMAGE-J (Abramoff et al. 2004) on the same day specimens were collected. Their length, width and area were measured. Leaves were then dried to a constant mass in a drying oven at 30°C, and leaf mass was measured to the nearest 0.01 mg with an electronic

(a) Coarse-scale:100 × 60 m plot gridded in fifteen 20 × 20 m quadrats
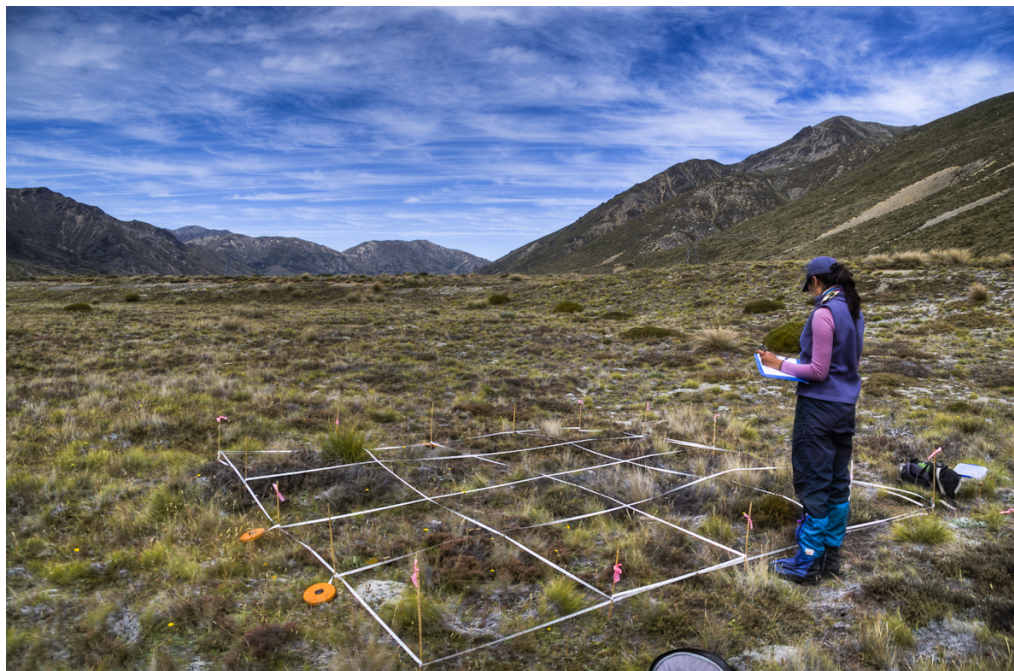
(b) Fine-scale: sixteen 1 × 1 m subquadrats

Figure 16: Example of different plot scales used.

balance. From these measures, two leaf traits were calculated: total leaf area (mm²) and specific leaf area (mm² mg⁻¹). For measurements seed and dispersule mass, seeds and leaves were dried in the laboratory oven at a constant temperature of 30°C for 48 h, and collectively weighed to the nearest 1 mg. Height of each plant was measured using an extendible ruler.

3.2.4 *Data analysis – types of rarity*

For each of the 46 species, area of occupancy was calculated as the proportion of quadrats in which a species was found. AOO was calculated at two different scales: the local-scale using all four hundred and five 20×20 m quadrats (400 m²) and regional scale using the twenty-seven 100×60 m plots (6000 m²) quadrats. Out of the 46 species, abundance was measured for twenty-four of them because of their unitary growth habit. Abundance was defined as the mean count of individuals per plot (regional-scale) and per quadrat (local-scale).

AOO *area of occupancy*

In order to determine the different rarity categories (Figure 13, Table 5 and 6), I computed the median of occupancy and abundance. *LO* species were considered as those having an occupancy lower than the median of all species and higher abundance than the median; *WS* species were those having an occupancy higher than the median and lower abundance than the median; in *R* species both occupancy and abundance were lower than the median; in *W* species both occupancy and abundance were higher than the median.

Although different types of rarity can be recognised and categorised as above, my major interest was to find out what distinguishes species with a localised form of rarity (*LO*) from those that are widespread but scattered (*WS*), for a given number of individuals. For this purpose I used a continuous measure of aggregation which reflects this diagonal axis of variation on abundance-occupancy plots (Hartley 1998).

3.2.5 *Data analysis – aggregation*

Aggregation was quantified for three different spatial scales; at a local-scale, at a regional-scale and across all scales. Aggregation for local and regional-scale was measured as the parameter *k* from the NBD, since is one of the probabilistic models most frequently used to describe spatial distribution in biological populations (He & Gaston

2000). When individuals of a species are distributed according to a
NBD, the probability of the presence of a species in sampling area a is

$$p = 1 - \left(1 + \left(\frac{\mu}{k}\right)\right)^{-k}$$
(3.1)

where k is a "clumping" parameter and μ is the mean abundance
across sampling unit a. From that equation we can derive k such that:

$$k = \frac{\mu^2}{\sigma^2} - \mu$$
(3.2)

where μ is the mean abundance across the entire area including non-
occupied sites and σ^2 is the variance.

For the purposes of this study, the inverse of k ($1/k$) was used.
Large $1/k$ values indicate high degree of spatial clumping in the
distribution pattern at a particular scale of analysis; values close
to zero are equivalent to random distribution and negative values
indicate a distribution that is more even than random. At the local
scale $1/k$ was calculated separately for each of the 27 sites, and then
the mean of $1/k$ was computed. For regional scale $1/k$ was calculated
directly using each 60×100 m plots as a single sample unit.

Aggregation across all scales was calculated as the parameter $1/k$ of
the negative binomial 2 (NBD2) model of the occupancy-area relation-
ship, fitted using the non-linear method from Chapter 2. This method
allows the use of occupancy (proportional) data instead of count data:

$$occ = A \cdot 1 - (1 + \frac{N \cdot a}{A \cdot k})^{-k}$$
(3.3)

where occ is occupancy (the proportion of sample units of the occu-
pied area), a is the area of a single sample unit (quadrat size), k is
a parameter generally regarded to reflect degree of aggregation or
positive spatial autocorrelation among occupied quadrats. Because
occupancy is measured as a proportion of the total available (hence in
the range 0-1) the total area of the study area is 1 and is removed from
the original formulation of He and Gaston 2000a, b. This causes the pa-
rameter N, in the negative binomial model 2 (NBD2) to be proportional
to the estimated number of individuals across the total study area
rather than a direct estimate. Proportional occupancy was measured
at four scales (resolutions): 1×1 m subquadrats ($n = 432$), 5×5 m

subquadrats ($n = 432$), 20×20 m quadrats ($n = 405$), and 100×60 m plots ($n = 27$).

For both methods the estimated parameter k (or $1/k$) was used as an index of aggregation, but it does not imply that the species necessarily have distributions that match a theoretical NBD.

Since the k parameter obtained from the NBD2 using occupancy-area nonlinear regressions is a novel method for measuring aggregation proposed in Chapter 2, the results were compared using the more traditional method for measuring aggregation (Local and regional-scale k from NBD1 count data, Eq. 3.1), hence the use of $1/k$ in all analyses.

3.2.6 *Relationships between plant traits and aggregation – cross-species and PICs analyses*

Association between plant traits and distribution patterns were analysed using cross-species correlations and phylogenetically independent contrasts. Cross-species analyses have traditionally been used in comparative studies to examine inter-specific correlations among biological and ecological traits. However, because these analyses treat species as independent data points they have been criticised (Felsenstein 1985). One method to overcome this problem is to use techniques such as phylogenetically independent contrasts, which control for the non-independence of taxa within phylogenies (Felsenstein 1985; Harvey & Pagel 1991).

Cross-species relationships between life-history traits and degree of aggregation (measured by the negative binomial NBD1 and NBD2 aggregation index $1/k$) were analysed using simple regression analysis for univariate models. Analysis of variance (ANOVA) was performed to compare different groups. Data were logarithmic (ln) transformed to conform to assumptions when necessary (Table 7).

Since some traits are correlated, the interaction between them could affect distribution patterns (Pocock et al. 2006). Multiple regression analyses were performed for testing this interaction in multivariate models. The full model (all predictor variables) was considered first, and then the Minimum Adequate Model (MAM) was fitted following the principle of parsimony (Sober 1991). The most parsimonious model was assessed as the one with the lowest AIC (Burnham & Anderson 2002). All statistical analyses were conducted in R v.2.10 (R

`Development Core Team 2011`) using the `MASS` package (Venables &
Ripley 2002).

A phylogeny was obtained using PHYLOMATIC software (Webb &
Donoghue 2005). PHYLOMATIC includes data from the Angiosperm
Phylogeny Group (APG) III Group 2009; Bremer et al. 2009). Unre-
solved nodes and subtrees in Phylomatic were dated according to
complementarily references (Wagstaff et al. 2002; Magallón & Sander-
son 2005; Hedges et al. 2006; Bremer 2009; Forest & Chase 2009b, a;
Magallón 2009; see Table 15 in Appendix S3 for estimated age and
nodes) and the *bladj* algorithm (Webb & Donoghue 2005). The phy-
logeny is given in Figure 17. Phylogenetically independent contrasts
were analysed using `adephylo` (Jombart et al. 2010), `ape` (Paradis et
al. 2004), `phylobase` (Hackathon et al. 2010) and `picante` (Kembel
et al. 2010) packages in `R` v.2.10 (`R Development Core Team 2011`). A
detailed explanation of the usage of these packages can be found in
Paradis (2006). Regressions were passed through the origin as recom-
mended by Legendre & Desdevises (2009). There were not enough
nodes to performed PIC analyses on binary data. Instructions on how
to calculate Phylogenetic independent contrast (PIC) in `R` are given in
Appendix C. In order to detect which divergences in the phylogenetic
tree were important and where the highest divergences in plant traits
were, the contribution index was computed in phylocom (Webb et al.
2008).

3.2.7 *Comparison between the aggregation parameter 1/k from the negative binomial 1 and the negative binomial 2*

Model II regression (also known as major axis (MA) regression) was
used to quantify the relationship between the parameter $1/k$ from the
negative binomial 1 (e.g., count data) and $1/k$ from the negative bino-
mial 2 (e.g., occupancy data) for the 24 species in common. This type of
regression is recommended when the two variables are not controlled
by the researcher (Legendre & Legendre 1998). Model I regression
using least squares underestimates the slope of the linear relation-
ship between the variables when they both contain error (Legendre
& Legendre 1998). Three regressions were performed: first between
the $1/k$ local-scale and the $1/k$ regional-scale both from the negative
binomial 1; secondly, between $1/k$ local-scale and $1/k$ all scales from
the negative binomial 2; and finally between $1/k$ regional-scale (NBD1,
Eq. 3.1) and $1/k$ all scales (NBD2, Eq. 3.3).
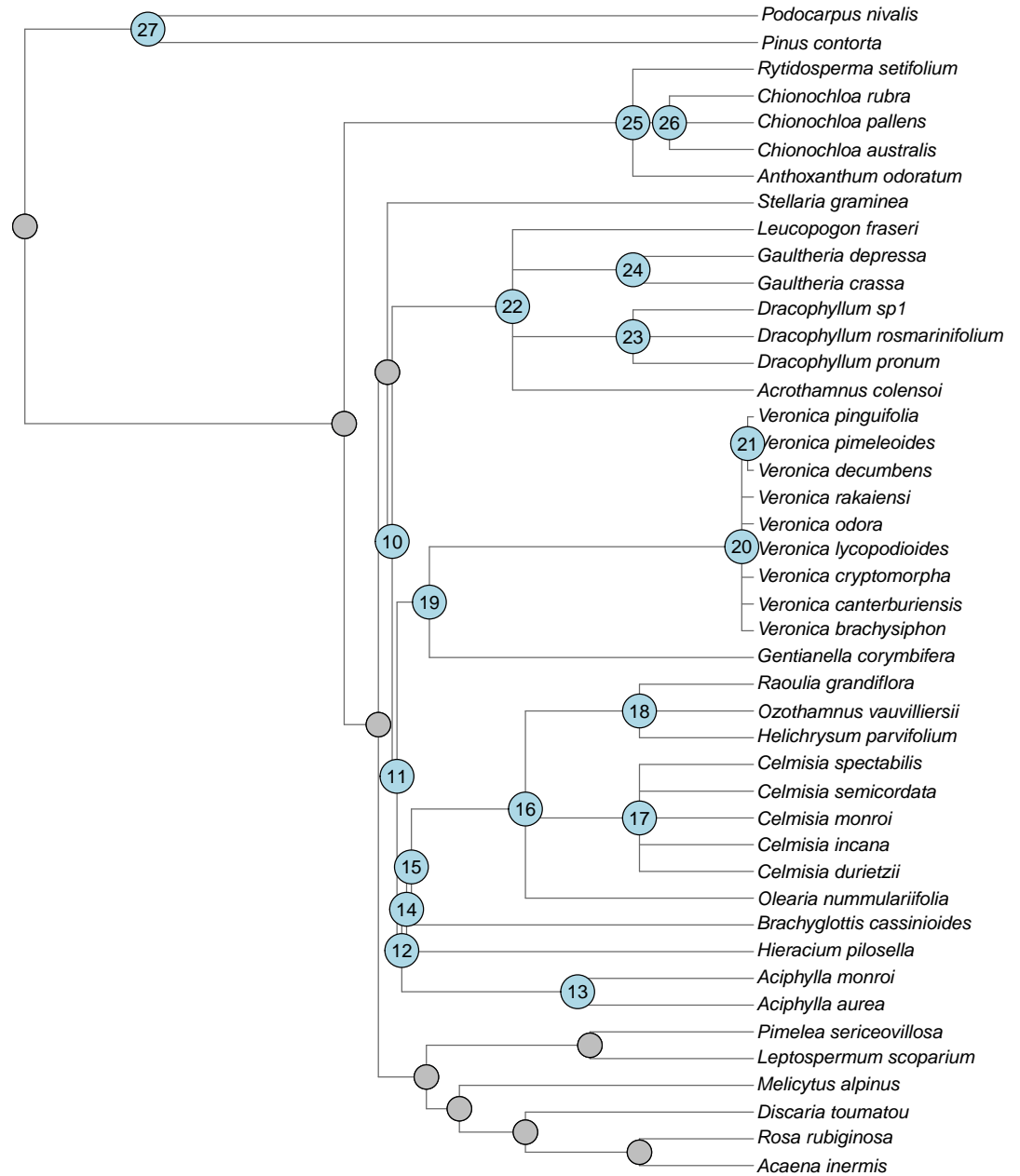
Figure 17: Phylogenetic tree of the species surveyed from Molesworth Station. Dated nodes are numbered. Estimated ages are given on Table 15.

The relationship between the aggregation $1/k$ from negative binomial 1 (local and regional $1/k$, count data, Eq. 3.1) and the aggregation $1/k$ from the negative binomial 2 (NBD2, occupancy data, Eq. 3.3) was tested for the 24 species in common.

3.3 RESULTS

3.3.1 *Types of distribution and rarity*

The type of distribution and rarity is scale-dependant. Species displayed different types of distribution and rarity depending of the spatial scale analysed. Species such as *Veronica odora* and *Celmisia monroi* moved from type *R* at a local-scale to type *WS* at a regional-scale.

Average abundance in 20×20 m quadrats ranged from less than one individual per quadrat, to almost 70 individuals per quadrat from species of the assemblage.

At a local scale (20×20 m quadrats), most species displayed rarity type *R* (relatively small AOO with low densities) or were common and widespread (*W*) with large AOO and high densities. Only four species had localised *(LO)* or widespread but scattered *(WS)* rarity forms (Figure 18 and Table 5).

At a local-scale, spatial structure, measured as $1/k$ from the NBD, varied among species. Results varied from larger $1/k$ values meaning highly aggregated patterns (e.g., *Veronica pimeleoides*, *Aciphylla glaucescens*), to lower values equivalent to dispersed and random distribution patterns (e.g., *Chionochloa pallens*, *Veronica decumbens*, Table 5; isolines of $1/k$ in Figure 18).

Average abundance in 100×60 m plots ranged from less than one individual per quadrat, to almost 1000 individuals per quadrat from species of the assemblage.

At a regional-scale (100×60 m plots) *R* and *W* types of rarity were more common. Species such as *Veronica pimeleoides* and *Helichrysum parvifolium* were very restricted in range but locally abundant (*LO*). In contrast, *Veronica decumbens* and *Celmisia monroi* were more widespread in range but at low densities (*WS*; Table 6, Figure 19).

Similar to local scale, spatial structure ($1/k$) varied from very aggregated (e.g., *Veronica pimeleoides*) to more dispersed patterns (e.g., *Gentianella corymbifera*, isolines of $1/k$ in Figure 19).

R, species with small populations and in low densities; LO, species locally abundant but with narrow ranges; WS, species that are widespread but with small populations; W, species that are both regionally widespread and locally abundant

Table 5: Local scale summary (within 20×20 m quadrats). The parameter $1/k$ of the negative binomial distribution, assumes larger values as the aggregation increases.

SPECIES	SITES OCCUPIED	MEAN ABUN-DANCE	AOO	MEAN VARIANCE	MEAN $1/k$	TYPE OF RARITY
Aciphylla aurea	203	10.12	0.52	101.48	0.49	*W*
Aciphylla glaucescens	9	0.44	0.02	1.09	5.45	*R*
astelia nervosa	12	1.17	0.03	4.69	3.23	*R*
Brachyglottis cassinioides	82	31.43	0.21	1399.01	2.79	*W*
Celmisia monroi	37	0.28	0.10	0.32	0.19	*R*
Chionochloa pallens	176	68.25	0.45	4106.77	0.11	*W*
Chionochloa rubra	108	14.36	0.28	265.99	1.91	*W*
Chionochloa sp1	44	66.24	0.11	827.19	2.04	*LO*
Discaria toumatou	107	16.25	0.28	538.40	1.54	*W*
Gentianella corymbifera	261	44.68	0.67	2374.19	1.04	*W*
Helichrysum parvifolium	41	7.59	0.11	211.77	2.23	*LO*
Leptospermum scoparium	13	3.33	0.03	83.77	2.60	*R*
Olearia nummulariifolia	83	4.78	0.21	4.78	0.54	*W*
Ozothamnus vauvilliersii	149	5.44	0.38	146.34	1.24	*W*
Pinus contorta	35	1.36	0.09	3.38	1.63	*R*
Rosa rubiginosa	21	1.70	0.05	7.08	3.51	*R*
Veronica brachysiphon	80	5.09	0.21	131.00	0.56	*W*
Veronica canterburiensis	69	2.13	0.21	151.02	1.97	*WS*
Veronica cryptomorpha	3	0.27	0.01	0.35	1.21	*R*
Veronica decumbens	42	0.71	0.11	6.84	0.09	*R*
Veronica lycopodioides	28	0.38	0.07	0.89	0.79	*R*
Veronica odora	35	0.95	0.09	6.56	0.28	*R*
Veronica pimeleoides	16	13.09	0.04	552.42	2.85	*LO*
Veronica pinguifolia	17	0.26	0.04	0.32	1.17	*R*
Veronica rakaiensis	14	0.43	0.04	0.60	-0.06	*R*
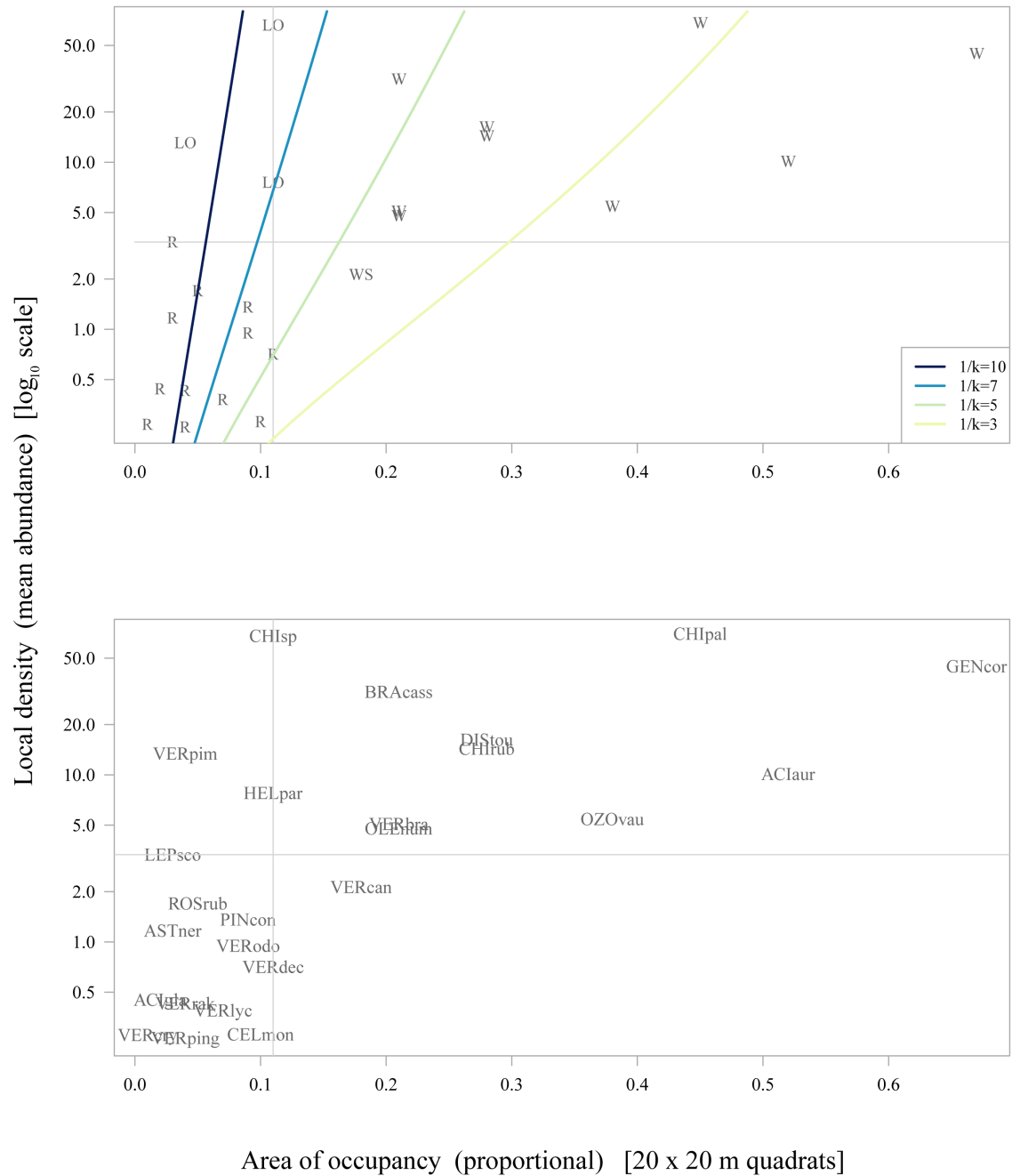
Figure 18: Relationship between the mean abundance and the proportional occupancy at local-scale (20 × 20 m quadrats). Light grey lines indicate the medians of occupancy (vertical) and local density (horizontal) which were used as a criterion to define the different types of rarity for the species surveyed (*LO, R, WS* and *W*, see Figure 13 for a description). Isolines represent different values of the parameter $1/k$ from the negative binomial distribution. The degree of aggregation increases as the values become larger. The top plot shows the actual rarity group and the bottom plot the species that belong to the rarity group. List of species' abbreviations are given in Appendix B.

Table 6: Regional scale summary (between plots 100 × 60 m). The parameter
1/k of the negative binomial distribution, assumes larger values as
the aggregation increases.

SPECIES	TOTAL IN-DIVIDUALS	SITES OCCUPIED	MEAN ABUN-DANCE	AOO	VARIANCE	1/k	TYPE OF RARITY
Aciphylla aurea	3778	24	139.93	0.89	68185.61	3.46	*W*
Aciphylla glaucescens	20	3	0.74	0.11	6.51	10.50	*R*
Astelia nervosa	35	2	1.30	0.07	22.29	12.46	*R*
Brachyglottis cassinioides	4715	10	174.63	0.37	348860.90	11.42	*W*
Celmisia monroi	47	11	1.74	0.41	11.35	3.17	*WS*
Chionochloa pallens	20474	20	758.30	0.74	2618294.00	20.07	*W*
Chionochloa rubra	3015	14	111.67	0.52	209153.70	4.54	*W*
Chionochloa sp1	2981	3	110.41	0.11	244599.90	16.73	*LO*
Discaria toumatou	3656	15	135.41	0.56	241228.70	13.18	*W*
Gentianella corymbifera	14742	22	546.00	0.81	602140.50	2.02	*W*
Helichrysum parvifolium	674	6	24.96	0.22	6753.73	10.82	*LO*
Leptospermum scoparium	150	3	5.56	0.11	735.10	23.62	*R*
Olearia nummulariifolia	931	12	34.48	0.44	19360.64	16.24	*W*
Ozothamnus vauvilliersii	1386	17	51.33	0.63	17909.08	6.76	*W*
Pinus contorta	102	5	3.78	0.19	184.56	12.67	*R*
Rosa rubiginosa	102	4	3.78	0.15	171.87	11.76	*R*
Veronica brachysiphon	1146	15	3.78	0.56	22532.87	12.46	*W*
Veronica canterburiensis	351	11	13.00	0.41	2342.92	13.76	*W*
Veronica cryptomorpha	4	1	0.15	0.04	0.59	20.25	*R*
Veronica decumbens	128	12	4.74	0.44	291.58	12.74	*WS*
Veronica lycopodioides	51	9	1.89	0.33	22.41	5.76	*R*
Veronica odora	156	11	5.78	0.41	639.10	19.01	*WS*
Veronica pimeleoides	589	3	21.81	0.11	12713.54	26.63	*LO*
Veronica pinguifolia	23	6	0.85	0.22	5.28	6.10	*R*
Veronica rakaiensis	26	4	0.96	0.15	17.88	18.23	*R*
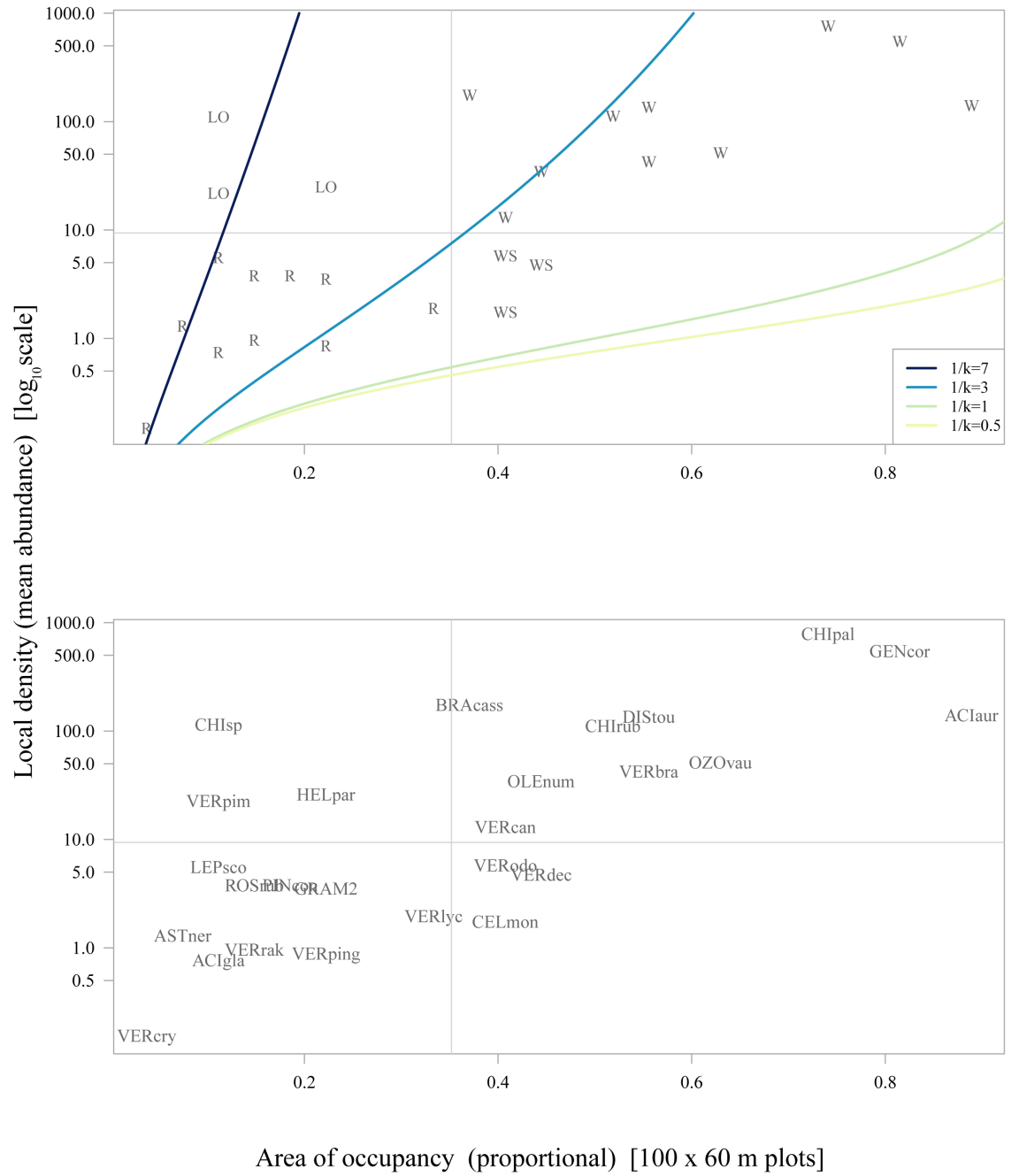
Figure 19: Relationship between the mean abundance and the proportional occupancy at regional scale (100×60 m quadrats). Light grey lines indicate the medians of occupancy (vertical) and local density (horizontal) which were used as a criterion to define the different types of rarity for the species surveyed (*LO, R, WS* and *W*, see Figure 13 for a description). Isolines represent different values of the parameter $1/k$ from the negative binomial distribution. the degree of aggregation increases as the values become larger. The top plot shows the actual rarity group and the bottom plot the species that belong to the rarity group. List of species' abbreviations are given in Appendix B.

3.3.2   *Relationships between plant traits and aggregation – cross-species analyses*

The relationship between aggregation and plant life-history trait varied across the scales used for measuring aggregation.

Vegetative traits

At a local-scale, SLA was the only plant trait that was significantly positively correlated with local aggregation $1/k$ ($R^2 = 0.25$; $n = 24$ species; $P = 0.01$; Figure 20b, Table 7). This means that species comprising plants with thin leaves and low density tissues tend to have more aggregated patterns.

At a regional-scale, the average plant height of different species exhibited a significant positive association with aggregation ($R^2 = 0.20$; $n = 24$ species; $P = 0.03$; Figure 21e, Table 7). Species comprising taller plants had more aggregated patterns. In the same way, there is a tendency for species comprising woody plants to have more regional aggregated patters, although this was not significant ($P = 0.07$; Figure 21f).

Reproductive traits

When aggregation was measured across all four scales using the occupancy-area model ($1/k$ NBD2), dispersal mode was associated with aggregation ($n = 46$ species; $P = 0.04$). Species that lack any specialised system dispersal were more aggregated (Figure 22g).

Figure 20: Scatter and boxplots of variables used as predictors of distribution patterns (aggregation $1/k$) at local scale for cross-species. In the scatter plots each point represents one species. In the boxplots the dark line is the median, the box encloses the interquartile range and the whiskers show the full range, with outliers shown as circles.
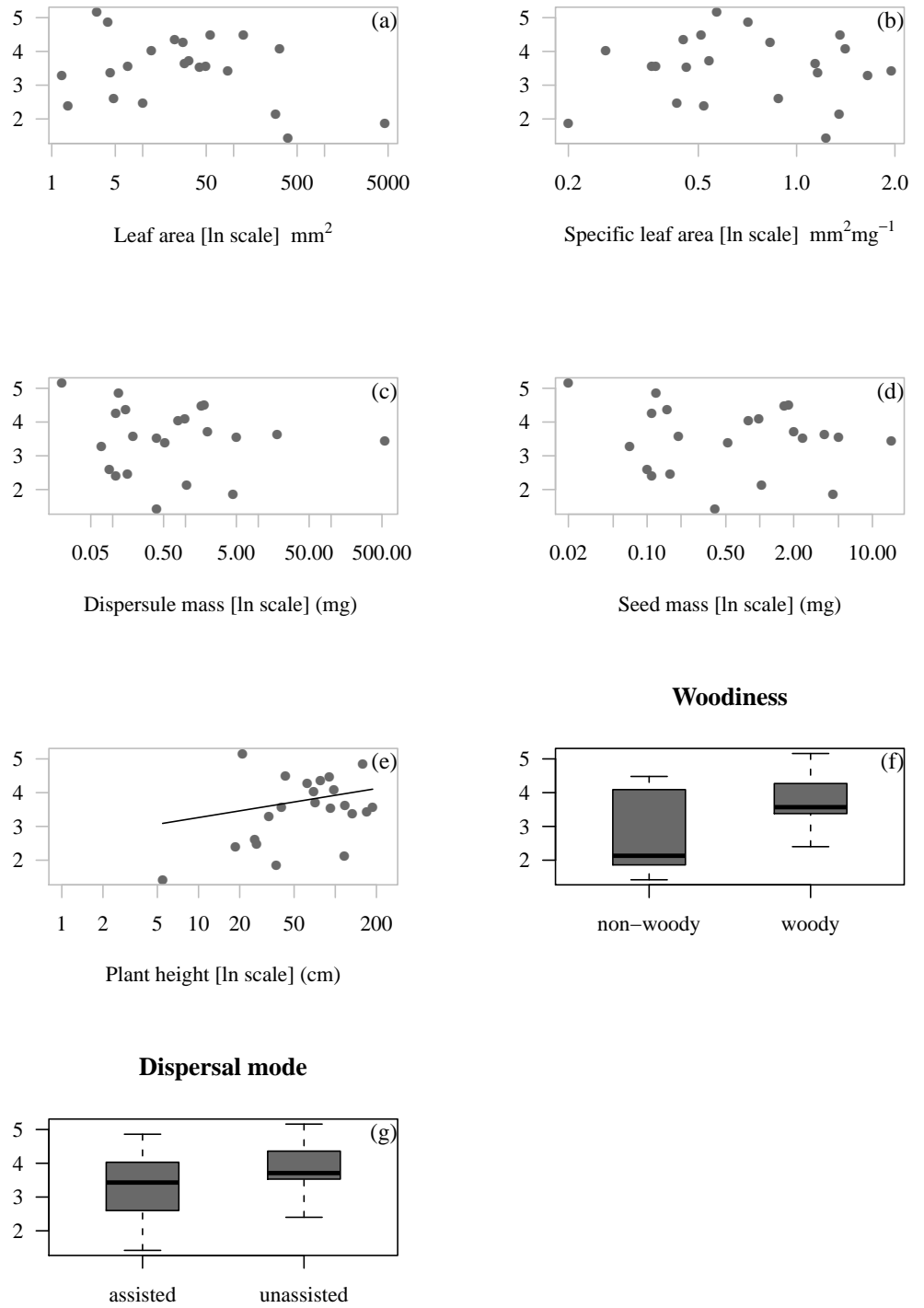
Figure 21: Scatter and boxplots of variables used as predictors of distribution patterns (aggregation $1/k$) at regional scale for cross-species. In the scatter plots each point represents one species. In the boxplots the dark line is the median, the box encloses the interquartile range and the whiskers show the full range, with outliers shown as circle.
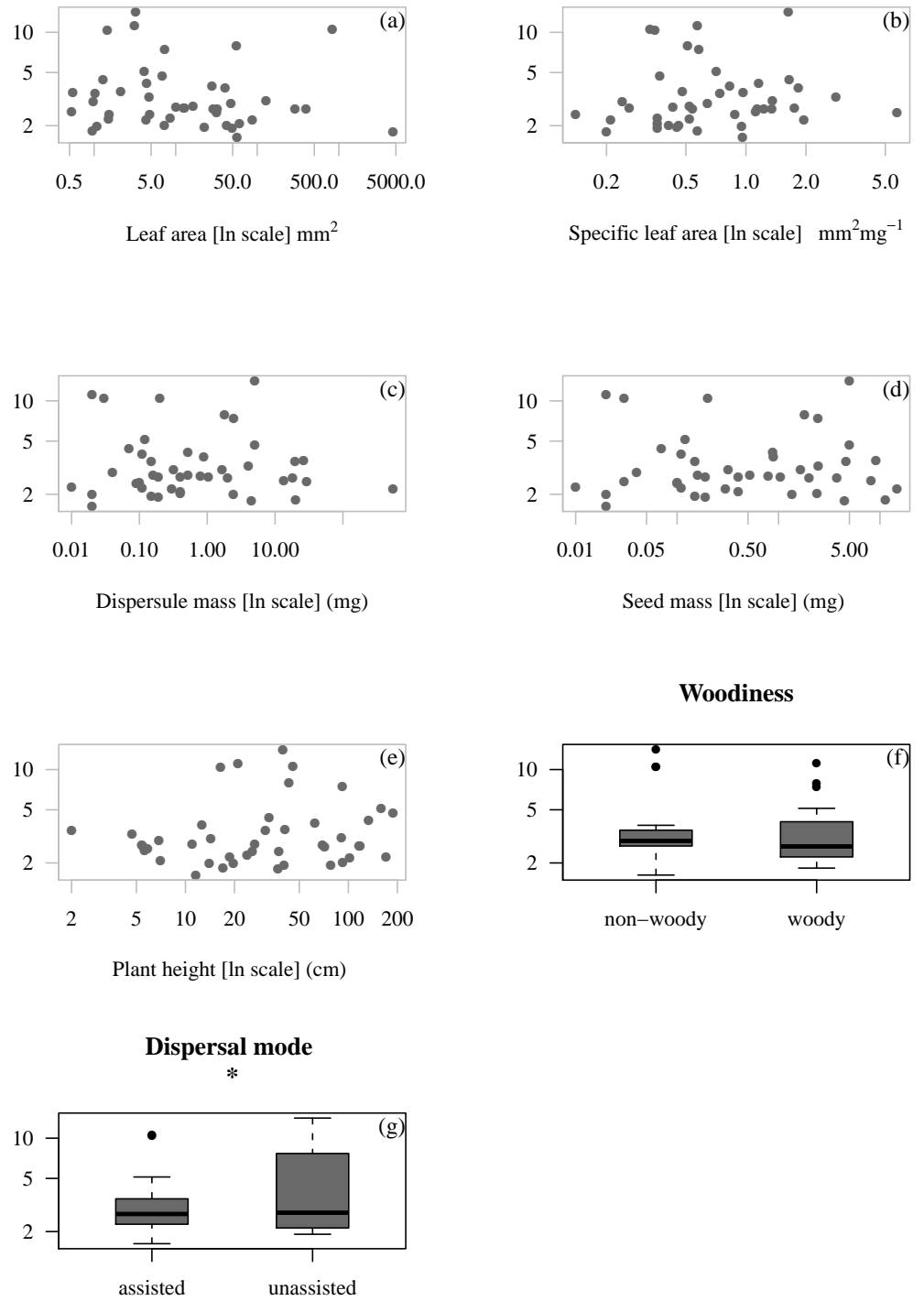
Figure 22: Scatter and boxplots of variables used as predictors of distribution patterns (aggregation $1/k$) across all scales for cross-species. In the scatter plots each point represents one species. In the boxplots the dark line is the median, the box encloses the interquartile range and the whiskers show the full range, with outliers shown as circles. Significant results are indicated as * $P < 0.05$

Table 7: Regression coefficients for univariate traits for cross-species analyses at local, regional scale and across scales. Trait type codes: C. Continuous; N. Nominal.

TRAIT	TYPE	ATTRIBUTES	$1/k$ LOCAL SCALE			$1/k$ REGIONAL SCALE			$1/k$ NBD2		
			b	R^2	P	b	R^2	P	b	R^2	P
Leaf area	C	log transformed (ln)	-0.13	0.06	0.23	-0.15	0.10	0.13	-0.03	0.016	0.40
Specific leaf area (SLA)	C	log transformed (ln)	1.03	0.25	**0.01**	-0.16	0.01	0.70	0.04	0.003	0.69
Dispersule mass	C	log transformed (ln)	0.11	0.06	0.25	-0.01	0.01	0.61	-0.01	0.004	0.67
Seed mass	C	log transformed (ln)	0.04	0.01	0.76	-0.08	0.02	0.50	-0.01	0.001	0.96
Plant height	C	log transformed (ln)	0.31	0.07	0.21	0.50	0.20	**0.036**	-0.01	0.001	0.37
Woodiness	N	2 attributes: non-woody and woody species			0.50			0.07			0.38
Dispersal mode	N	2 attributes: assisted and unassisted			0.12			0.26			**0.04**

Multivariate models

Multivariate analyses with all the predictor plant traits in the model suggested not a single significant variable at local ($P = 0.19$, $R^2 = 0.18$) and regional scale ($P = 0.14$, $R^2 = 0.22$).

MAM *minimum adequate model*

At local-scale, the best model contained the leaf area, SLA and dispersule mass variables ($P = 0.017$, $R^2 = 0.32$, Table 8). Leaf area and SLA had an additive effect that affected local aggregation, especially when taking into account dispersule mass. At regional scale, two traits interacted to affect regional aggregation ($P < 0.01$, $R^2 = 0.34$, Table 8); seed mass and plant height were the traits that contributed to the best model.

Multivariate analyses indicated that all life-history traits analysed, except for SLA, contributed and had an additive effect in explaining aggregation when measured across all four scales ($P = 0.02$, $R^2 = 0.20$; NBD2; Table 8).

Table 8: Regression coefficients for multivariate analyses at local and regional scale. ΔAIC is the difference in the Akaike's information criterion between the full model (upper) and the minimum adequate model (lower) at each scale.

DEPENDENT VARIABLE	PREDICTOR VARIABLE	ΔAIC	P - PREDICTOR VARIABLE	MODEL CO-EFFICIENTS
Local scale: mean aggregation ($1/k$) within plots, (20×20 m) sample units	Leaf area, SLA, dispersule, seed, height, woodiness, dispersal mode		none of the terms were significant (NS)	$P = 0.19$ $R^2 = 0.18$
	Leaf area, SLA, dispersule	6.83	Leaf area < 0.05 SLA < 0.05 Dispersule $=$ NS	$P = 0.017$ $R^2 = 0.32$
Regional scale: aggregation ($1/k$) among plots, (100×60 m) sample units	Leaf area, SLA, dispersule, seed, height, woodiness, dispersal mode		none of the terms were significant (NS)	$P = 0.14$ $R^2 = 0.22$
	Seed, height	6.70	Seed < 0.05 Height < 0.01	$P = 0.0074$ $R^2 = 0.34$
NBD2: aggregation ($1/k$)	Leaf area, SLA, dispersule, seed, height, woodiness, dispersal mode		SLA $=$ NS Dispersule $=$ NS Seed $=$ NS Leaf area $= 0.01$ Height $= 0.02$ Woodiness $= 0.002$ Dis.mode $= 0.003$	$P = 0.038$ $R^2 = 0.18$
	Leaf area, dispersule, seed, height, woodiness, dispersal mode	1.97	Leaf area $= 0.01$ Dispersule $=$ NS Seed $=$ NS Height $= 0.02$ Woodiness $= 0.001$ Dis.mode $= 0.003$	$P = 0.020$ $R^2 = 0.20$

3.3.3   *Relationships between plant traits and aggregation – phylogenetically independent contrast analyses*

At a local-scale, the PIC analyses identified plant traits associated with local aggregation. Divergence in leaf area exhibited a significant negative association ($R^2 = 0.27$; $P < 0.05$; Figure 23a), while the divergence in dispersule had a significant positive relationship ($R^2 = 0.40$; $P < 0.001$; Figure 23c). Species with small leaves were more aggregated locally. Conversely, species with smaller dispersule mass tended to have more even or random aggregation. There was also a tendency for short plants to have more aggregated patterns, although it was not significant ($P = 0.09$; Figure 23e and Table 9).

None of the plant traits was an important or significant explanatory variable in predicting regional aggregation (Figure 24 and Table 9), although there was a strong tendency for species comprising plants with heavier dispersules to be aggregated ($P < 0.08$; Figure 23c).

Aggregation ($1/k$ from the NBD2) measured across the four scales was negatively related to divergences in leaf area ($R^2 = 0.09$, $P = 0.03$; Figure 25a), dispersule mass ($R^2 = 0.11$, $P = 0.02$; Figure 25c) and seed mass ($R^2 = 0.14$, $P < 0.01$; Figure 25d).

The results of the phylogenetically independent contrast analyses were dissimilar to the cross-species analyses. This suggests that the cross-species relationships were an artefact generated by one or more divergences deep in the tree.

The most important divergence for the leaf area trait was at the dichotomy in the genus *Aciphylla* with larger leaves (contribution = 0.49). It has been suggested that large spine-like leaflets of the *Aciphylla* genus are xerophytic adaptations to dry climatic conditions (Oliver 1956). The biggest contributing divergences for SLA were the divergence at the base of *Gaultheria* and at the Ericaceae. The two gaultherias surveyed belong to the few that occur in alpine communities (Dawson 1988). Their leaves are short, broad and and have low SLA, probably another adaptation to alpine conditions (contribution = 0.37).

For seed mass, the highest divergence contribution was between angiosperms (smaller seeds) and gymnosperms (larger seeds, contribution = 0.34). This result is consistent with what Moles et al. (2005) found. The ability of angiosperms to produce smaller seeds than gymnosperms might be associated with the reduction in the gametophyte and early abortion (Haig & Westoby 1991).
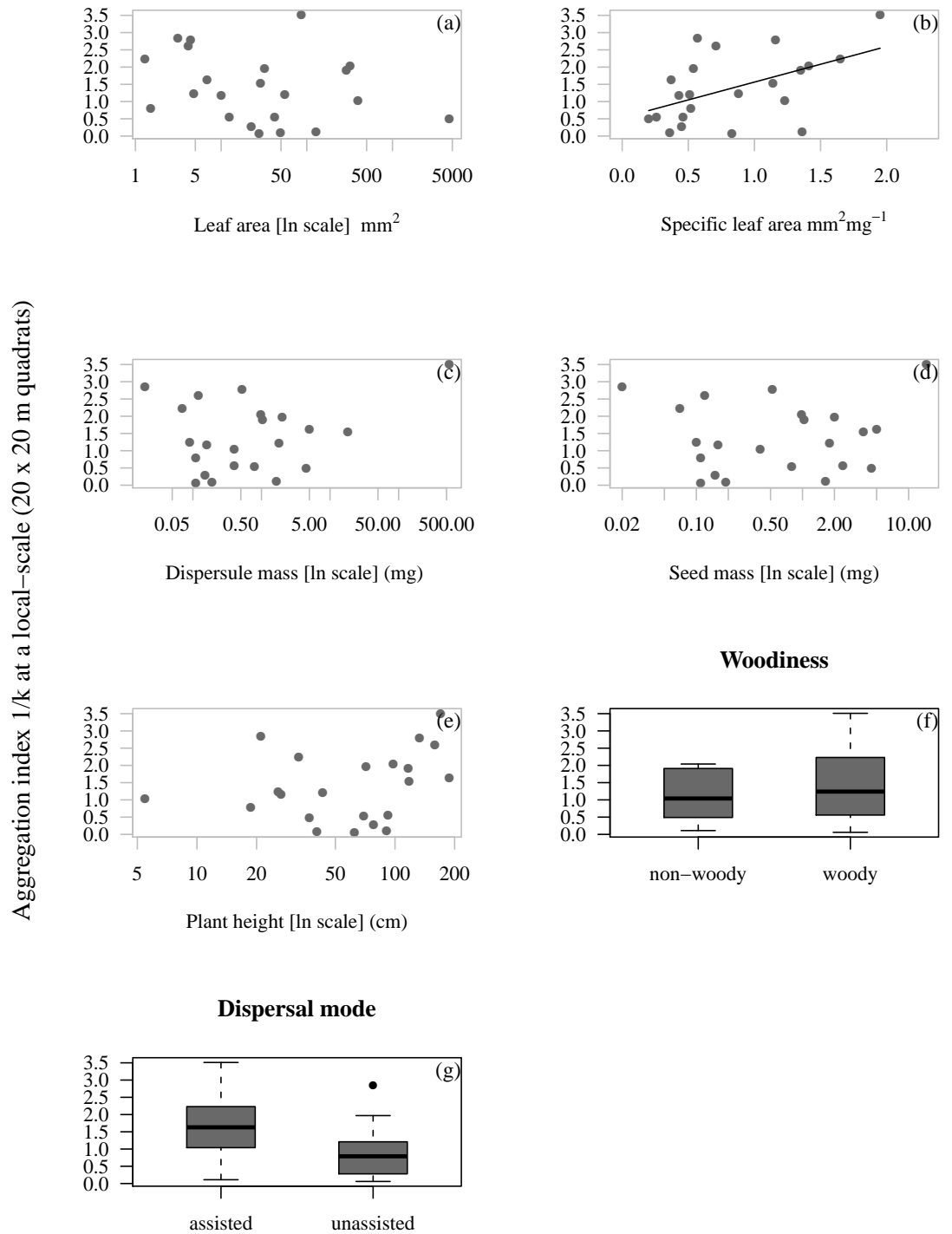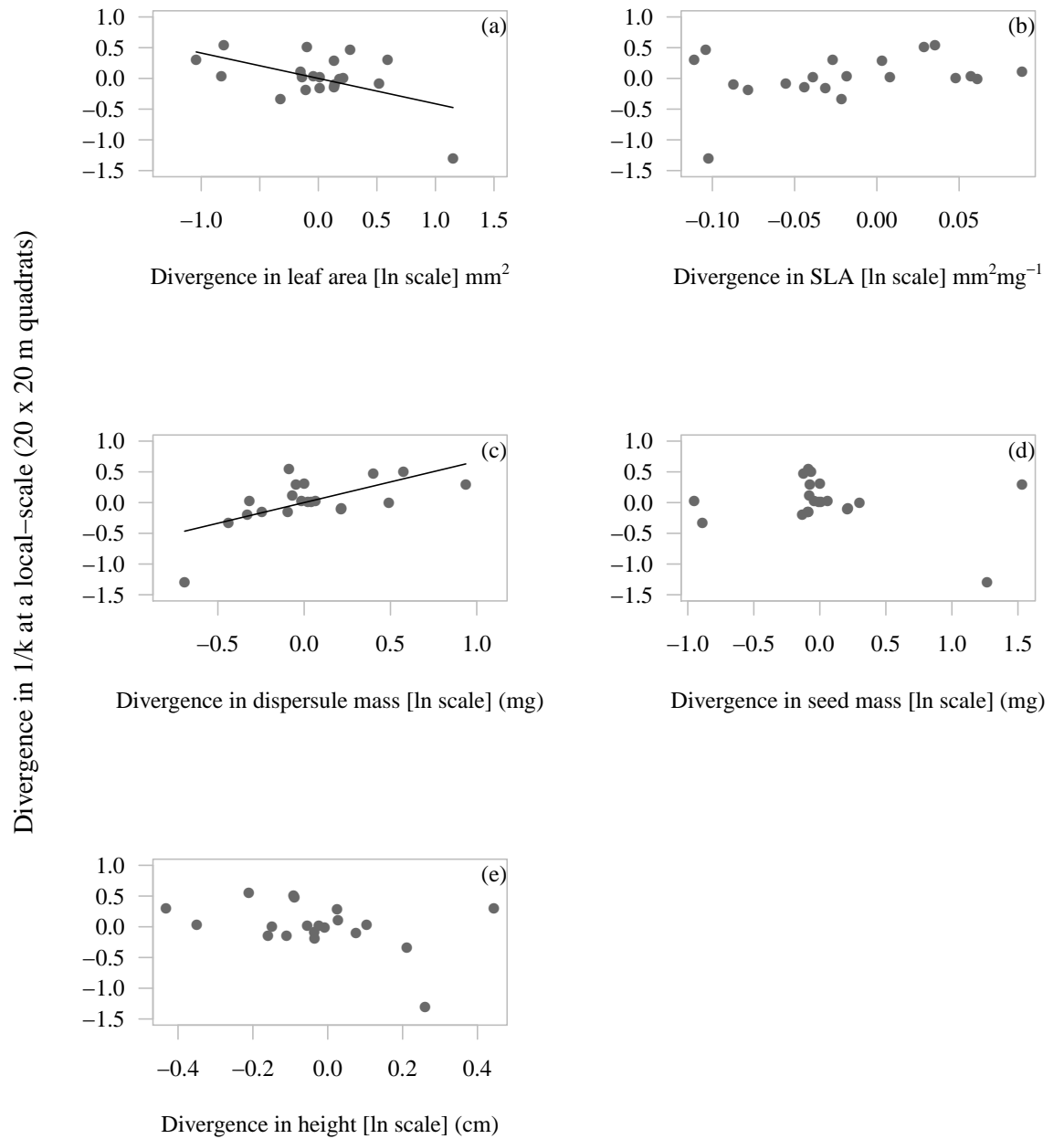
Figure 23: Scatter plots of variables used as predictors of distribution patterns (aggregation $1/k$) at local-scale for PIC analysis. Solid regression lines represent significant linear relationships, and dashed lines represent strong trends.
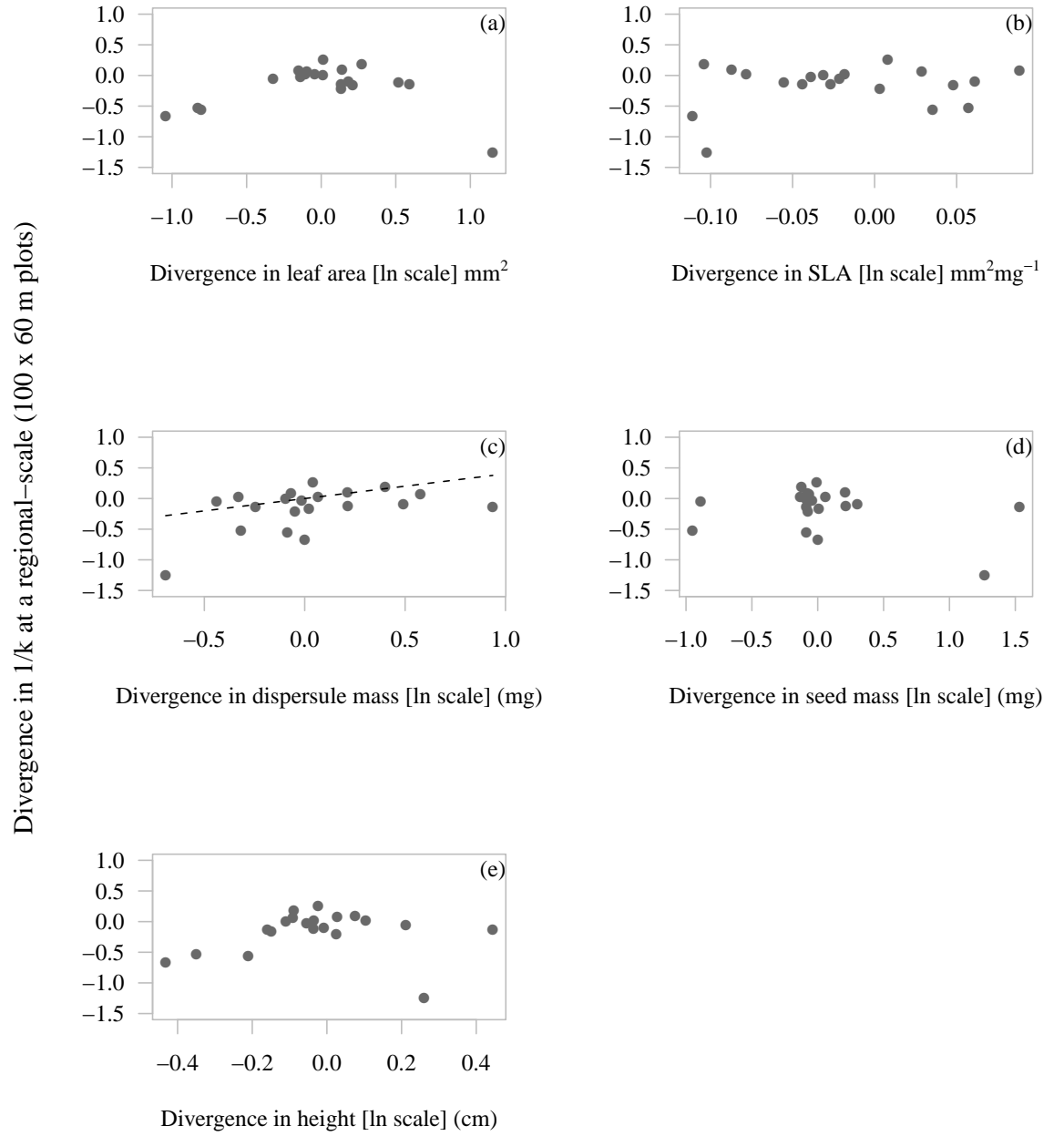
Figure 24: Scatter plots of variables used as predictors of distribution patterns (aggregation $1/k$) at regional-scale for PIC analysis. Solid regression lines represent significant linear relationships, and dashed lines represent strong trends.
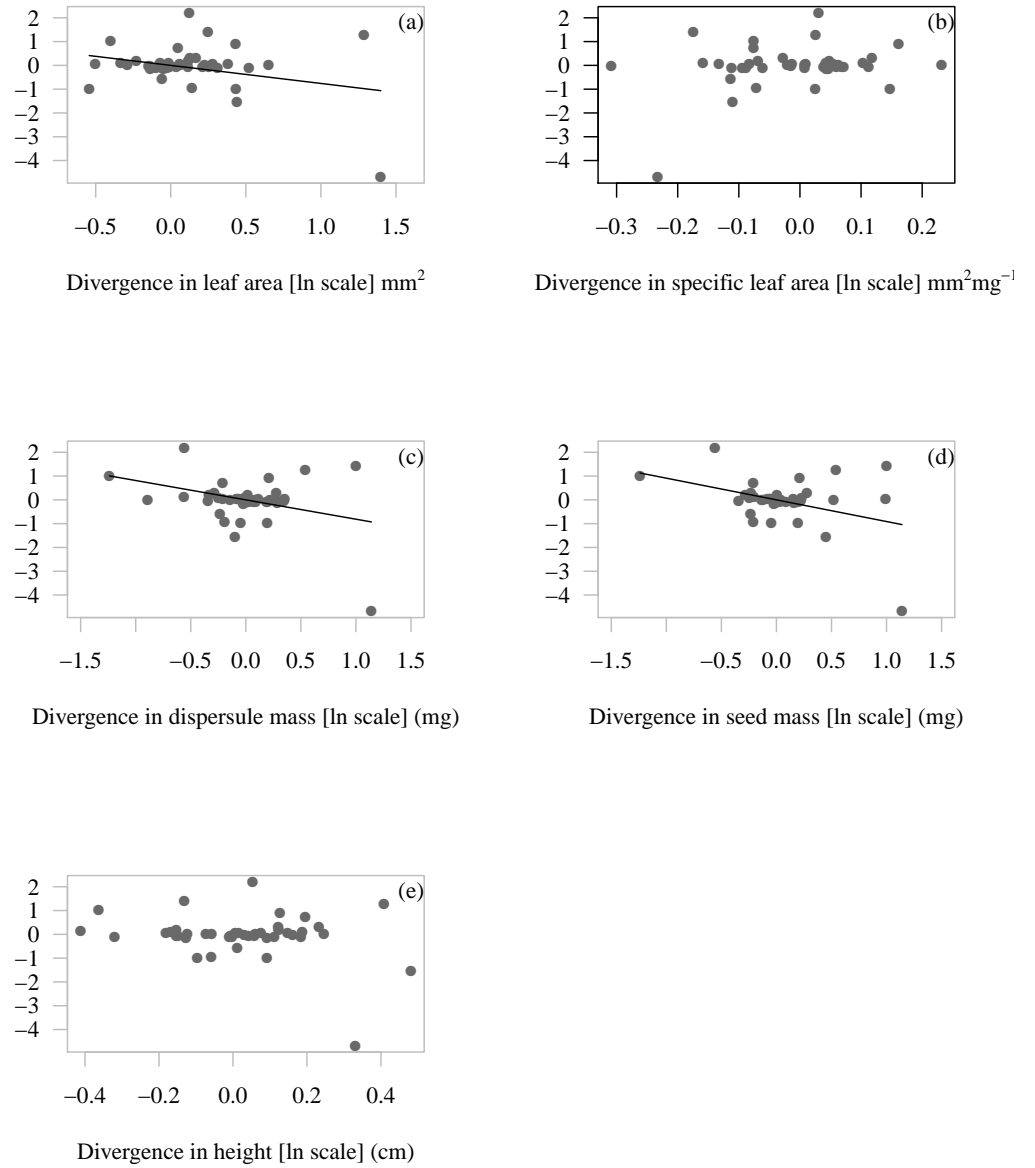
Figure 25: Scatter plots of variables used as predictors of distribution patterns (aggregation $1/k$) across all four scales for PIC analysis. Solid regression lines represent significant linear relationships, and dashed lines represent strong trends.

Table 9: Regression coefficients for univariate traits for PIC analyses at local and regional scale. Trait type codes: C. Continuous.

TRAIT	TYPE	ATTRIBUTES	$1/K$ LOCAL SCALE			$1/K$ REGIONAL SCALE			$1/K$ NBD2		
			b	R^2	P	b	R^2	P	b	R^2	P
Leaf area	C	log transformed (ln)	-0.41	0.27	**0.015**	-0.01	0.01	0.97	-0.75	0.09	**0.03**
Specific leaf area (SLA)	C		1.78	0.08	0.21	1.8	0.08	0.19	2.31	0.07	0.08
Dispersule mass	C	log transformed (ln)	0.67	0.40	**0.001**	0.40	0.15	0.08	-0.81	0.11	**0.02**
Seed mass	C	log transformed (ln)	-0.18	0.06	0.25	-0.21	0.09	0.18	-0.91	0.14	**0.009**
Height	C	log transformed (ln)	-0.74	0.14	0.09	-0.29	0.02	0.54	-1.2	0.05	0.11

Table 10: Important divergences in plant life-history traits that make the largest contribution to present-day variation for the data set used

TRAIT	RANK	DIVERGENCES	NODE
Leaf area	1	Dichotomy in *Aciphylla*	13
	2	Asterales versus Apiales	12
Specific leaf area	1	Dichotomy in *Gaultheria*	24
	2	Polytomy in Ericaceae	22
Dispersule mass	1	Dichotomy in Rosaceae	7
	2	Polytomy in Rosales	6
	3	Eudicots versus monocots	2
Seed mass	1	Angiosperms versus gymnosperms	1
	2	Polytomy in Ericaceae	22
	3	Dichotomy in Rosaceae	7
Height	1	Angiosperms versus gymnosperms	1
	2	Dichotomy in Rosaceae	7

The most important contribution to the divergence in dispersule mass was at the base of the Rosaceae (contribution = 0.50) followed by the Rosales (contribution = 0.15). At this node two species were surveyed and had large dispersules. The naturalised *Rosa rubiginosa* and the native *Acaena inermis* are both adapted for dispersal by birds and domestic stock. Although, *Rosa rubiginosa* is dispersed internally and *Acaena inermis* externally. The full table of contribution index values is given in Appendix D.

The highest contribution in the divergence of plant height was between the angiosperms and gymnosperms (contribution = 0.31), followed by divergence in the Rosaceae (contribution = 1.2). Divergences in height were in the same position as those in seed and dispersule mass, and might be a result of the association between larger dispersule mass with larger growth forms.

3.3.4 *Comparison between the aggregation parameter 1/k from the negative binomial 1 and the negative binomial 2 method*

The local and regional aggregation ($1/k$) relationship was plotted in order to detect if local scale $1/k$ predicts regional scale $1/k$; nevertheless they were not correlated (Figure 26a and Table 11).

The aggregation parameter $1/k$ measured across all four scales using the occupancy-area model from the negative binomial 2 was correlated

Table 11: Relationship and coefficients from model II regression in Figure 26.

	$1/k$ LOCAL			$1/k$ REGIONAL		
	P		R^2	P		R^2
	2-tailed	1-tailed		2-tailed	1-tailed	
$1/k$ REGIONAL	0.58	0.29	0.014			
$1/k$ ALL SCALES (NBD2)	**0.03**	**0.01**	0.21	**0.05**	**0.02**	0.16

positively with $1/k$ local and regional scale from the negative binomial 1 (Figure 26b and 26c, Table 11). This means that aggregation measured from occupancy-area models could be an option for measuring distribution patterns (aggregation) when data is only available in the presence/absence format.
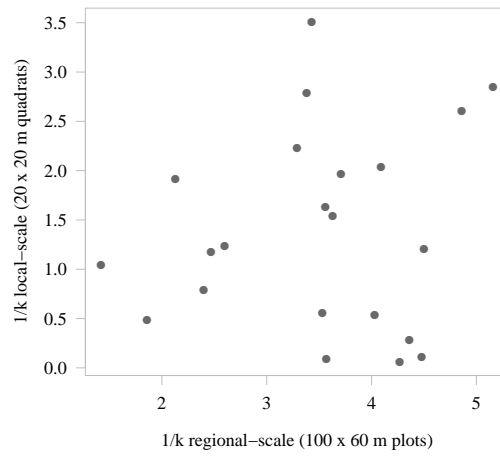
(a)
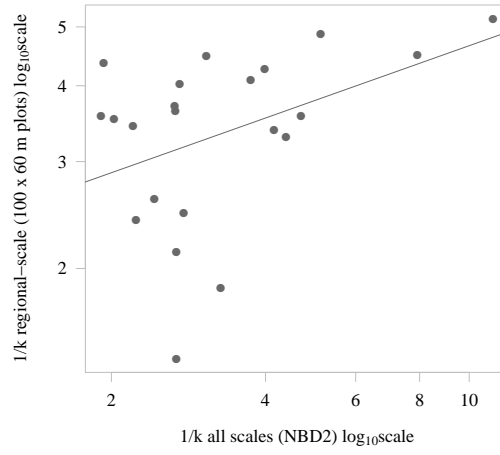
(b)
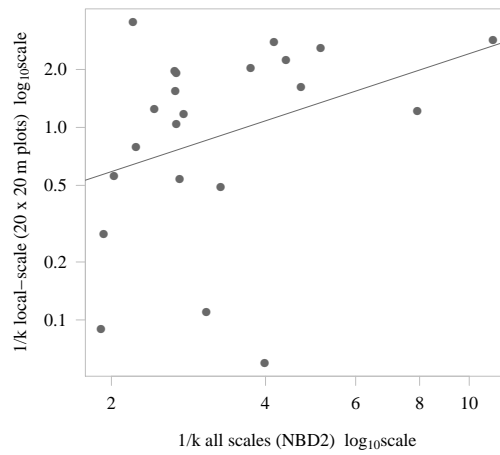
(c)

Figure 26: Regressions model II for the aggregation parameter $1/k$ between the negative binomial 1 and the negative binomial 2.

3.4 DISCUSSION

These aggregation patterns were correlated with some plant life-history traits and the significance of these correlations varied with the scale and the type of analyses used (Table 12).

3.4.1 *Plant traits related to aggregation and to spatial scales*

In general, plant traits had low predictive ability in explaining the degree of aggregation for cross-species analyses. The only plant traits considered that had a significant positive effect on aggregation were SLA (at local-scale), height (at regional-scale) and dispersal mode (across all scales). But these relationships may be spurious if generated by an alternative mechanism associated with one or more divergences deep in the phylogenetic tree. We cannot ignore this possibility since results from phylogenetic independent contrast analyses were inconsistent with cross-species relationships. Traits that were significant in cross-species analyses (SLA, height and dispersal mode) moved from significant to non-significant in PIC analyses. And *vice versa*, traits that were significant in PIC analyses were non-significant in cross-species analyses. One possibility is that there are one or more divergences deep in the phylogeny that underlie the present-day associations. Important divergences were identified in the phylogeny in comparison to the overall variance in trait values across all present-day species in the data set (the contribution index, Moles et al. 2005, Table 10).

SLA *specific leaf area*

   In the cross-species analyses SLA was positively correlated with aggregation at a local-scale. The forces that maintain the present-day specific leaf area variation might be a result of herbivory (Moles et al. 2011), or of abiotic conditions (Close et al. 2003). The results of this study showed that species comprising plants with lower SLA were more scattered. It is likely that these patterns were driven by environmental factors. It is known that exposure to UV and colder conditions (as those found in Molesworth Station) influence leaf toughness (Jordan et al. 2005). These conditions may explain the scattered patterns of species comprising plants with lower SLA which are very likely adapted to high UV exposure and other aggressive abiotic factors, and therefore can spread and survive in wider ranges. Whereas plants with aggregated distribution and high SLA values, can only persist in restricted, but favourable places in such hostile environment.

Table 12: Summary of the plant life-history traits that were related to the aggregation index $1/k$ and the type of analysis. Signs+ and − indicates the direction of the relationship, NS were non-significant relationships and NA were the cases where the analysis could not be performed .

PLANT TRAIT	$1/k$ LOCAL			$1/k$ REGIONAL			$1/k$ MULTI-SCALE (NBD2)		
	CROSS-SPECIES	PICs	MULTIVARIATE (MAM)	CROSS-SPECIES	PICs	MULTIVARIATE (MAM)	CROSS-SPECIES	PICs	MULTIVARIATE (MAM)
Leaf area	NS	-	-	NS	NS	NS	NS	-	-
Specific leaf area	+	NS	+	NS	NS	NS	NS	NS	NS
Dispersule mass	NS	+	NS	NS	NS	NS	NS	-	NS
Seed mass	NS	NS	NS	NS	NS	+	NS	-	NS
Plant height	NS	NS	NS	+	NS	+	NS	NS	+
Woodiness	NS	NA	NS	NS	NA	NS	NS	NA	+
Dispersal mode	NS	NA	NS	NS	NA	NS	+	NA	+

Species comprising tall plants tend to have more scattered distribution at regional-scale. Plant height is associated with competitive vigour and plant fecundity; also some tall plants may successfully avoid fire reaching the green parts in the canopy (Cornelissen et al. 2003). The findings of this Chapter are consistent with previous reports from Australia and Europe (Murray et al. 2002b; Van der Veken et al. 2007) which have noted positive relationships between plant height and range size.

PIC analyses identified positive associations between dispersule mass and aggregation at local-scale and negative associations between dispersule and seed mass and aggregation across all scales. Seed mass affects distribution and abundance patterns in different ways: plants with smaller seeds produce more propagules per unit canopy area than do plants with larger seeds; smaller seeds disperse more readily than larger seeds; large seeds spend longer time exposed to predispersal seed predators (Moles & Westoby 2003) and larger seeds could suffer greater and more intensive selective predation by animals (Eriksson & Jakobsson 1998; Guo et al. 2000). However, in a more recent study (Moles et al. 2003) no relationship was found between either seed mass and pre-dispersal or post-dispersal survivorship. Nevertheless, the most obvious explanation is that lighter dispersules can travel further distances and hence generate a more scattered distribution of seedlings (but see Thomson et al. in press).

PIC *phylogenetic independent contrast*

PIC and multivariate analyses identified negative relationships between leaf area and aggregation. Leaf area gave the more consistent negative association. The forces that maintain the present-day leaf area variation might be a result of herbivory (Moles et al. 2011). Species comprising plants larger leaves were more aggregated. This may be explained by the Janzen Connell effect (Janzen 1970), which suggests that species that suffer high levels of herbivory and parasite attack will suffer density-dependent mortality, this counteracts the tendency for seedlings to be aggregated close to their parent.

3.4.2 *Types of distribution patterns and rarity*

Different types of rarity were recognised for the surveyed data set and more importantly, the continuous measure of aggregation used reflected this variation. At a regional-scale (100×60 m), locally abundant species with small populations (*LO*) tended to have $1/k$ values
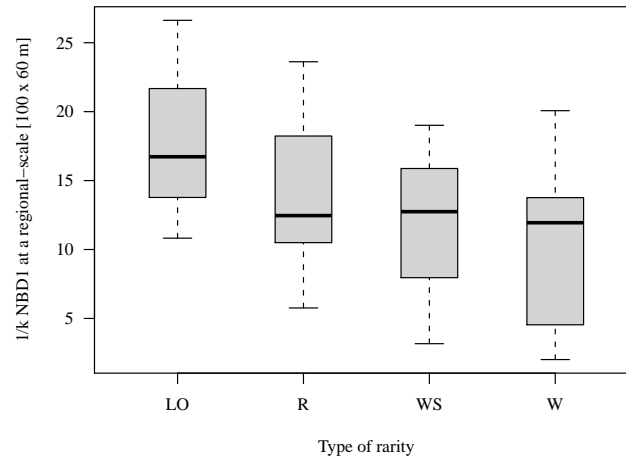
Figure 27: Boxplots of different types of distribution and rarity and aggregation $1/k$ from the NBD1. *R*, species with small populations and in low densities; *LO*, species locally abundant but with narrow ranges; *WS*, species that are widespread but with small populations; *W*, species that are both regionally widespread and locally abundant. Box = interquartile range, containing 50% of values; line across box = median; whiskers = highest and lowest values.

ranging from 11 to 26, while widespread but scattered species (*WS*) tended to have values ranging from 3 to 19 from the NBD1. Although there were no significant differences ($P = 0.30$) between the different types of distribution at regional-scale and the aggregation parameter $1/k$ (NBD1), there was a tendency for *LO* and *R* to have more aggregated patterns, and for *WS* and *W* to have more scattered patterns (Figure 27). There was a discrepancy between $1/k$ values from the NBD1(Table 6) and values from the NBD2 (isolines in Figure 19) for the different types of distribution and rarity. This may be caused because some species did not follow the negative binomial distribution. In order to identify this discrepancy I compared the frequency between the expected counts from the negative binomial versus the counts observed (Figure 28). Examples of this are *Aciphylla aurea* (Figure 28a) and *Discaria toumatou* (Figure 28b) which followed the expected NBD, while *Gentianella corymbifera* (Figure 28c) followed the NBD to lesser extent. These differences between expected and observed counts yielded different values from NBD1 and NBD2 and the types of distribution and rarity.

This study demonstrates that even closely related species exhibit different distribution patterns and fit in different rarity groups (e.g.,
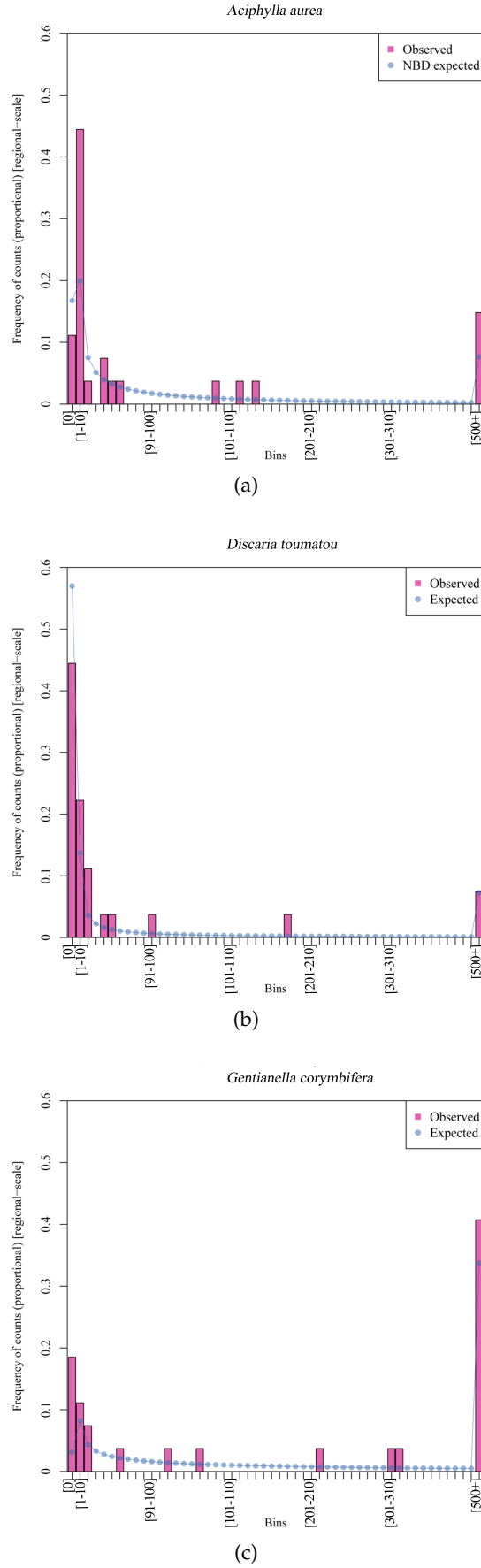
Figure 28: Frequency of counts between observed and expected (NBD model) counts across sample units (regional-scale) for (a) *Aciphylla aurea*; (b) *Discaria toumatou*; and (c) *Gentianella corymbifera*. Classes of bins range from 0 to 500+ in breaks of 10 individuals per bin. See Table 6 for the observed values of mean, variance and occupancy, and the expected $1/k$ at a regional-scale for the NBD1.

Veronica spp, *Chionochloa* spp; Figure 18 y 19, Table 5 and 6). This was consistent with the findings of Lloyd et al. (2003) who analysed the variation on distribution patterns of nine species of *Chionochloa* from range-restricted to widespread distribution patterns. Another study of *Acaena* species demonstrated that the distribution pattern varied greatly within the genus (Lee et al. 2001).

Distribution pattern also varied across scales. For instance, *Celmisia monroi* at a local-scale was highly restricted (*R*) and at a regional-scale the AOO was higher but still in low densities (*WS*). These findings suggest that several scales are needed to fully describe the distribution of even a single species.

3.4.3 *Comparison between NBD1 and NBD2 for measuring aggregation*

The new method for measuring aggregation that considered occupancy (NBD2) was comparable to the more traditional method from the negative binomial that took count data into account (NBD1). In order to correctly model ecological patterns and infer mechanisms of species coexistence, there is a fundamental need to accurately describe the spatial distributions of species (Zillio & He 2010). The NBD1 and NBD2 were used for measuring aggregation ($1/k$). The NBD1 model is widely used in the literature and has been proved to work well; nevertheless there are some cautions to bear in mind when using it. For example when the sampling area is large the NBD is clearly inappropriate as the model takes into account an infinite area, this is demonstrated by Zillio & He (2010). The NBD1 tended to erroneously detect rare species as being random, this was the case of *Veronica rakaiensis* at local scale which had a very low mean abundance and was present only in 14 subquadrats out of 405; the $1/k$ value of -0.06 suggested a random pattern, however it was highly aggregated. This is due to the fact that $1/k$ of the NBD1 can only measure spatially implicit aggregation based on the values of variance and mean; it is blind to the actual locations of quadrats so cannot detect spatially explicit aggregation at the broader scale. Interestingly, the new method for measuring aggregation from the negative binomial 2 (NBD2) opens new possibilities for calculating aggregation from occupancy, including occupancies of zero at small scales, as might occur when plots are subsmapled at finer scales. In the majority of the larger extent data sets, species distributions are often summarised as presences and absences at a particular resolution. This method might have the advantage of using this type of data over the other negative binomial 1 (NBD1) that requires count data, which

NBD *negative binomial model*

are more costly to get and not always appropriate to all plant growth habits at fine scales. This new method for measuring aggregation could have the potential to compare studies of the different distribution patterns achieved by declining or expanding species and may provide general guidelines for the management and monitoring of rare and invasive species in the absence of longer term ecological studies.

3.4.4 *Methodological issues*

Over very large areas of the South Island the vegetative cover is a product of many years of repeated burning (McGlone 1989; McGlone & Basher 1995). Although the area of study has been under protection during the last decade by the Department of Conservation (DoC), the land suffered from loss of vegetation and severe erosion caused by overgrazing by sheep and rabbits and repeated burning of tussocklands (Moore 1976). Hence, present-day distribution patterns reflect the interplay of species intrinsic properties with the abiotic environment, the sympatric plant community and the prevailing disturbance regimes. Additionally, introduced and naturalised plants have had a profound influence on the present-day distribution of tussock grasslands. It is likely that a combination of factors is important for their successful invasion (Platt et al. 1995). However, there is still a need of more complete understanding of spatial patterns, the processes and the interactions driving distribution patterns. Here, it was of my interest to test if exotic and native plants display different distribution patterns; nevertheless the number ($n = 5$) for exotic plant species was low to make comparisons. Also, to detect which species are expanding ranges it would have been ideal to measure the temporal scale. Permanent quadrats were left in the study site to monitor these changes in long-term studies. Such knowledge will develop more robust predictions of all factors that would have an important influence on shaping distribution patterns in tussock grasslands.

3.4.5 *Implications and conclusions*

Knowledge about spatial distribution and the mechanisms that are causing these patterns across scales is essential for predicting the distribution of species in areas where there is no information available (Raxworthy et al. 2003) for planning priorities in the conservation of

the species across scales (Myers et al. 2000) and for identifying which processes and mechanisms are more important at the different resolutions. In this Chapter, I have demonstrated that plant traits do help explain different distribution patterns at different scales. One of the main conclusions was that leaf area was the most robust life-history trait overall in predicting the degree of local aggregation. Leaf area gave a consistent negative association in four of the nine analyses. These findings suggest that leaf area should be considered when relating plant traits to distribution patterns. A Multi-scale approach helps to generate refined hypotheses about which processes are important at which scales. Multi-scale studies have the advantage of being better able to detect scale-specific associations between pattern and process that are not always apparent at a single scale of study. Furthermore, cross-species and PIC analyses can help increase the robustness of cross-scale predictions.

3.5 REFERENCES

- Abràmoff M.D., Magalhaes P. & Ram S. (2004). Image processing with ImageJ. Biophotonics international, 11, 36-43.

- Allan H. (1961). Flora of New Zealand, Vol. I. Government Printer, Wellington.

- Blackburn, T.M. (2004) Method in macroecology. Basic and Applied Ecology, 5, 401-412.

- Bradshaw, C. J. A., Giam, X. L., Tan, H. T. W., Brook, B. W. & Sodhi, N. S. (2008) Threat or invasive status in legumes is related to opposite extremes of the same ecological and life-history attributes. Journal of Ecology, 96, 869-883.

- Bremer B. (2009). Asterids. In: The TimeTree of Life (eds. Hedges SB & Kumar S). Oxford University Press New York, pp. 177-187.

- Bremer B., Bremer K., Chase M., Fay M., Reveal J., Soltis D., Soltis P. & Stevens P. (2009). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. Botanical Journal of the Linnean Society, 161, 105-121.

- Cadotte M.W. (2006). Dispersal and species diversity: A meta-analysis. American Naturalist, 167, 913-924.

- Close, D., McArthur, C., Paterson, S., Fitzgerald, H., Walsh, A. & Kincade, T. (2003) Photoinhibition: a link between effects of the environment on eucalypt leaf chemistry and herbivory. Ecology, 84, 2952-2966.

- Condit, R., Ashton, P.S., Baker, P., Bunyavejchewin, S., Gunatilleke, S., Gunatilleke, N., Hubbell, S.P., Foster, R.B., Itoh, A., LaFrankie, J.V., Lee, H.S., Losos, E., Manokaran, N., Sukumar, R. & Yamakura, T. (2000) Spatial patterns in the distribution of tropical tree species. Science, 288, 1414-1418.

– Cornelissen J.H.C., Lavorel S., Garnier E., Diaz S., Buchmann N., Gurvich D.E., Reich P.B., ter Steege H., Morgan H.D., van der Heijden M.G.A., Pausas J.G. & Poorter H. (2003). A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. Australian Journal of Botany, 51, 335-380.

– Coulter J.D. (1969). The climate of Molesworth. Meteorological Service Miscellaneous Publication, New Zealand.

– Diaz S., Hodgson J.G., Thompson K., Cabido M., Cornelissen J.H.C., Jalili A., Montserrat-Marti G., Grime J.P., Zarrinkamar F., Asri Y., Band S.R., Basconcelo S., Castro-Diez P., Funes G., Hamzehee B., Khoshnevi M., Perez-Harguindeguy N., Perez-Rontome M.C., Shirvany F.A., Vendramini F., Yazdani S., Abbas-Azimi R., Bogaard A., Boustani S., Charles M., Dehghan M., de Torres-Espuny L., Falczuk V., Guerrero-Campo J., Hynd A., Jones G., Kowsary E., Kazemi-Saeed F., Maestro-Martinez M., Romo-Diez A., Shaw S., Siavash B., Villar-Salvador P. & Zak M.R. (2004). The plant traits that drive ecosystems: Evidence from three continents. Journal of Vegetation Science, 15, 295-304.

– Eriksson O. & Jakobsson A. (1998). Abundance, distribution and life histories of grassland plants: a comparative study of 81 species. Journal of Ecology, 86, 922-933.

– Feeny P. (1976). Plant apparency and chemical defense. Recent advances in phytochemistry, 10, 1-40.

– Felsenstein J. (1985). Phylogenies and the comparative method. American Naturalist, 125, 1-15.

– Forest F. & Chase M.W. (2009a). Eudicots. In: The TimeTree of Life (eds. Hedges SB & Kumar S). Oxford University Press New York, pp. 169-176.

– Forest F. & Chase M.W. (2009b). Eurosid I. In: The TimeTree of Life (eds. Hedges SB & Kumar S). Oxford University Press New York, pp. 188-196.

– Fortin M.J. & Dale M.R.T. (2005). Spatial Analysis: A guide for ecologist. Cambridge University Press., New York, United States of America.

– Gaston K.J. (1994). Rarity. Chapman and Hall, London, UK.

– Gaston K.J., Blackburn T.M., Greenwood J.J.D., Gregory R.D., Quinn R.M. & Lawton J.H. (2000). Abundance-occupancy relationships. J. Appl. Ecol., 37, 39-59.

– Gaston K.J. & Kunin W.E. (1997). Rare-common differences: an overview. In: The Biology of Rarity (eds. Kunin WE & Gaston KJ). Chapman & Hall London, pp. 3-11.

– Guisan A. & Thuiller W. (2005). Predicting species distribution: offering more than simple habitat models. Ecology Letters, 8, 993-1009.

– Guo Q.F., Brown J.H., Valone T.J. & Kachman S.D. (2000). Constraints of seed size on plant distribution and abundance. Ecology, 81, 2149-2155.

– Hackathon R., Bolker B., Butler M., Cowan P., Vienne D.d., Eddelbuettel D., Holder M., Jombart T., Kembel S., Michonneau F., Orme D., O'Meara B., Paradis E., Regetz J. & Zwickl D. (2010). phylobase: Base package for phylogenetic structures and comparative data. URL http://CRAN.R-project.org/package=phylobase

– Harvey P.H. & Pagel M.D. (1991). The comparative method in evolutionary biology. Oxford University Press, Inc, Oxford, England.

- He, F., Gaston, K.J. & Wu, J. (2002) On species occupancy-abundance models. Ecoscience, 9, 119-126.

- He F.L. & Gaston K.J. (2000). Estimating species abundance from occurrence. American Naturalist, 156, 553-559.

- He, F., Zhou, J. & Zhu, H. (2003) Autologistic regression model for the distribution of vegetation. Journal of agricultural, biological, and environmental statistics, 8, 205-222.

- Healy A.J. & Edgar E. (1980). Flora of New Zealand; Vol III. Government Printer, Wellington.

- Hedges S.B., Dudley J. & Kumar S. (2006). TimeTree: a public knowledge-base of divergence times among organisms. Bioinformatics, 22, 2971-2972.

- Hurlbert, A. H. & Haskell, J. P. (2003) The effect of energy and seasonality on avian species richness and community composition. The American Naturalist, 161, 83-97.

- Hwang, W. H. & He, F. (in press) Estimating abundance from presence/absence maps. Methods in Ecology and Evolution.

- Janzen, D.H. (1970) Herbivores and the number of tree species in tropical forests. American Naturalist, 501-528.

- Johnson P. & Brooke P. (1989). Wetland plants of New Zealand. DSIR Publishing, Wellington.

- Jombart T., Balloux F. & Dray S. (2010). adephylo: new tools for investigating the phylogenetic signal in biological traits. Bioinformatics, 26, 1907-1909.

- Jordan, G. J., Dillon, R. A. & Weston, P. H. (2005) Solar radiation as a factor in the evolution of scleromorphic leaf anatomy in Proteaceae. American Journal of Botany, 92, 789.

- Kembel S.W., Cowan P.D., Helmus M.R., Cornwell W.K., Morlon H., Ackerly D.D., Blomberg S.P. & Webb C.O. (2010). Picante: R tools for integrating phylogenies and ecology. Bioinformatics, 26, 1463-1464.

- Kunin W.E. & Gaston K.J. (1993). The Biology of Rarity - Patterns, Causes and Consequences. Trends in Ecology & Evolution, 8, 298-301.

- Lee, D.E., Lee, W.G. & Mortimer, N. (2001) Where and why have all the flowers gone? Depletion and turnover in the New Zealand Cenozoic angiosperm flora in relation to palaeogeography and climate. Australian Journal of Botany, 49, 341-356.

- Legendre P. & Desdevises Y. (2009). Independent contrasts and regression through the origin. Journal of Theoretical Biology, 259, 727-743.

- Lennon, J., Greenwood, J. & Turner, J. (2000) Bird diversity and environmental gradients in Britain: a test of the species–energy hypothesis. Journal of Animal Ecology, 69, 581-598.

- Linnaeus C. (1737). Flora Lapponica. Amstelaedami., 1759. Flora Alpina. Amoeniates Acadamicae, 4.

- Lloyd K.M., Wilson J.B. & Lee W.G. (2003). Correlates of geographic range size in New Zealand Chionochloa (Poaceae) species. Journal of Biogeography, 30, 1751-1761.

— Magallón S. (2009). Flowering plants (Magnoliophyta). In: The TimeTree of Life (eds. Hedges SB & Kumar S). Oxford University Press New York, pp. 161-165.

— Magallón S.A. & Sanderson M.J. (2005). Angiosperm divergence times: the effect of genes, codon positions, and time constraints. Evolution, 59, 1653-1670.

— Mathews, S. & Bonser, S.P. (2005) Life histories, ecological tolerance limits, and the evolution of geographic range size in Eucalyptus (Myrtaceae). Australian Journal of Botany, 53, 501-508.

— McGlone M. (1989). The Polynesian settlement of New Zealand in relation to environmental and biotic changes. New Zealand Journal of Ecology, 12, 115-130.

— McGlone M.S. & Basher L.R. (1995). The deforestation of the upper Awatere catchment, Inland Kaikoura Range, Marlborough, South Island, New Zealand. New Zealand Journal of Ecology, 19, 53-66.

— Moles A.T., Wallis I.R., Foley W.J., Warton D.I., Stegen J.C., Bisigato A.J., Cella Pizarro L., Clark C.J., Cohen P.S., Cornwell W.K., Edwards W., Ejrnæs R., Gonzales-Ojeda T., Graae B.J., Hay G., Lumbwe F.C., Magaña-Rodríguez B., Moore B.D., Peri P.L., Poulsen J.R., Veldtman R., von Zeipel H., Andrew N.R., Boulter S.L., Borer E.T., Campón F.F., Coll M., Farji-Brener A.G., De Gabriel J., Jurado E., Kyhn L.A., Low B., Mulder C.P.H., Reardon-Smith K., Rodríguez-Velázquez J., Seabloom E.W., Vesk P.A., van Cauter A., Waldram M.S., Zheng Z., Blendinger P.G., Enquist B.J., Facelli J.M., Knight T., Majer J.D., Martínez-Ramos M., McQuillan P. & Prior L.D. (2011). Putting plant resistance traits on the map: a test of the idea that plants are better defended at lower latitudes. New Phytologist, 191, 777-788.

— Moles A.T., Warton D.I. & Westoby M. (2003). Do small-seeded species have higher survival through seed predation than large-seeded species? Ecology, 84, 3148-3161.

— Moles A.T. & Westoby M. (2000). Do small leaves expand faster than large leaves, and do shorter expansion times reduce herbivore damage? Oikos, 90, 517-524.

— Moles A.T. & Westoby M. (2003). Latitude, seed predation and seed mass. Journal of Biogeography, 30, 105-128.

— Moore L.B. (1976). The changing vegetation of Molesworth Station, New Zealand, 1944 to 1971. New Zealand Department of Scientific and Industrial Research, Wellington.

— Moore L.B. & Edgar E. (1970). Flora of New Zealand; Vol II. Government Printer, Wellington.

— Muller Landau H.C., Wright S.J., Calderón O., Condit R. & Hubbell S.P. (2008). Interspecific variation in primary seed dispersal in a tropical forest. Journal of Ecology, 96, 653-667.

— Murray B.R., Thrall P.H., Gill A.M. & Nicotra A.B. (2002a). How plant life-history and ecological traits relate to species rarity and commonness at varying spatial scales. Austral Ecology, 27, 291-310.

— Murray B.R., Thrall P.H. & Lepschi B.J. (2002b). Relating species rarity to life history in plants of eastern Australia. Evolutionary Ecology Research, 4, 937-950.

– Myers N., Mittermeier R.A., Mittermeier C.G., da Fonseca G.A.B. & Kent J. (2000). Biodiversity hotspots for conservation priorities. Nature, 403, 853-858.

– Paradis E., Claude J. & Strimmer K. (2004). APE: analyses of phylogenetics and evolution in R language. Bioinformatics, 20, 289.

– Platt, K.H., Rose, A.B., & Frampton, C.M. (1995) Vegetation change over 25 years in a New Zealand short tussock grassland: effects of sheep grazing and exotic invasions. New Zealand Journal of Ecology, 19, 163-174.

– Pocock M.J.O., Hartley S., Telfer M.G., Preston C.D. & Kunin W.E. (2006). Ecological correlates of range structure in rare and scarce British plants. Journal of Ecology, 94, 581-596.

– Pole M. (1994). The New Zealand flora-entirely long-distance dispersal? Journal of Biogeography, 625-635.

– Quinn R.M., Gaston K.J. & Arnold H.R. (1996). Relative measures of geographic range size: empirical comparisons. Oecologia, 107, 179-188.

– R Development Core Team (2011). R: A Language and Environment for Statistical Computing, Vienna, Austria. URL http://www.R-project.org

– Rabinowitz D. (1981). Seven forms of rarity. In: The Biological Aspects of Rare Plant Conservation (ed. Synge H). Wiley Great Britain, pp. 205-217.

– Rabinowitz D., Cairns S. & Dillon T. (1986). Seven forms of rarity and their frequency in the flora of the British Isles. In: Conservation Biology The Science of Scarcity and Diversity (ed. Soulé ME). Sinauer Associates, INC Publishers Sunderland, Massachusetts, pp. 182-204.

– Réjou Méchain, M., Flores, O., Bourland, N., Doucet, J.L., Fétéké, R.F., Pasquier, A. & Hardy, O.J. (2011) Spatial aggregation of tropical trees at multiple spatial scales. Journal of Ecology, 99, 1373-1381.

– Raxworthy C.J., Martinez-Meyer E., Horning N., Nussbaum R.A., Schneider G.E., Ortega-Huerta M.A. & Peterson A.T. (2003). Predicting distributions of known and unknown reptile species in Madagascar. Nature, 426, 837-841.

– Van der Veken S., Bellemare J., Verheyen K. & Hermy M. (2007). Life-history traits are correlated with geographical distribution patterns of western European forest herb species. Journal of Biogeography, 34, 1723-1735.

– Schulze, E.-D., Beck, E. & Müller-Hohenstein, K. (2005) Plant ecology. Springer, New York.

– Sefton, C.A., Montagu, K., Atwell, B.J. & Conroy, J.P. (2002) Anatomical variation in juvenile eucalypt leaves accounts for differences in specific leaf area and CO_2 assimilation rates. Australian Journal of Botany, 50, 301-310.

– Sober E. (1991). Reconstructing the past: parsimony, evolution, and inference. The MIT Press, London, England.

– Thompson, K., Hodgson, J.G., Grime, J.P. & Burke, M.J.W. (2001) Plant traits and temporal scale: evidence from a 5-year invasion experiment using native species. Journal of Ecology, 89, 1054-1060.

– Thomson F.J., Moles A.T., Auld T.D. & Kingsford R.T. (in press). Seed dispersal distance is more strongly correlated with plant height than with seed mass. Journal of Ecology.

— Tilman, D. & Kareiva, P.M. (1997) Spatial ecology: the role of space in population dynamics and interspecific interactions. Princeton Univ Pr.

— Townsend A.J., de Lange P.J., Duffy C.A.J., Miskelly C.M., Molloy J. & Norton D.A. (2008). New Zealand threat classification system manual. In. New Zealand Department of Conservation, PO Box 10420, The Terrace, Wellington, 6143, New Zealand.

— Tremlova K. & Munzbergova Z. (2007). Importance of species traits for species distribution in fragmented landscapes. Ecology, 88, 965-977.

— Van der Veken, S., Bellemare, J., Verheyen, K. & Hermy, M. (2007) Life-history traits are correlated with geographical distribution patterns of western European forest herb species. Journal of Biogeography, 34, 1723-1735.

— Venable D.L., Flores-Martinez A., Muller-Landau H.C., Barron-Gafford G. & Becerra J.X. (2008). Seed dispersal of desert annuals. Ecology, 89, 2218-2227.

— Venables W.N. & Ripley B.D. (2002). Modern Applied Statistics with S. Fourth edn. Springer, New York.

— Wagstaff S.J., Bayly M.J., Garnock-Jones P.J. & Albach D.C. (2002). Classification, origin, and diversification of the New Zealand hebes (Scrophulariaceae). Annals of the Missouri Botanical Garden, 89, 38-63.

— Wardle P. (1991). Vegetation of New Zealand. Cambridge University Press, Cambridge.

— Webb C., Johnson P. & Sykes B. (1990). Flowering Plants of New Zealand. DSIR Botany, Christchurch, N.Z.

— Webb C.J. & Simpson M.J.A. (2001). Seeds of New Zealand gymnosperms and dicotyledons. Manuka Press, Christchurch, New Zealand.

— Webb C.J., Sykes W.R. & Garnock-Jones P.J. (1988). Flora of New Zealand; Vol IV. Scientific and Industrial Research, Christchurch.

— Webb C.O., Ackerly D.D. & Kembel S.W. (2008). Phylocom: software for the analysis of phylogenetic community structure and trait evolution. Bioinformatics, 24, 2098-2100.

— Webb C.O. & Donoghue M.J. (2005). Phylomatic: tree assembly for applied phylogenetics. Mol. Ecol. Notes, 5, 181-183.

— Westoby M., Leishman M.R. & Lord J.M. (1995). On misinterpreting the phylogenetic correction. Journal of Ecology, 83, 531-534.

— Westoby M. & Wright I.J. (2006). Land-plant ecology on the basis of functional traits. Trends in Ecology & Evolution, 21, 261-268.

— Wiens, J. (1989) Spatial scaling in ecology. Functional Ecology, 3, 385-397.

— Woodward, F.I. (1987) Plant Ecology - Trends and Traits. Trends in Ecology & Evolution, 2, 252-254

— Wright I.J., Ackerly D.D., Bongers F., Harms K.E., Ibarra-Manriquez G., Martinez-Ramos M., Mazer S.J., Muller-Landau H.C., Paz H., Pitman N.C.A., Poorter L., Silman M.R., Vriesendorp C.F., Webb C.O., Westoby M. & Wright S.J. (2007). Relationships among ecologically important dimensions of plant trait variation in seven Neotropical forests. Annals of Botany, 99, 1003-1015.

– Wright, I.J., Westoby, M. & Reich, P.B. (2002) Convergence towards higher leaf mass per area in dry and nutrient-poor habitats has different consequences for leaf life span. Journal of Ecology, 90, 534-543.

– Wu J., Jones K.B., Li H. & Loucks O.L. (2006). Scaling and uncertainty analysis in ecology: methods and applications. Springer Verlag.

– Young A.G. & Brown A.H.D. (1998). Comparative analysis of the mating system of the rare woodland shrub *Daviesia suaveolens* and its common congener *D. mimosoides*. Heredity, 80, 374-381.

– Zillio T. & He F.L. (2010). Modelling spatial aggregation of finite populations. Ecology, 91, 3698-3706.

3.6 APPENDIX S1

Table 13: Summary of the plant traits included in the data set.

TRAIT	DESCRIPTION	SCALE	MAIN SOURCES
Vegetative traits			
Leaf Area	One-sided projected surface area of an average leaf lamina, expressed in mm²	Continuous	Own measurements
SLA	One-sided area of fresh leaf divided by its dry weight, expressed in mm² mg⁻¹	Continuous	Own measurements
Woodiness	2 classes: 1, woody; 2, non woody	Nominal	(Webb et al. 1988; Allan 1961)
Average height	Mean height of plant individual in cm	Continuous	Own measurements, (Webb et al. 1990)
Reproductive trait			
Seed mass	Weight of one seed in mg	Continuous	Own measurements, (Webb & Simpson 2001)
Dispersule mass	Weight of one seed including the fruit or achene in mg	Continuous	Own measurements, (Webb & Simpson 2001)
Dispersal mode	2 classes: 1,assisted ; 2, unassisted	Nominal	(Webb et al. 1988; Webb et al. 1990)

3.7 APPENDIX S2

Table 14: Average leaf area, specific leaf area (SLA), dispersule and seed mass for 46 plant species (±SD). The number of individuals sampled are shown in the parenthesis followed by the total number of leaves or seeds sampled.

SPECIES	LEAF AREA (mm²)	SLA (mm²mg⁻¹)	DISPERSULE (mg)	SEED (mg)	HEIGHT (cm)
Acaena inermis	4.74 ± 1.32 (5,50)	2.83 ± 0.99 (5,50)	4.04 ± 1.13 (5,100)	2.47 ± 1.13 (5,100)	4.7 ± 1.8 (5)
Aciphylla aurea	4595.75 ± 214.01 (4,8)	0.2 ± 0.01 (4,8)	4.48 ± 0.71 (4,120)	4.48 ± 0.71 (4,120)	37 ± 4.97 (4)
Aciphylla monroi	16.32 ± 3.07 (5,10)	0.52 ± 0.05 (5,10)	0.52 ± 0.9 (5,50)	0.52 ± 0.9 (5,50)	11 ± 2.92 (5)
Acrothamnus colensoi	0.95 ± 0.21 (5,50)	0.57 ± 0.05 (5,50)	20.11 ± 4.94 (5,50)	11.27 ± 2.57 (5,50)	16.98 ± 4.83 (5)
Anthoxanthum odoratum	40.23 ± 15.56 (5,13)	1.83 ± 0.23 (5,13)	0.89 ± 0.04 (5,50)	0.89 ± 0.04 (5,50)	12.6 ± 1.82 (5)
Brachyglottis cassinioides	0.88 ± 0.19 (5,50)	1.16 ± 0.16 (5,50)	2.06 ± 0.14 (5,100)	0.88 ± 0.14 (5,100)	133 ± 16.91 (5)
Celmisia du-rietzii	56.35 ± 3.61 (4,20)	0.96 ± 0.06 (4,20)	0.02 ± 0.08 (3,60)	0.02 ± 0.08 (3,60)	11.6 ± 1.76 (2)
Celmisia incana	47.67 ± 14.84 (3,30)	0.64 ± 0.05 (3,30)	0.04 ± 0.01 (3,60)	0.04 ± 0.01 (3,60)	6.87 ± 0.57 (3)
Celmisia monroi	7.3 ± 1.04 (3,30)	0.41 ± 0.03 (3,30)	0.02 ± 0.01 (5,50)	0.02 ± 0.01 (5,50)	19.7 ± 0.78 (5)
Celmisia semicordata	829.44 ± 17.16 (4,20)	0.33 ± 0.01 (4,20)	0.03 ± 0.29 (4,80)	0.03 ± 0.29 (4,80)	45.75 ± 5.12 (4)
Celmisia spectabilis	60.36 ± 10.53 (5,50)	0.36 ± 0.05 (5,50)	0.4 ± 0.2 (5,50)	0.4 ± 0.2 (5,50)	7.02 ± 1.29 (5)
Chionochloa sp1	321.4 ± 35.09 (5,50)	1.41 ± 0.17 (5,50)	0.98 ± 0.07 (5,50)	0.98 ± 0.07 (5,50)	97.8 ± 13.34 (5)
Chionochloa australis	0.98 ± 0.18 (5,50)	0.24 ± 0.03 (5,50)	0.32 ± 0.02 (5,50)	0.32 ± 0.02 (5,50)	14.34 ± 4.16 (5)
Chionochloa pallens	128.3 ± 33.87 (5,50)	1.36 ± 0.51 (5,50)	1.65 ± 0.08 (5,50)	1.65 ± 0.08 (5,50)	90.6 ± 10.75 (2)
Chionochloa rubra	290.42 ± 9.26 (5,50)	1.35 ± 0.27 (5,50)	1.04 ± 0.08 (5,50)	1.04 ± 0.08 (5,50)	116.6 ± 8.65 (5)
Discaria toumatou	28.82 ± 1.5 (5,50)	1.14 ± 0.13 (5,50)	18.11 ± 3.56 (5,25)	3.76 ± 1.48 (5,25)	118 ± 33.77 (5)

SPECIES	LEAF AREA (mm²)	SLA (mm²mg⁻¹)	DISPERSULE (mg)	SEED (mg)	HEIGHT (cm)
Dracophyllum pronum	1.53 ± 0.43 (5,50)	0.14 ± 0.03 (5,50)	0.1 ± 0.01 (5,50)	0.1 ± 0.01 (5,50)	37.5 ± 7.04 (5)
Dracophyllum rosmarinifolium	4.33 ± 0.79 (5,50)	0.21 ± 0.03 (5,50)	0.3 ± 0.01 (5,50)	0.3 ± 0.01 (5,50)	102 ± 18.28 (5)
Dracophyllum sp1	7.32 ± 1.06 (5,50)	0.58 ± 0.03 (5,50)	2.45 ± 0.01 (2,100)	2.45 ± 0.01 (2,100)	91.8 ± 14.65 (5)
Gaultheria crassa	8.55 ± 1.82 (5,50)	0.36 ± 0.07 (5,50)	0.01 ± 0 (5,300)	0.01 ± 0 (5,300)	23.88 ± 2.01 (5)
Gaultheria depressa	31.82 ± 24.7 (5,50)	5.72 ± 4.63 (5,50)	29.19 ± 4.55 (5,50)	0.03 ± 0.01 (5,250)	5.62 ± 1.01 (5)
Gentianella corymbifera	396.3 ± 48.03 (5,25)	1.23 ± 0.2 (5,25)	0.4 ± 0.09 (5,250)	0.4 ± 0.09 (5,250)	5.46 ± 2.32 (5)
Helichrysum parvifolium	1.29 ± 0.13 (5,50)	1.65 ± 0.44 (5,50)	0.07 ± 0.01 (5,50)	0.07 ± 0.01 (5,50)	32.64 ± 6.84 (5)
Hieracium pilosella	12.86 ± 2.45 (5,25)	1.75 ± 0.12 (5,25)	0.19 ± 0.02 (5,50)	0.19 ± 0.02 (5,50)	5.4 ± 2.07 (5)
Leptospermum scoparium	4.13 ± 0.21 (5,100)	0.71 ± 0.01 (5,100)	0.12 ± 0.01 (5,130)	0.12 ± 0.01 (5,130)	158.8 ± 12.74 (5)
Leucopogon fraseri	0.53 ± 0.13 (5,50)	1.12 ± 0.61 (5,50)	13.49 ± 0.36 (5,50)	8.18 ± 0.38 (5,49)	5.86 ± 1.59 (5)
Melicytus alpinus	1.03 ± 0.14 (2,20)	0.74 ± 0.12 (2,20)	19.75 ± 3.84 (2,20)	4.61 ± 0.4 (2,20)	31 ± 1.41 (2)
Olearia nummulariifolia	12.5 ± 0.63 (5,50)	0.26 ± 0.05 (5,50)	0.79 ± 0.1 (5,100)	0.79 ± 0.1 (5,100)	69.44 ± 18.41 (5)
Ozothamnus vauvilliersii	4.8 ± 0.21 (5,50)	0.88 ± 0.1 (5,50)	0.09 ± 0.09 (4,36)	0.1 ± 0.08 (4,34)	25.72 ± 4.88 (5)
Pimelea sericeovillosa	1.08 ± 0.21 (5,50)	0.95 ± 0.12 (5,50)	2.42 ± 0.45 (5,25)	1.36 ± 0.21 (5,25)	14.02 ± 6.98 (5)
Pinus contorta	6.85 ± 0.56 (5,50)	0.37 ± 0.02 (5,50)	5 ± 1.1 (5,25)	5 ± 1.1 (5,25)	187.7 ± 10.43 (5)
Podocarpus nivalis	2.13 ± 0.4 (5,50)	0.48 ± 0.09 (5,50)	26.26 ± 12.35 (5,25)	9.12 ± 2.89 (5,25)	40.8 ± 27.26 (5)
Raoulia grandiflora	0.55 ± 0.1 (5,100)	0.97 ± 0.09 (5,100)	0.15 ± 0.02 (5,50)	0.15 ± 0.02 (5,50)	2 ± 0.71 (5)
Rosa rubiginosa	17.32 ± 1.96 (5,50)	1.95 ± 0.06 (5,50)	545.88 ± 144.44 (5,15)	14.71 ± 3.89 (5,50)	65.4 ± 15.16 (5)
Rytidosperma setifolium	1.45 ± 0.71 (5,50)	0.35 ± 0.12 (5,50)	0.2 ± 0.01 (5,50)	0.2 ± 0.01 (5,50)	16.4 ± 3.36 (5)

SPECIES	LEAF AREA (mm²)	SLA (mm²mg⁻¹)	DISPERSULE (mg)	SEED (mg)	HEIGHT (cm)
Stellaria graminea	3.25 ± 0.99 (5,40)	1.63 ± 1.46 (5,40)	5 ± 0.9 (5,50)	5 ± 0.9 (5,50)	39.6 ± 5.5 (5)
Veronica brachysiphon	42.17 ± 2.75 (5,50)	0.46 ± 0.03 (5,50)	0.4 ± 0.01 (5,50)	2.4 ± 0.01 (5,50)	92.2 ± 24.35 (5)
Veronica canterburiensis	32.23 ± 1.08 (5,50)	0.54 ± 0.03 (5,50)	2 ± 0.07 (5,50)	2 ± 0.07 (5,50)	71.4 ± 14.4 (5)
Veronica cryptomorpha	55.4 ± 1.71 (5,50)	0.51 ± 0.01 (5,50)	1.8 ± 0.07 (5,50)	1.8 ± 0.07 (5,50)	43.2 ± 1.48 (5)
Veronica decumbens	49.4 ± 1.51 (5,50)	0.36 ± 0.03 (5,50)	0.19 ± 0.01 (3,97)	0.19 ± 0.01 (3,97)	40.36 ± 20.56 (5)
Veronica lycopodioides	1.51 ± 0.17 (3,30)	0.52 ± 0.06 (3,30)	0.11 ± 0.09 (5,50)	0.11 ± 0.09 (5,50)	18.67 ± 4.73 (3)
Veronica odora	22.53 ± 1.06 (5,50)	0.45 ± 0.06 (5,50)	0.15 ± 0.03 (5,50)	0.15 ± 0.03 (5,50)	78 ± 6.16 (5)
Veronica pimeleoides	3.12 ± 0.13 (5,50)	0.57 ± 0.03 (5,50)	0.02 ± 0 (2,41)	0.02 ± 0 (2,41)	21 ± 1.87 (5)
Veronica pinguifolia	10.07 ± 0.29 (5,50)	0.43 ± 0.04 (5,50)	0.16 ± 0.04 (5,117)	0.16 ± 0.04 (5,117)	26.6 ± 2.3 (5)
Veronica rakaiensis	27.79 ± 1.14 (5,50)	0.83 ± 0.06 (5,50)	0.11 ± 0.02 (4,87)	0.11 ± 0.02 (4,87)	62.38 ± 10.8 (5)

End

3.8 APPENDIX S3

Table 15: Estimated age for the 8 nodes numbered in Figure 17. References
are indicated for the estimated age.

NODE	ESTIMATED AGE (MYA)	REFERENCE
1	325	(Soltis et al. 2002)
2	147.8	(Magallón 2009)
3	125	(Forest & Chase 2009a)
4	107	(Forest & Chase 2009b)
5	98	(Forest & Chase 2009b)
6	76	(Forest & Chase 2009b)
7	38	Phylocom *(bladj)*
8	54.5	Phylocom *(bladj)*
9	122	(Forest & Chase 2009a)
10	118	(Bremer 2009)
11	90	(Bremer 2009)
12	115	Phylocom *(bladj)*
13	58	Phylocom *(bladj)*
14	123	(Bremer 2009)
15	114	(Bremer 2009)
16	76	Phylocom *(bladj)*
17	39	Phylocom *(bladj)*
18	38	Phylocom *(bladj)*
19	108	(Bremer 2009)
20	3.9	(Wagstaff et al. 2002)
21	1.95	Phylocom *(bladj)*
22	80	Phylocom *(bladj)*
23	41	Phylocom *(bladj)*
24	41	Phylocom *(bladj)*
25	40	Phylocom *(bladj)*
26	28	(Magallón & Sanderson 2005)
27	298	(Renner 2009)

4

LIFE-HISTORY TRAITS RELATED TO NATIONAL, REGIONAL AND LOCAL DISTRIBUTION PATTERNS IN THE *VERONICA* SECT. *HEBE* (PLANTAGINACEAE)

4.1 INTRODUCTION

Why are some species aggregated and localised while others are widely distributed or highly fragmented? This is a crucial question in ecology. To provide complete answers for even a single species in a single location, we would ideally need, a knowledge of the degree of aggregation, physicochemical conditions, the level of resources available, the organism's life cycle and the influence of competitors, predators and parasites (Begon et al. 2006). The range of potential processes shaping species' distribution patterns is nearly as varied and complex as the diversity of life (Holt et al. 2005). Processes inherent to the species (e.g., dispersal, competition) and processes independent of the species (e.g., disturbance, environmental conditions) can generate different types of distribution patterns (Fortin & Dale 2005). Quantifying aggregation levels of spatial patterns might be a first step in studies that aim to relate patterns to ecological processes (He et al. 2000). From the different distribution pattern indices that have been developed, the mass fractal dimension (D_m) was used in this study as an index of spatial aggregation since it can be applied over a specified range of scales and can describe major aspects of the range structure over several scales (Hartley & Kunin 2003).

The relationship between pattern and process is of great interest in all natural sciences, and scale is an integral part of this relationship. Scale usually refers to the spatial or temporal dimension of a phenomenon, and scaling is the transfer of information between scales (Wu et al. 2006); or in other words, the extent to which information at one scale can predict properties of pattern at another scale. Several studies have shown that patterns at coarse-scale are likely to be controlled by climatic regulators (Guisan & Thuiller 2005) and patterns at fine-scale are controlled by interspecific interactions. Here, I tested the hypothesis that different processes (inherent to the species) are generating different species distribution patterns at particular scales of analysis.

The *Veronica* sect. *Hebe* complex offers an ideal system to be studied due to the distinctive diversity in life-history traits across New Zealand. The group studied here is part of *Veronica* sect. *Hebe* that was formerly classified as *Hebe* and *Leonohebe* (Bayly & Kellow 2006). These groups were referred to as the hebe clade and the semiwhipcord hebe clade by Albach & Meudt (2010). The study is confined to these two clades because of the availability of morphological data contained in Bayly & Kellow (2006) was assembled. Additionally the phylogeny of these species has been studied by Low (2005) using DNA sequence data and reanalysed by Garnock-Jones (pers. comm.).

The observation of different present-day distribution patterns exhibited within *Veronica* sect. *Hebe* (Species taxonomy follows Bayly & Kellow 2006; nomenclature follows Garnock-Jones et al. 2007) at different scales is likely to be due to the geological history of New Zealand (Pole 1994; Lee et al. 2001; Cieraad & Lee 2006), and also influenced by their intrinsic life-history traits. A less explored relationship is that on-going processes like colonisation dynamics, reproduction and dispersal may correlate with statistical properties of the distribution patterns of these taxa that have dispersed to and diversified there. Life-history traits can be used as a correlate to those mechanisms, given that mechanisms such as colonisation, reproduction and competition, are difficult to measure (Cornelissen et al. 2003).

So far, discussion of the present-day distribution of the mountain flora in New Zealand including the *Veronica* complex, has been focused on the evolution, radiation and long-distance dispersal of these groups (McGlone et al. 2001; Winkworth et al. 2005). Here, I carried out a less explored approach investigating if intrinsic plant life-history traits correlate with present-day distribution. For example, it is known that the predominance of self-pollination in many populations of plants is one mechanism that can reproductively isolate populations, thus leading to an aggregated distribution (Raven et al. 1976). In New Zealand, researchers have found high predominance of self-pollination (Newstrom & Robertson 2005). On the other hand dioecy and a lack of specialised pollinators may act to limit species radiations by promoting gene flow between populations and thus leading to a more scattered distribution (Lloyd 1985; Newstrom & Robertson 2005). In this Chapter, I tested these and other life-history traits. I hypothesised some relationships between life-history traits and distribution patterns, as follows:

1. *Seed size* (mass). It is expected that species with larger seeds will tend to have shorter dispersal distances (Thomson et al. in press) and hence, more aggregated distributions leading to lower values of the aggregation index D_m. Also, large seeds spend longer time exposed to pre-dispersal seed predators (Moles & Westoby 2003); and larger seeds could suffer highly and intensive selective predation by animals (Eriksson & Jakobsson 1998; Guo et al. 2000). However, in a more recent study (Moles et al. 2003) no relationship was found between either seed mass and pre-dispersal or post-dispersal survivorship.

   D_m *The mass fractal dimension*

2. *Presence or absence of seed wings*. Species with extra investment in dispersal structure (wind adapted) will tend to have wider geographic range (Edwards & Westoby 1996) than will species without specialised adaptations for dispersal. Hence, species with presence of wings should have more diffuse distributions than those lacking this capacity, leading to higher D_m values.

3. *Breeding system*. Local abundance of a species is linked to its breeding system. Species that are cosexual and self-compatible could form and persist in more fragmented ranges (Young & Brown 1998) leading to lower values of D_m.

4. *Growth habit*. Growth habit is related to longevity (Pocock et al. 2006). Species with increase longevity can persist more time, thus leading to more scattered distributions (Kelly 1996; Kelly & Woodward 1996). Trees are expected to have relatively large and more scattered ranges.

5. *Leaf area and specific leaf area (SLA)*. These traits could be related to herbivory. Species with large leaves have been shown to experience greater degrees of herbivory (Moles & Westoby 2000). SLA is related to physical toughness which is an important form of resistance against herbivores (Hanley et al. 2007; Clissold et al. 2009). Herbivory could lead to specific distribution patterns. I predict that species with small and thicker leaves (small SLA values) will tend to have more aggregated patterns as they may be less susceptible to density-dependent herbivory (Janzen 1970).

6. *Plant height*. It is hypothesised that taller species would lead to higher values of D_m (diffuse distributions), since taller species could increase dispersal capacity into previously unoccupied habitats (Mathews & Bonser 2005; Thomson et al. in press).

7. *Inflorescence length.* The hypothesis was that the species with longer inflorescences would lead to high values of D_m. Similar to the hypothesis of the height of the plants, where it is expected that longer inflorescences might disperse seeds further from the mother plant (Lee et al. 2001), leading to more scattered distributions.

Aggregation patterns of plant species can provide information about the biological processes that shape these distribution patterns. In the present Chapter, life-history traits of the hebe and semiwhipcord hebe clades of *Veronica* sect. *Hebe* were utilised as correlates of the mechanisms of colonisation, reproduction, competition and they were related to the degree of spatial aggregation exhibited by the present-day distribution across multiple scales.

In summary, I asked: a) what plant life-history traits are related with different distribution patterns? and b) which traits are more important at local, regional or national resolution?

The spatial pattern we observe can be the result of the interaction between a number of factors including herbivory, competition and other interactions with neighbouring plants (Dale 2000). In this study, I analysed and compared the results of two approaches: the cross species analysis and the phylogenetically informed method.

4.2 METHODS

Two data sets were analysed. The first data set comprised the national distribution of all 84 species of the *Veronica* complex present in New Zealand, and the second data set comprised the fine-scale distribution of 8 of these *Veronica* species across 27 plots in Molesworth Station, New Zealand (Figure 14). For the *Veronica* species distributed in Molesworth Station history-life traits were obtained both from Bayly & Kellow (2006) and also from field measurements. These traits were plant height, seed mass, inflorescence length, leaf area, SLA, presence or absence of seed wings, growth habit and breeding system.

4.2.1 *Distributional data and study area*

In order to quantify aggregation at different scales, two data sets were analysed. To quantify aggregation at local-scale, the distribution of eight *Veronica* species was obtained from surveys in Molesworth Station, New Zealand (Figure 29). Molesworth station is situated in the South Island, behind the Inland Kaikoura Mountains in South Marlborough. The average annual rainfall ranges from less than 760 mm in the East to more than 2,540 mm in the West (Coulter 1969). It is common to find gravelfield and scree communities in dry places and in places where the rainfall is heavier, exotic grassland species and short native tussocks are common. The range of elevation for the quadrats was from 700 to 1460 m. Twenty-seven 100×60 m quadrats were established. Georeferenced position for the eight *Veronica* species from Molesworth was recorded in the field using a differential Global Positioning System (GPS) Trimble GeoXT which can yield precise measurements (30 cm error in post processing in combination with a base station for differential correction).

To quantify aggregation at regional and national scales, distribution records for 84 species of hebes and semiwhipcord hebes (*Veronica* sect. *Hebe*) distributed across New Zealand were obtained from Bayly & Kellow (2006). They were recorded as presence/absence at a ~2.5 km resolution (Figure 14).

4.2.2 *Aggregation index: the mass fractal dimension D_m*

Popular aggregation indices like Ripley's *K* statistic (Haase 1995) or Ω (Condit et al. 2000) have been used widely in previous studies. However, fractal dimension *D* was chosen in the present chapter since it has been proven that many distribution patterns have (as fractals do) relevant features on a variety of different scales (Halley et al. 2004). The calculations of mass fractal dimension (D_m) and Ω are closely related and follow many of the same steps in their calculation. Therefore, the results of analyses based on D_m can be comparable to the results of other studies based on Ω index. All the steps involved in the calculation of D_m and Ω are given in Appendix E.2.

The fractal dimension can be derived from the slope (*b*) of an occupancy-area curve relationship, and can be thought of as a measure of sparseness of any set embedded in an Euclidean space. Distributions that are highly fragmented and clustered tend to have $D \approx 0$;
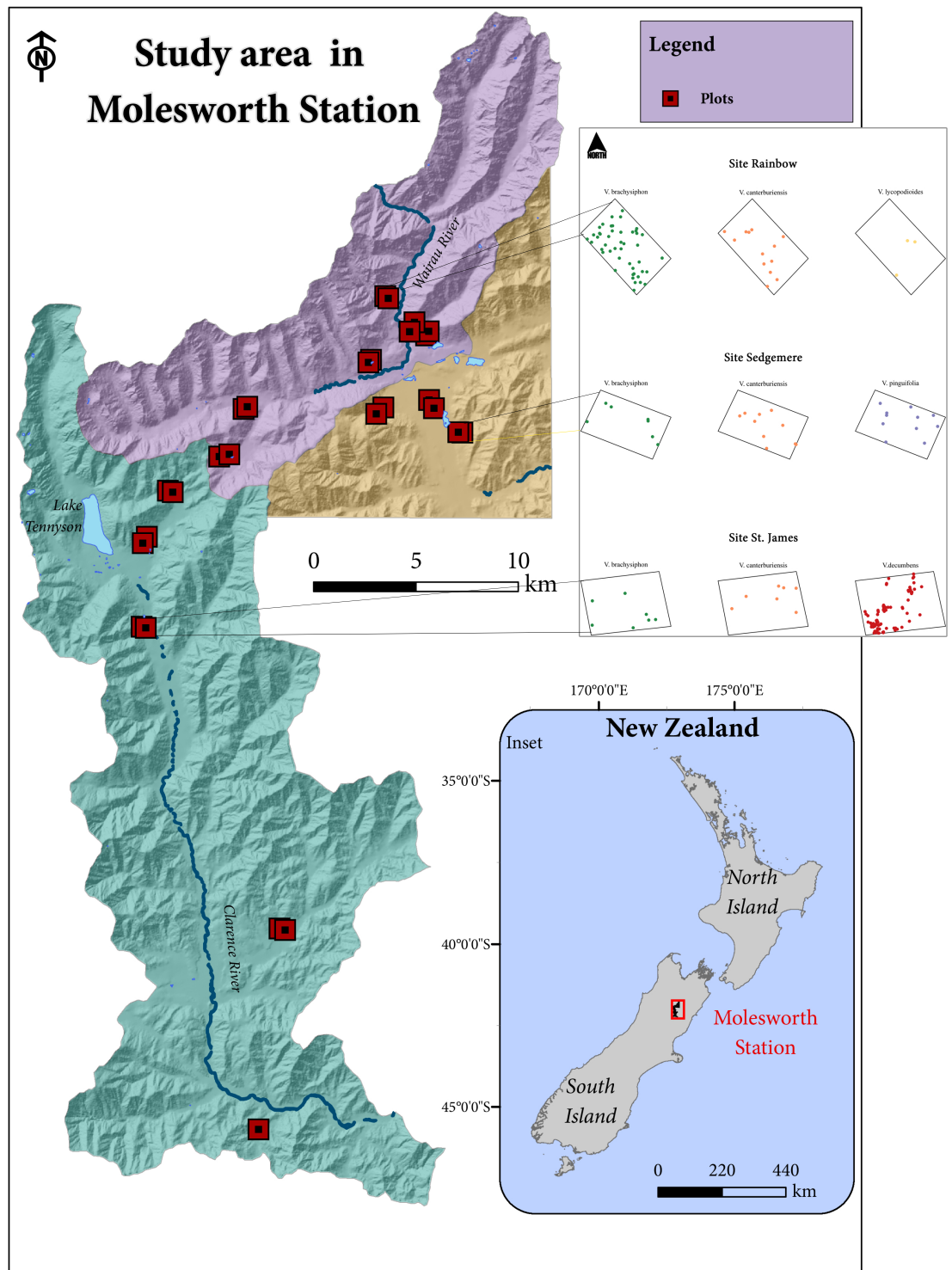
Figure 29: Study area and plot location in Molesworth Station. Examples of point patterns of *Veronicas* recorded with the differential GPS in three quadrats are shown in the right box: Green points depict *Veronica brachysiphon* individuals; orange, *V. canterburiensis*; purple, *V. pinguifolia*; red, *V. decumbens* and yellow, *V. lycopodioides*.
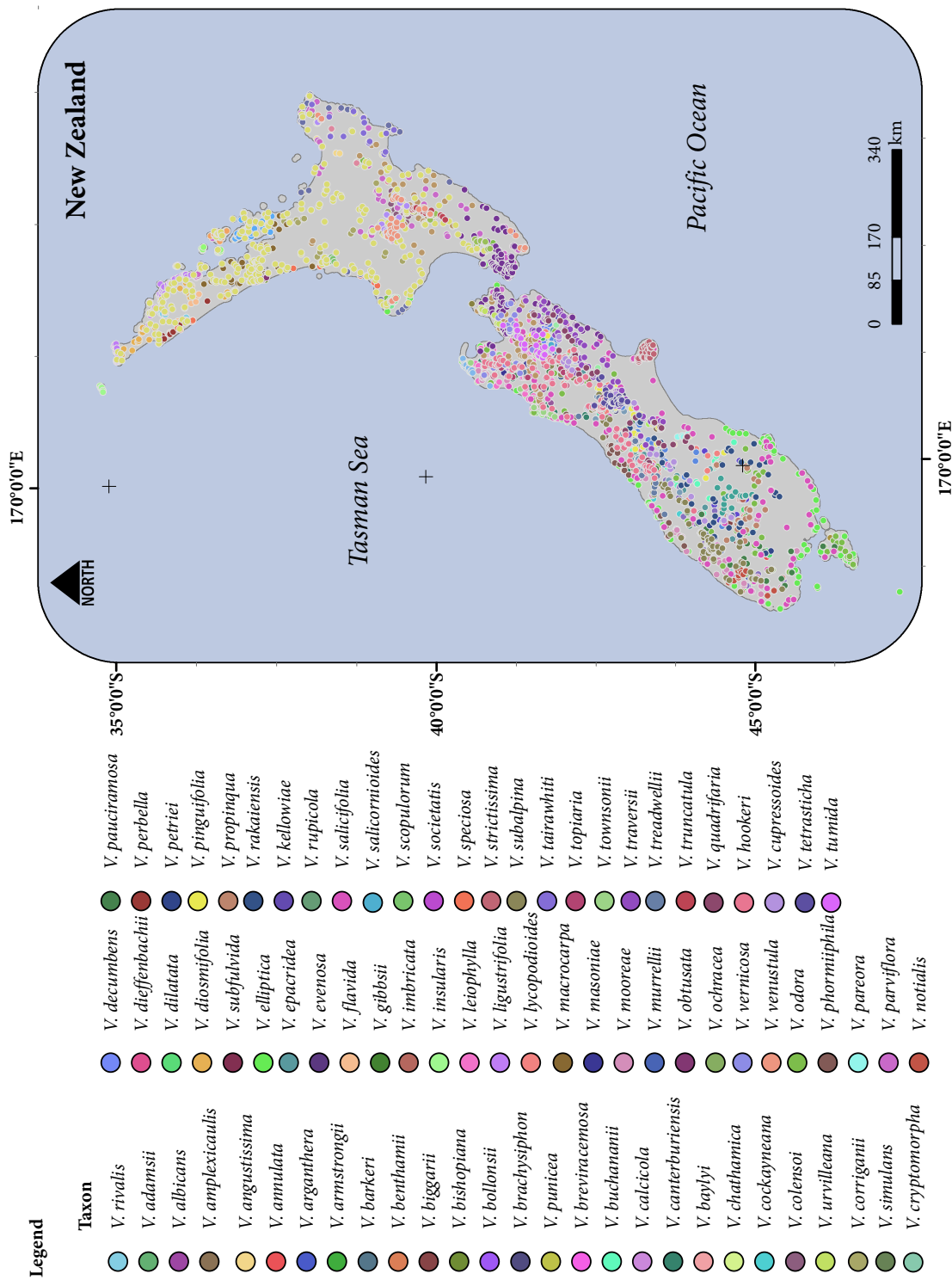
Figure 30: Distribution map of hebes and semiwhipcord hebes (*Veronica* sect. *Hebe*). This map was compiled by Benjamin Magana-Rodríguez on May 2011. Data sources: New Zealand Territorial Authority data, distribution data of *Veronica* spp. provided by (Bayly & Kellow 2006, as *Hebe* and *Leonohebe*). Projection and Datum: NZMG 1949. The distribution of one species is shown in Figure 32.

distributions with continuous coverage patterns tend to have $D \approx 2$ (Hartley & Kunin 2003; Figure 31).
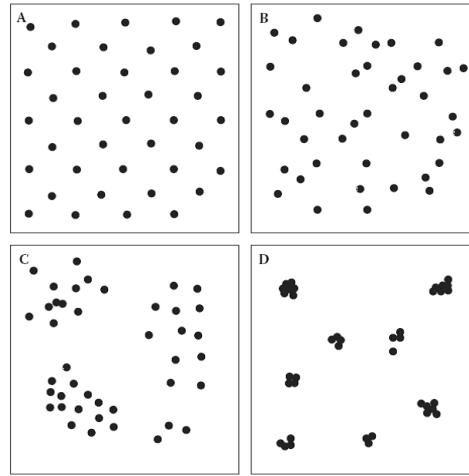
Figure 31: Fractal dimensions of different point patterns. A) Regular point pattern, $D = 2$, B) random point pattern, $D = 1.8$ C) random clumped point pattern, $D = 1.4$ and D) aggregated clumped point pattern, $D = 0.9$. Figure taken from Seuront (2009).

There are many ways to measure the fractal dimension depending on the nature of the data (Halley et al. 2004). The mass fractal dimension (D_m) method was used in this study since it has been developed to analyse point pattern data. This method averages the number of points located within a distance of circles of radius r of any random point in the point pattern. The method also makes an adjustment for the amount of area available within a radius r, hence areas of sea beyond the land boundary, or areas that were unsurveyed beyond the plot boundary are excluded when calculating the expected density of points. Practically, the mass fractal dimension can be estimated using squares or circles of increasing size r starting from the centre of the domain (Seuront 2009; Figure 32). Then, log-log axes of the slope are calculated to obtain the mass fractal dimension. Two ranges of resolution were used to quantify the distribution pattern, one at the regional resolution of 2.5 to 20 km ($D_{2.5-20km}$) and another at the national resolution of 20 to 100 km ($D_{20-100km}$). The distance classes used ranged from 2.5 to 100 km. An example is shown in Figure 33.

For the eight species surveyed in Molesworth Station distance classes ranged from 0.0025 km to 1 km. Two resolutions were used as well, one at fine resolution of 2.5 to 60 m ($D_{2.5-60m}$) and another at coarse resolution of 60 to 1000 m ($D_{60-1000m}$).
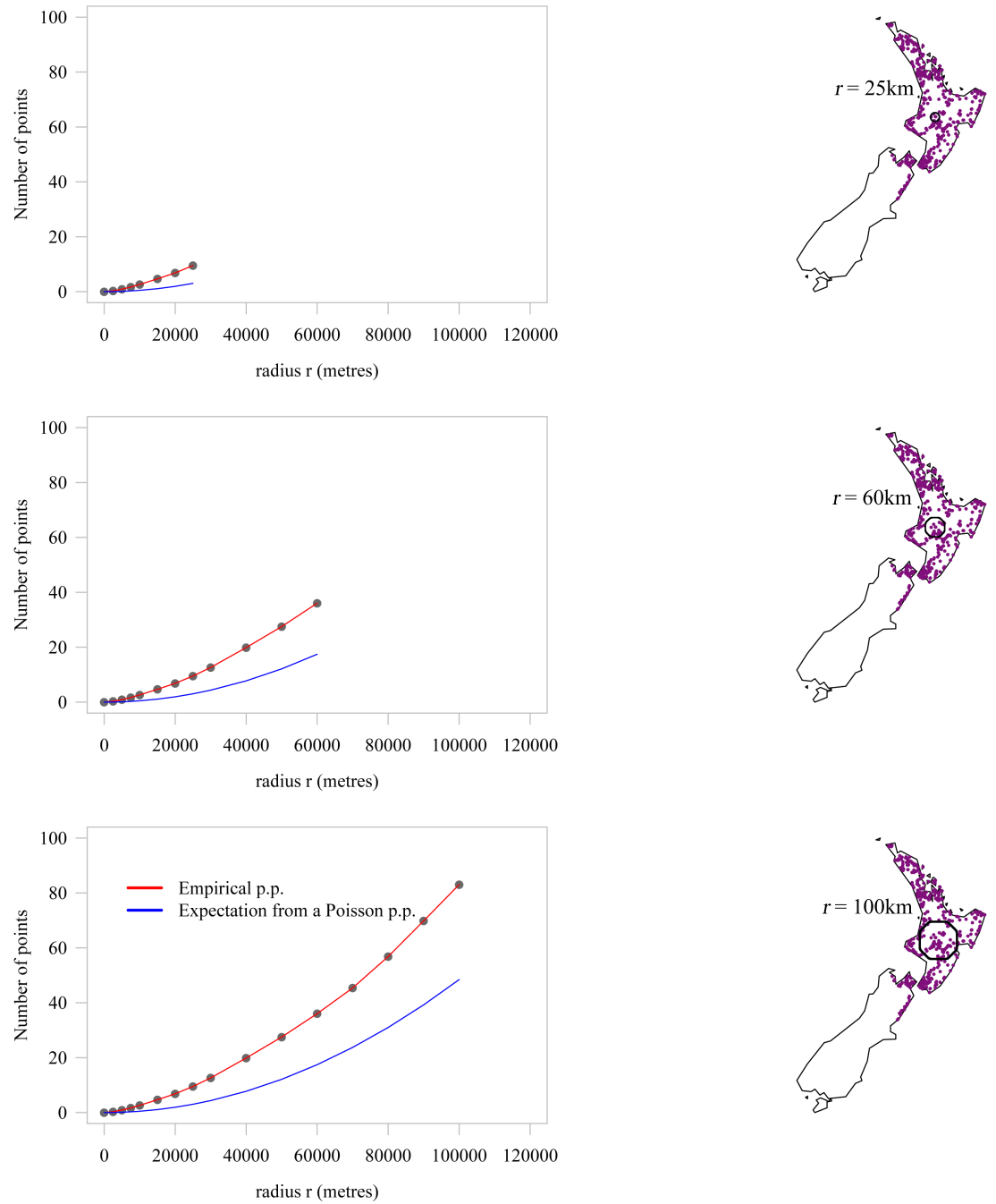
Figure 32: Schematic representation of the calculation of the mass fractal dimension D_m. The graphs indicate the average number of points located within a radius r of each point. The map represents the actual distribution of *Veronica stricta* in New Zealand. As r gets bigger, there are more points accumulated. Further explanation of the calculation of mass fractal dimension is provided in Figure 33.
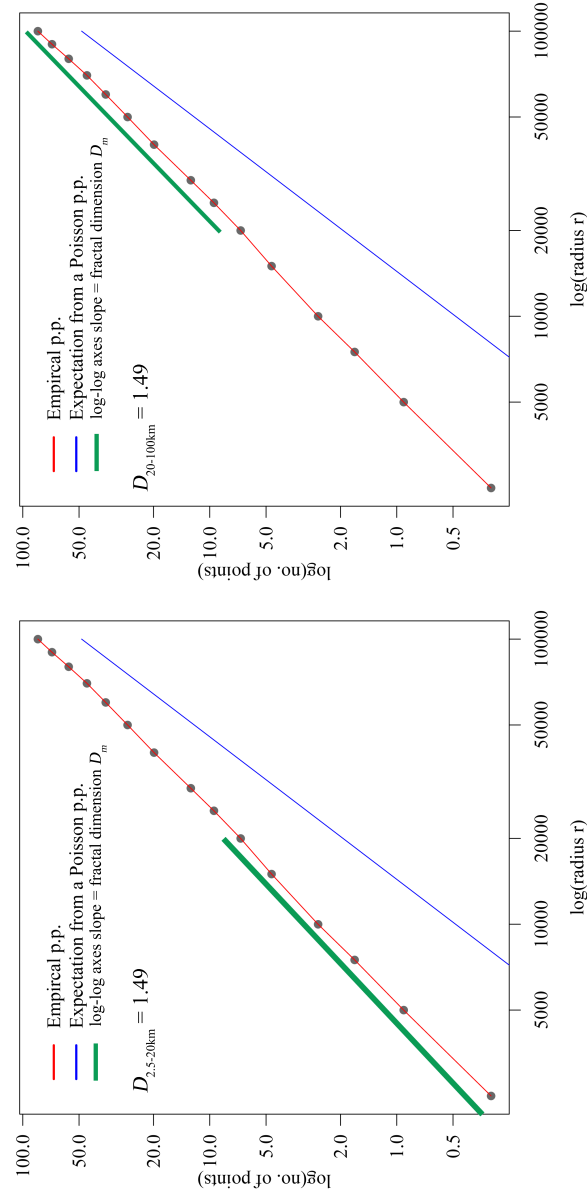
Figure 33: Calculation of mass fractal dimension D_m for *Veronica stricta*. Same graph as Figure 32 (bottom) but in log-log axes. Green lines are the slope of the empirical point pattern. The left graph is the calculation at regional resolution ($D_{2.5-20km} = 1.49$) and the right graph is at national resolution ($D_{20-100km} = 1.49$). This *Veronica* species has a more or less fractal (statistically self-similar) distribution pattern at both scales. The blue line is the expected cumulative density from a Poisson point process CSR and it always has a slope of 2, ($D_m = 2$).

4.2.3 *Plant life-history traits*

Life-history attributes of the study species for regional and national scales were obtained from Bayly & Kellow (2006) and included in the DELTA (Dallwitz 2000) data set. Besides, additional field measurements were taken from other life-history traits at a local-scale (*Veronica* species surveyed in Molesworth Station, see methods in Chapter 3).

4.2.4 *Data analysis*

Cross-species relationships between life-history traits and distribution patterns (defined as mass fractal dimension D_m) were analysed using simple regression analysis for univariate models. ANOVA were performed to compare different groups. Data were logarithmic (ln) transformed for normality when necessary.

Since some traits are dependent, the interaction between them could affect distribution patterns (Pocock et al. 2006). Multiple regression analyses were performed testing for these interactions. The full model (all predictor variables) was considered first, and then the MAM was fitted following the *principle of parsimony* (Sober 1991). The most parsimonious model was assessed to be the one with the lowest AIC (Burnham & Anderson 2002). Multivariate analyses were performed only in the *Veronica* species at the regional and national-scale since the dataset at local-scale contained only eight species, which gave insufficient degrees of freedom for these tests. All statistical analyses were conducted in R v.2.10 (R Development Core Team 2011) in association with the MASS package (Venables & Ripley 2002).

MAM *Minimum adequate model*

AIC *Akaike's Information Criterion*

4.2.5 *Phylogenetic dependence*

When correlating traits among related species, species cannot be considered as independent data points due to shared inheritance of traits (Felsenstein 1985). PIC analyses were performed to overcome this problem. I used the phylogeny of *Veronica* sect. *Hebe* given in Figure 34. This phylogeny was derived by a Bayesian phylogenetic analysis by Hamish Carson (P.J. Garnock-Jones pers. comm.) of combined nuclear and cpDNA sequences obtained by Evonne Low (Low 2005). The topology was obtained in Nexus format, and subsequent transformation was made in newick format to handle the file in R v.2.10 (R Development Core Team 2011) and PHYLOCOM. The root node was

fixed at 3.9 million years before present (mybp) which is the age that has been estimated for the hebe clade of *Veronica* (Wagstaff et al. 2002, as *Hebe*); then the remainig branch lengths were estimated using the *bladj* algorithm from the Phylocom software (Webb et al. 2008). Previous studies (Webb 2000) have shown that even with only a single dated node, the resulting phylogenetic distances calculated by the *bladj* algorithm can be a marked improvement upon the assumption of equal branch lengths. The resultant phylogenetic tree was analysed using PIC with the combination of adephylo (Jombart et al.), ape (Paradis et al. 2004), phylobase (Hackathon et al. 2010) and picante (Kembel et al.) packages in R v.2.10 (R Development Core Team 2011). Regressions were passed through the origin as recommended by Legendre & Desdevises (2009).

PIC *phylogenetic independent contrast*
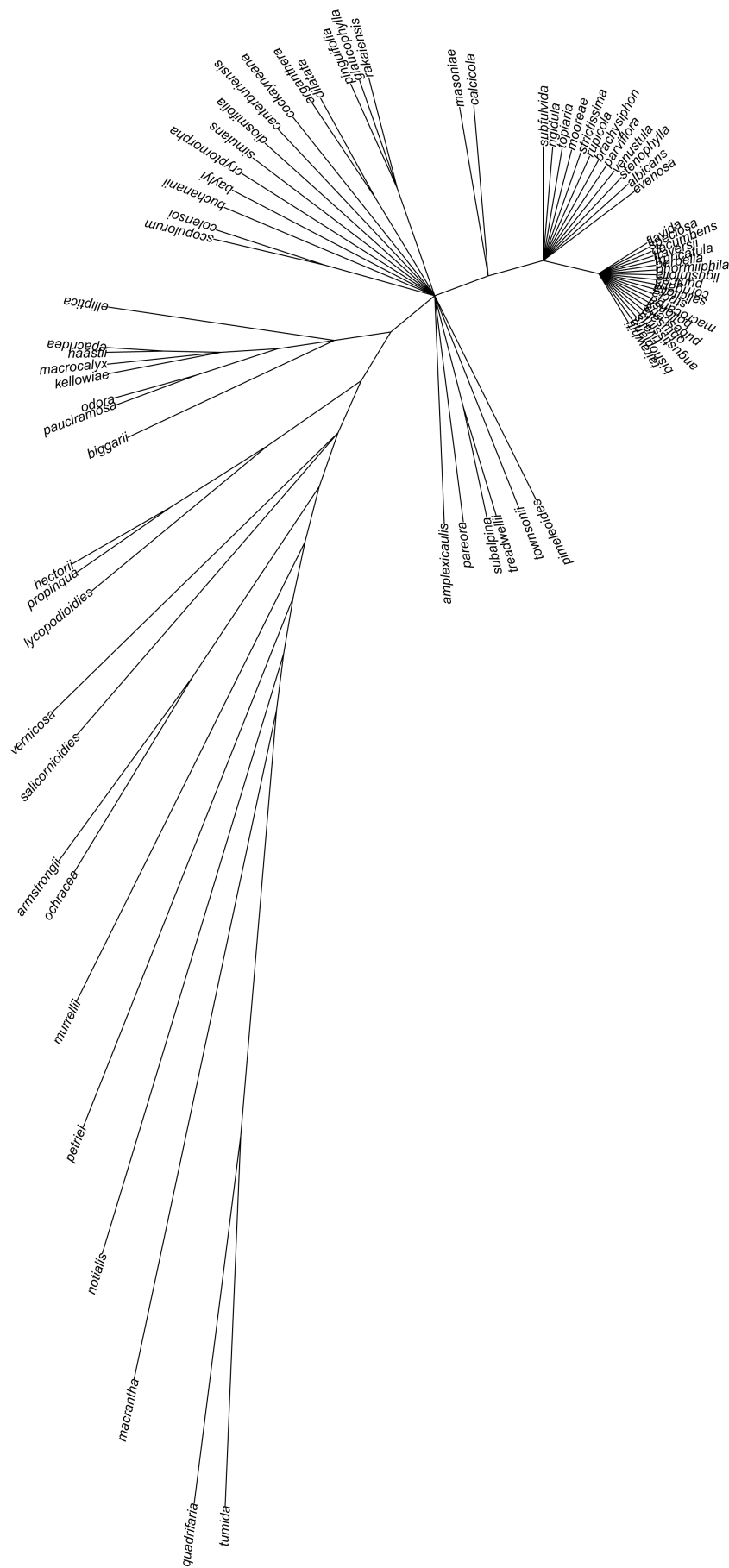
Figure 34: Modified phylogenetic tree for the hebe and semiwhipcord hebe clades of *Veronica* sect. *Hebe*.

4.3.1   *Distribution patterns*

Distribution patterns, defined as the mass fractal dimension, ranged from values close to 0 meaning one or more independent 'points' to the distributions, to values close to 2 meaning "space-filling" and continuous distributions (with no gaps greater than the minimum distance being considered). This range of values was present at both scales, with some variation among the species depending on the resolution (Figure 35).

For the *Veronica* spp. surveyed at a local-scale in Molesworth Station the values of D_m varied from low (*Veronica pimeleoides*) to high (*Veronica decumbens*) in the 2.5 to 60 m fine local-scale, and from low (*Veronica lycopodioides*) to high (*Veronica pinguifolia*) in the 60 to 100 m coarse local-scale.

4.3.2   *Correlations between variables*

Out of 21 possible pairwise correlations between the trait variables, there were significant correlations between six pairs, after applying the Bonferroni correction ($\alpha < 0.002$; Figure 36). There were significant relationships between height and seed size, inflorescence length and height, inflorescence length and leaf area, leaf area and height and between growth habit and plant height.

4.3.3   *Relationship between plant traits and aggregation*

Cross-species univariate analyses

The relationship between *Veronica* species life-history traits and aggregation patterns varied among the scale of analysis (local, regional and national-scale).

At the fine resolution ($2.5 - 60$ m) of the local-scale, leaf area and SLA were related to aggregation. Leaf area was weakly positively related to aggregation ($R^2 = 0.45$, $P = 0.06$; Figure 37c) and SLA was negatively related to D_m ($R^2 = 0.46$, $P = 0.05$; Figure 37d). That is, species with smaller and thinner leaves were associated with point-like distributions (Figure 37; Table 16). At the coarse resolution ($60 - 1000$ m) of the local-scale, the presence of wings on the seed was associated with less aggregation (high D_m; $P = 0.02$; Figure 38f). Weakly-winged

Figure 35: Mass fractal dimension D_m for the *Veronica* sect. *Hebe* in New Zealand at regional and national scale.
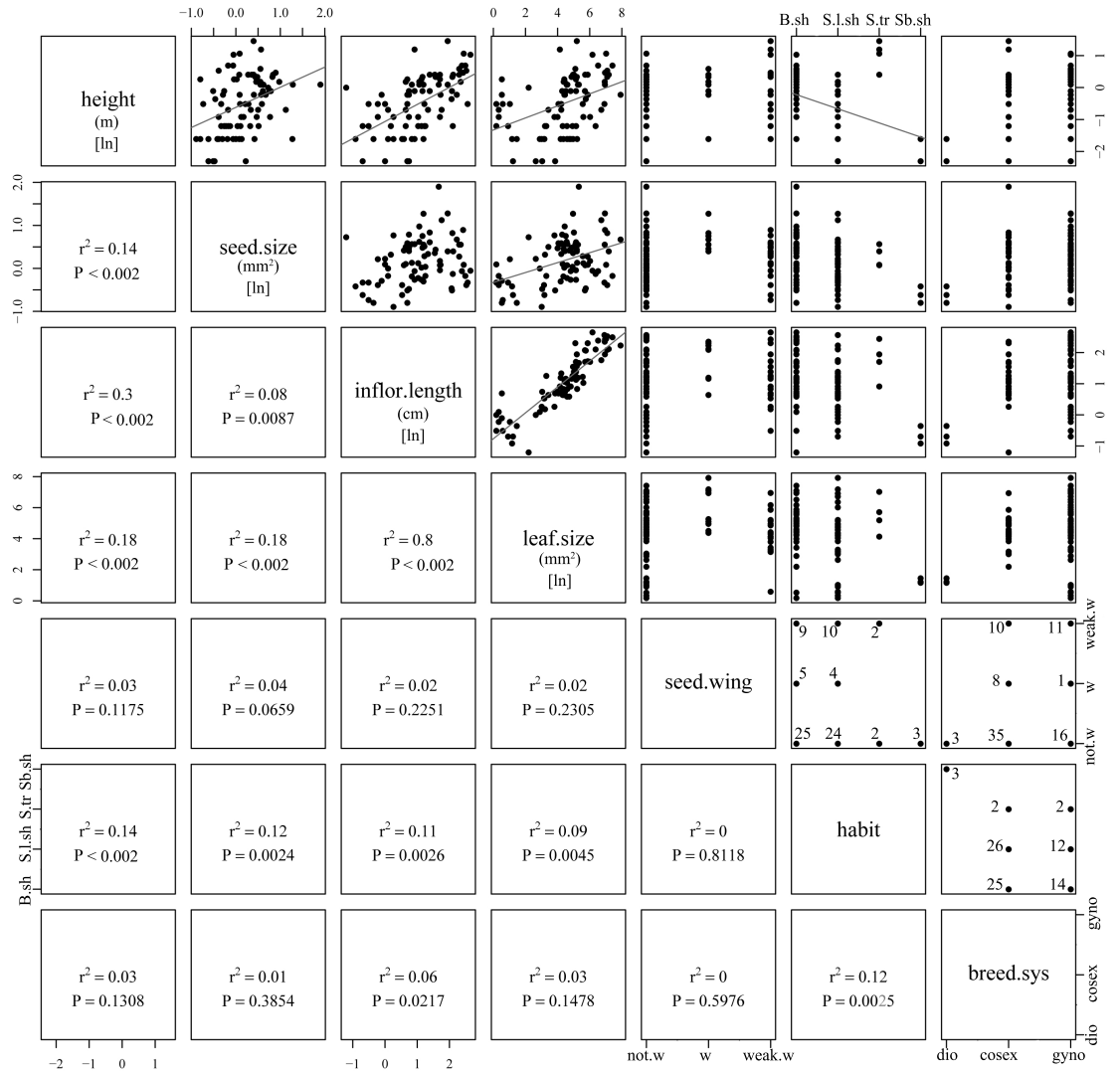
Figure 36: Correlations between the 7 variables tested for *Veronica* species at regional and national scale. Life-history traits are presented on the diagonal, scatter plots with the trend regression line are above the diagonal, and a statistics summary of the correlation are below the diagonal. All traits except the categorical were ln-transformed prior to analysis. The categorical variables labels are as follow: seed wings, not.w = not winged; weak.w = weakly winged; and w = winged. Growth habit, Sb.sh = subshrub; S.l.sh = spreading low shrub; B.sh = bushy shrub; S.tr = small tree. Breeding system, cosex = cosexual; gyno = gynodioecious; dio = dioecious. The number next to each point indicates the number of species with each combination of traits.

species had more space-filling distributions. Plant height, inflorescence length, leaf area, SLA, seed mass, growth habit and breeding system were not associated with aggregation at local coarse-scale (Figure 38a, b, c, d, e, g, h; Table 16).

At a regional-scale $(2.5 - 20$ km), of the seven plant traits tested, inflorescence length, leaf area and breeding system were significantly related to aggregation patterns (Table 17). Inflorescence length $(R^2 = 0.10, P < 0.01$; Figure 39c) and leaf area $(R^2 = 0.07, P = 0.012$; Figure 39d) showed negative correlations with mass fractal dimension. *Veronica* species comprising plants with bigger leaves and longer inflorescences tended to have more point-like distribution patterns. Breeding system showed a positive association with mass fractal dimension $(P = 0.04$, Figure 39g). Dioecious species had more continuous distributions, reflected by higher D_m values. Plant height, seed size, presence/absence of wings of seed and growth habit were not associated with aggregation at this particular scale (Figure 39a, b, e, f; Table 17).

At a national-scale $(20 - 100$ km) none of the plant traits were significant explanatory variables in predicting distribution patterns (Figure 40a, b, c, d, e, f, g; Table 17).

Cross-species multivariate analyses

At regional and national scales, multivariate analyses with all the predictor plant traits in the model (full model) did not suggest single significant life-history traits.

After simplifying the model following the MAM method, the best model at regional-scale, had three plant-history traits as important $(P = 0.03, R^2 = 0.10)$. The MAM suggested that leaf area, seed size and breeding system contributed significantly to the multiple regression. Leaf area was significant $(P = 0.05)$ when taking into account seed size and breeding system (Table 18). The MAM maintained the same direction of the relationship as the univariate analyses. For the leaf area the slope (b) was -0.06 and for breeding system it was -0.14.

MAM *minimum adequate model*

At a national-scale, the MAM suggested a very simple model with just an intercept (Table 18).

Can coarse-scale be predicted from fine-scale?

The regional and national-scale mass fractal dimension relationship was plotted using regression model II (Legendre & Legendre 1998),
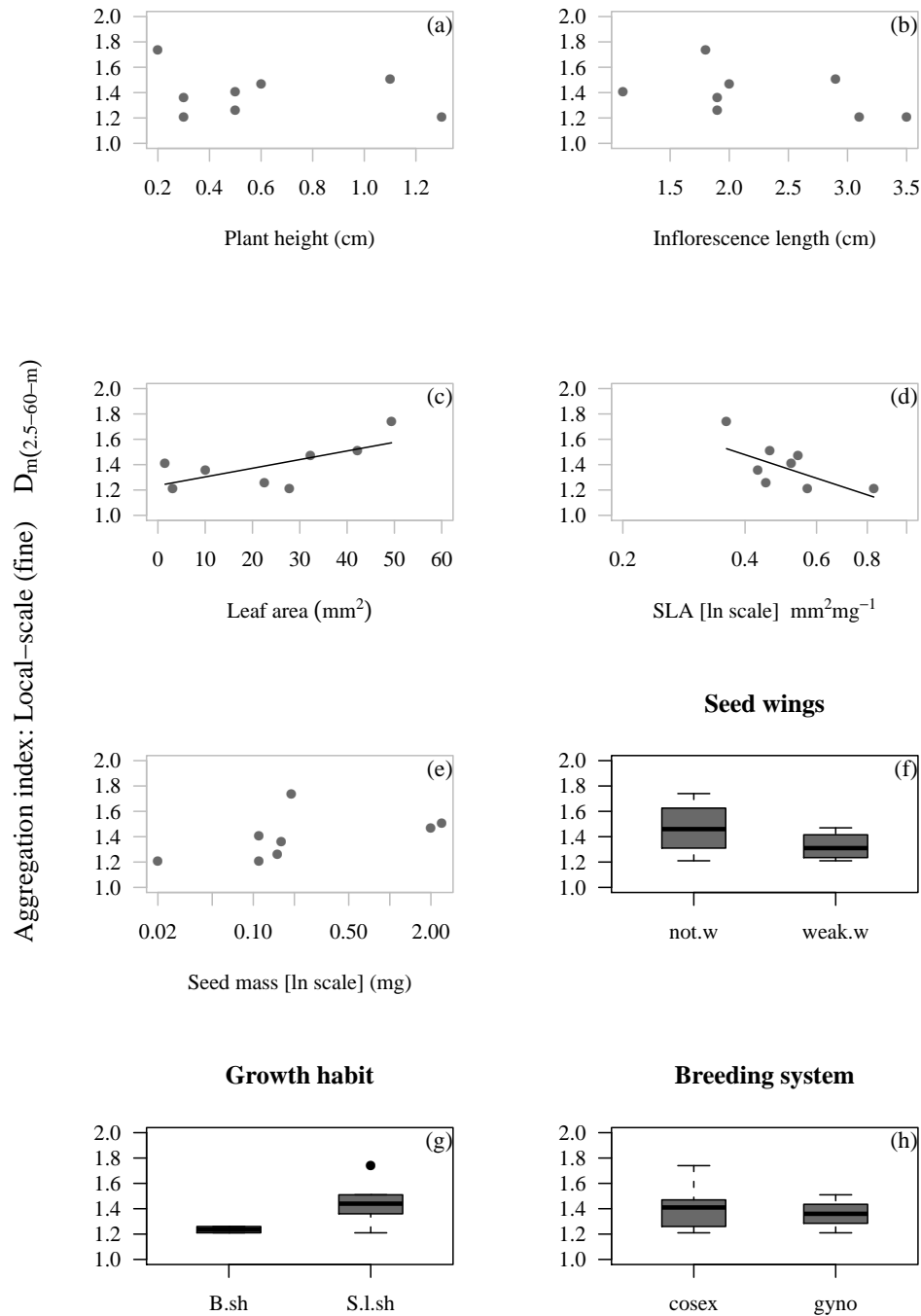
Figure 37: Scatter and boxplots of variables used as predictors of aggregation patterns (mass fractal dimension) at local-scale (fine resolution, 2.5-60 m). In the boxplots the categorical variables were: seed wings, not.w = not winged; weak.w = weakly winged. Growth habit, S.l.sh = spreading low shrub; B.sh = bushy shrub. Breeding system, cosex = cosexual; gyno = gynodioecious. Box = interquartile range, containing 50% of values; line across box = median; whiskers = highest and lowest values, outliers = cases with values > 1.5 box lengths from the upper or lower edge of the box. Regression lines only fitted when significant.
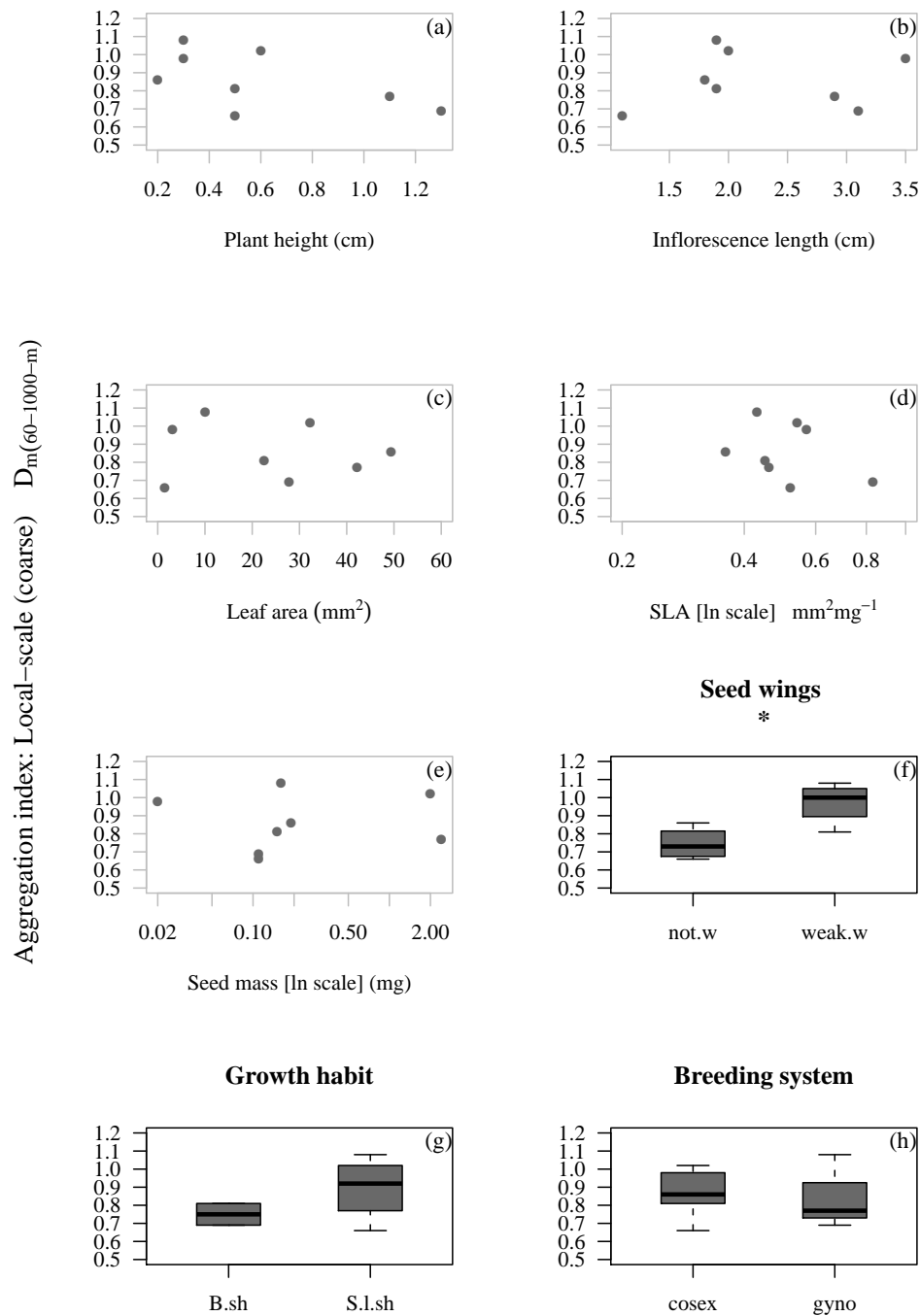
Figure 38: Scatter and boxplots of variables used as predictors of aggregation patterns (mass fractal dimension) at local-scale (coarse resolution, 60-1000 m). In the boxplots the categorical variables were: seed wings, not.w = not winged; weak.w = weakly winged. Growth habit, S.l.sh = spreading low shrub; B.sh = bushy shrub. Breeding system, cosex = cosexual; gyno = gynodioecious. Box = interquartile range, containing 50% of values; line across box = median; whiskers = highest and lowest values, outliers = cases with values > 1.5 box lengths from the upper or lower edge of the box. Significant results are indicated as * P < 0.05.
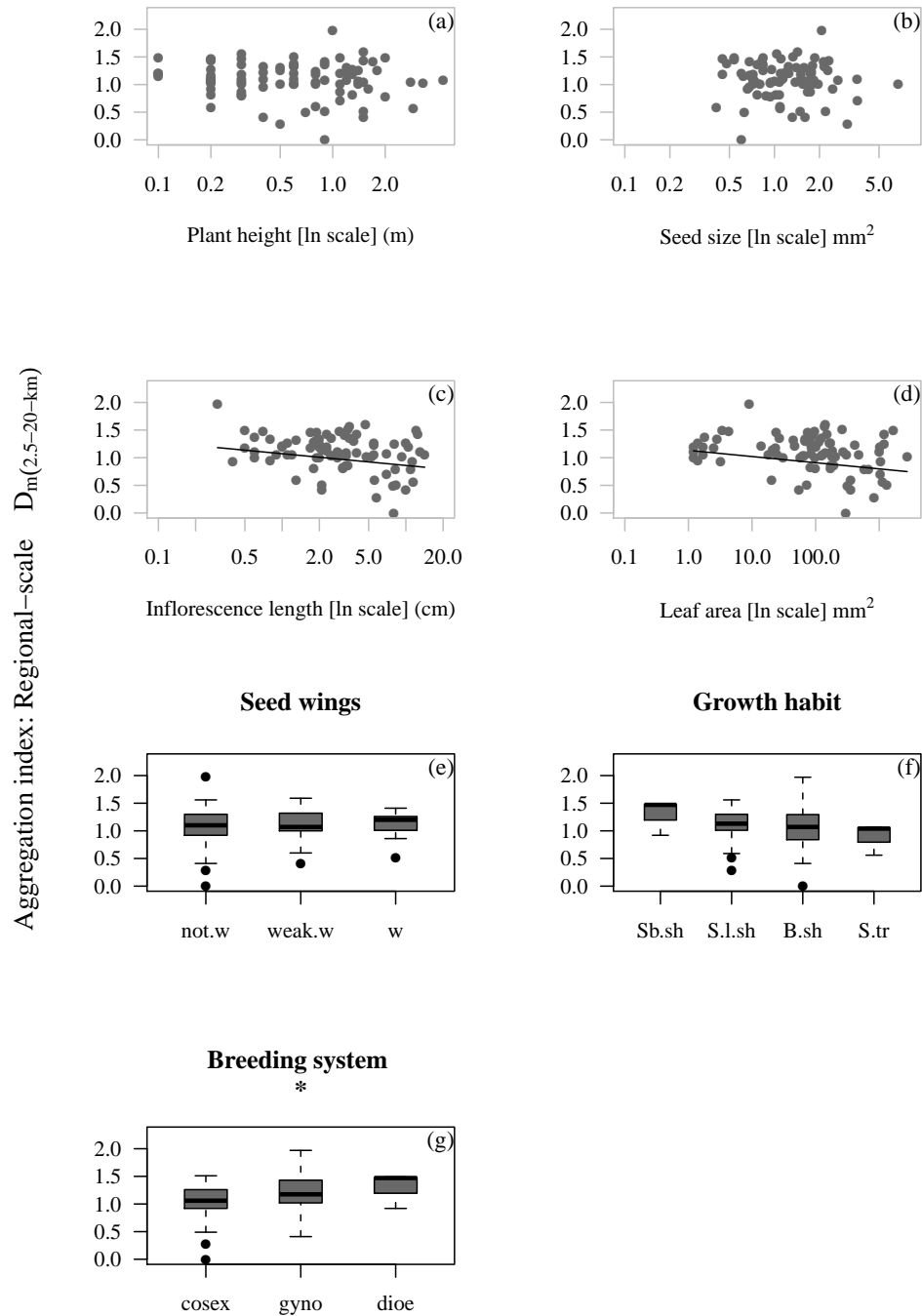
Figure 39: Scatter and boxplots of variables used as predictors of aggregation patterns (mass fractal dimension) at regional-scale 2.5-20 km. In the boxplots the categorical variables were: seed wings, not.w = not winged; weak.w = weakly winged; and w = winged. Growth habit, Sb.sh = subshrub; S.l.sh = spreading low shrub; B.sh = bushy shrub; S.tr = small tree. Breeding system, cosex = cosexual; gyno = gynodioecious; dio = dioecious. Box = interquartile range, containing 50% of values; line across box = median; whiskers = highest and lowest values, outliers = cases with values > 1.5 box lengths from the upper or lower edge of the box. Significant results are indicated as * P < 0.05. Regression lines only fitted when significant.
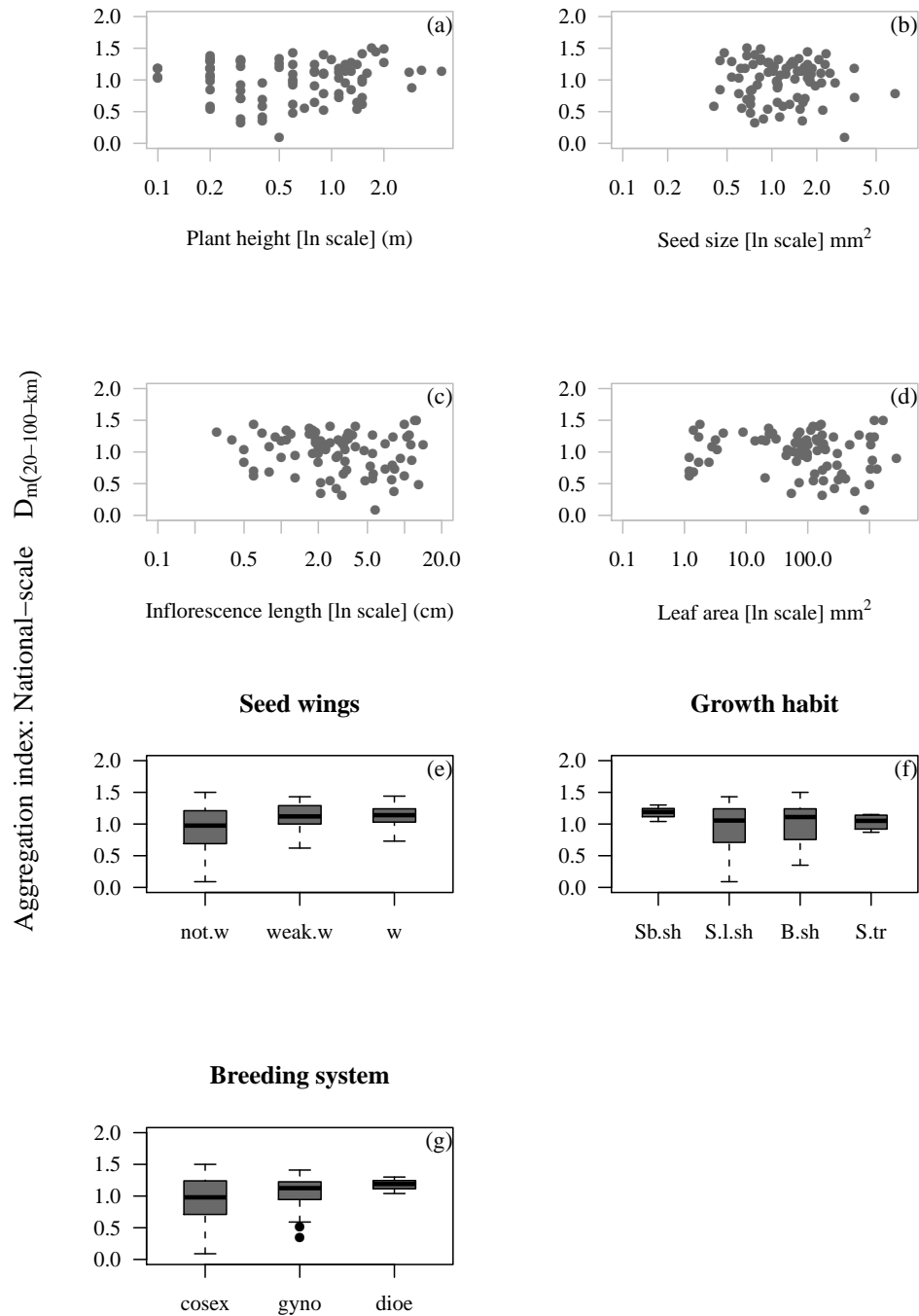
Figure 40: Scatter and boxplots of variables used as predictors of aggregation patterns (mass fractal dimension) at national-scale 20-100 km. In the boxplots the categorical variables were: seed wings, not.w = not winged; weak.w = weakly winged; and w = winged. Growth habit, Sb.sh = subshrub; S.l.sh = spreading low shrub; B.sh = bushy shrub; S.tr = small tree. Breeding system, cosex = cosexual; gyno = gynodioecious; dio = dioecious. Box = interquartile range, containing 50% of values; line across box = median; whiskers = highest and lowest values, outliers = cases with values > 1.5 box lengths from the upper or lower edge of the box.

Table 16: Regression coefficients for univariate traits for cross-species analyses at local-scale (fine and coarse resolutions), b is the direction of the slope. Trait type codes: C. Continuous; N. Nominal.

TRAIT	TYPE	ATTRIBUTES	LOCAL FINE RESOLUTION			LOCAL COARSE RESOLUTION		
			b	R^2	P	b	R^2	P
Plant height	C	Expressed in cm	-0.12	0.07	0.52	-0.21	0.31	0.14
Inflorescence length	C	Expressed in cm	-0.09	0.19	0.27	0.02	0.01	0.83
Leaf area	C	Expressed in mm^2	0.01	0.45	**0.06**	-0.01	0.01	0.80
SLA	C	Expressed in mm^2 mg^{-1}	-0.50	0.46	**0.05**	-0.20	0.10	0.43
Seed mass	C	Expressed in mg	0.05	0.26	0.19	0.03	0.01	0.93
Winging	N	2 attributes: not winged, weakly winged			0.29			**0.02**
Habit	N	2 attributes: bushy shrub, spreading low shrub			0.15			0.28
Breeding system	N	2 attributes: cosexual, gynodioecious			0.69			0.87

Table 17: Regression coefficients for univariate traits for cross-species analyses at regional-scale and national-scale. b is the direction of the slope. Trait type codes: C. Continuous; N. Nominal.

TRAIT	TYPE	ATTRIBUTES	REGIONAL-SCALE			NATIONAL-SCALE		
			b	R^2	P	b	R^2	P
Plant height	C	ln transformed	-0.03	0.10	0.34	0.02	0.01	0.44
Seed size	C	ln transformed	-0.02	0.01	0.71	-0.02	0.01	0.64
Inflorescence length	C	ln transformed	-0.11	0.10	**0.002**	-0.04	0.01	0.23
Leaf size	C	ln transformed	-0.04	0.07	**0.012**	-0.02	0.02	0.18
Winging	N	3 attributes: not winged, weakly winged and winged			0.93			0.12
Habit	N	4 attributes: bushy shrub, subshrub, spreading low shrub and small tree			0.44			0.72
Breeding system	N	3 attributes: cosexual, gynodioecious and dioecious			**0.04**			0.31

Table 18: Regression coefficients for multivariate analyses at regional-scale and national-scale. ΔAIC is the difference from the AIC full model and the MAM.

DEPENDENT.VAR	PREDICTOR.VAR	ΔAIC	p
Regional scale D_m (2.5 – 20 km) Full Model	Leaf size, seed size, height, inflorescence length, winging, habit, breeding system		NS
MAM	Leaf size, seed size, breeding system	9.60	Leaf area = 0.05
			Seed size = 0.11
			Breeding system = 0.08
National scale D_m (20 – 100 km) Full Model	Leaf size, seed size, height, inflorescence length, winging, habit, breeding system		NS
MAM	~ 1	6.81	NS

in order to detect if aggregation at regional-local scale can predict national-scale *Veronica* species or *vice versa*.

Regional-scale aggregation $D_{2.5\text{-}20km}$ predicted national-scale aggregation $D_{20\text{-}100km}$ ($R^2 = 0.38$, df $= 82$, $P < 0.001$, Figure 41a).

At a local-scale, aggregation at fine resolution did not predict coarse resolution aggregation ($R^2 = 0.001$, df $= 6$, $P = 0.46$; Figure 41b).
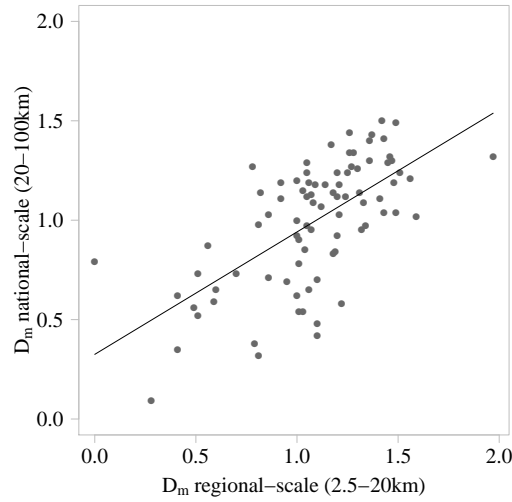
(a) *Veronica* spp. National Distribution (regional and national scale).
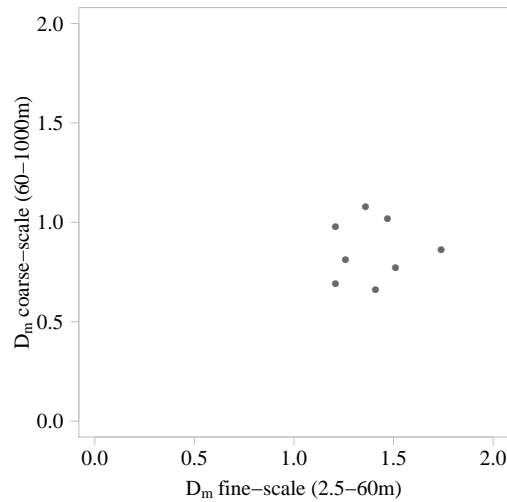
(b) *Veronica* spp. in Molesworth Station (local scale).

Figure 41: Relationship between coarse and fine-scale aggregation (regression model II). a) Regional-scale versus national-scale ($P < 0.001$, $R^2 = 0.38$). b) Local data, fine-scale versus coarse-scale ($P = 0.46$, $R^2 = 0.001$).

Table 19: Regression coefficients for univariate traits for PIC analyses at regional-scale and national-scale. b is the direction of the slope. Trait type codes: C, Continuous.

TRAIT	TYPE	ATTRIBUTES	REGIONAL-SCALE			NATIONAL-SCALE		
			b	R^2	P	b	R^2	P
Height	C	ln trans-formed	-0.005	0.001	0.92	0.130	0.074	**0.017**
Seed size	C	ln trans-formed	0.030	0.001	0.73	0.001	0.001	0.81
Inflorescence length	C	ln trans-formed	-0.013	0.050	**0.04**	-0.043	0.007	0.47
Leaf area	C	ln trans-formed	-0.090	0.050	**0.04**	-0.056	0.021	0.28

Phylogenetic independent contrast analyses

The phylogenetic tree of the study species (P.J. Garnock-Jones pers. comm.) is presented as an unrooted network (Figure 34). Of the seven life-history traits tested in this Chapter, only four were considered in phylogenetic analyses since the adephylo package can only handle continuous and binary data at present.

The PIC analyses at regional-scale identified negative associations between life-history traits and aggregation patterns (D_m) (Figure 43, Table 19). Inflorescence length ($R^2 = 0.05$, $P = 0.04$; Figure 43c) and leaf area ($R^2 = 0.05$, $P = 0.04$; Figure 43d) were negatively associated with aggregation (D_m) at regional-scales. These were consistent with the cross-species analyses.

At a national-scale, aggregation pattern was positively related to height ($P = 0.017$, $R^2 = 0.017$; Figure 44a). Taller *Veronica* species had more space-filling or random point pattern distribution. Seed size, inflorescence length and leaf area life-history traits were not related to aggregation patterns (Figure 44b, c, d; Table 19). These results were consistent with the cross-species analyses at national-scale except for the plant height life-history trait.
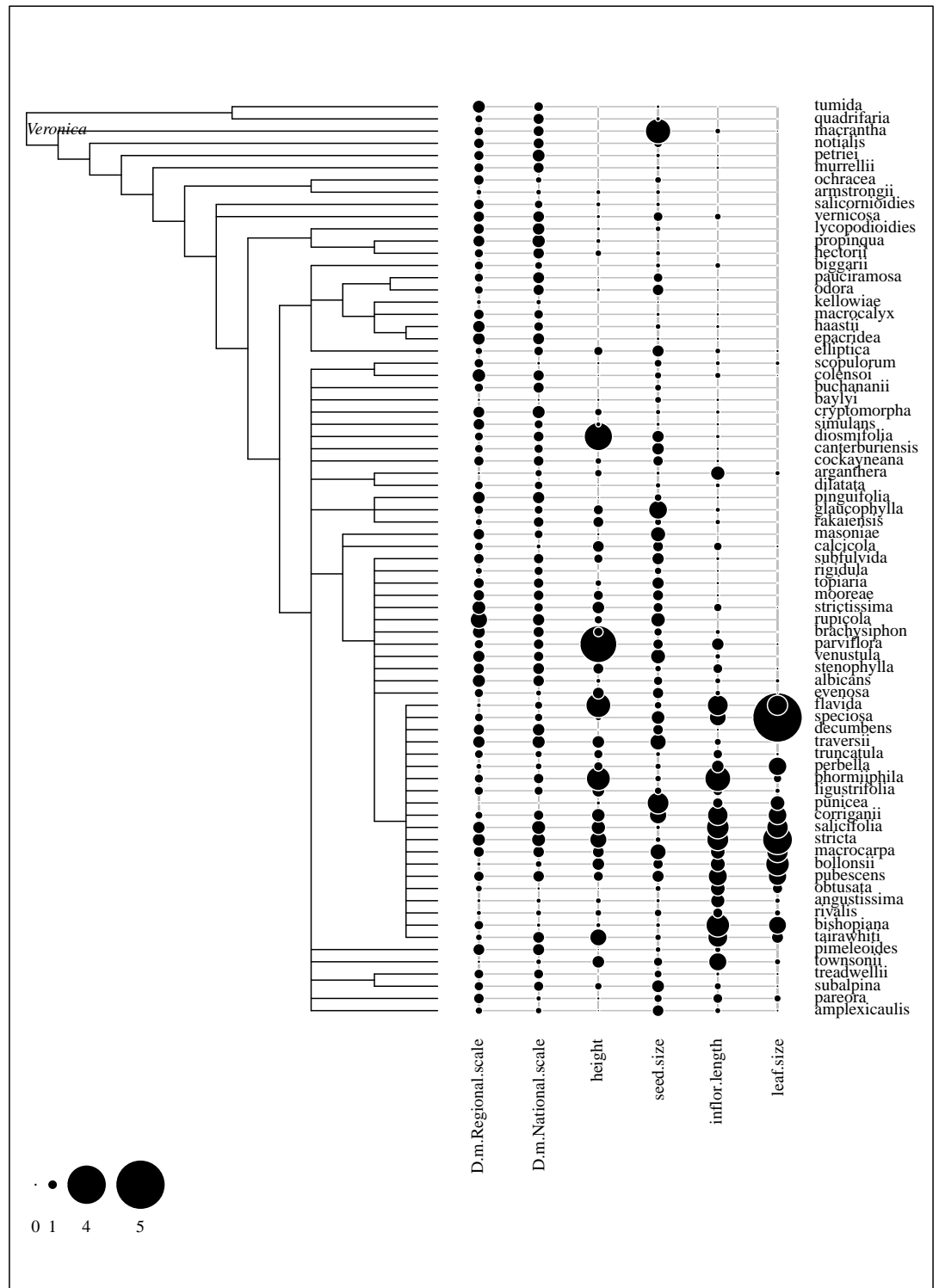
Figure 42: Representation of the life-history traits at the tips of the *Veronica* phylogeny. Circles of different size represent the ratio values of the life-history traits.
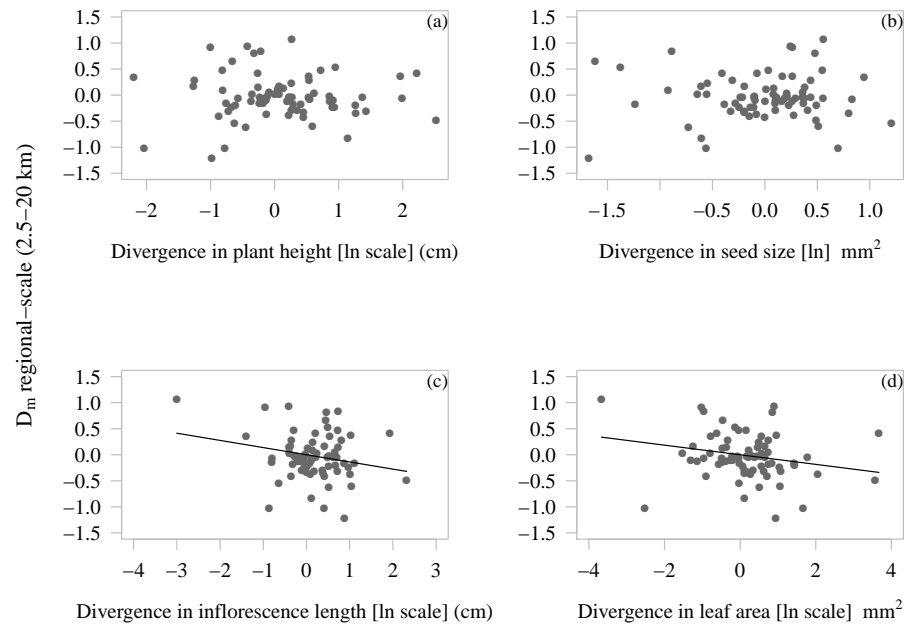
Figure 43: Scatter plots of variables used as predictors of distribution patterns (mass fractal dimension) at regional-scale for PIC analyses.
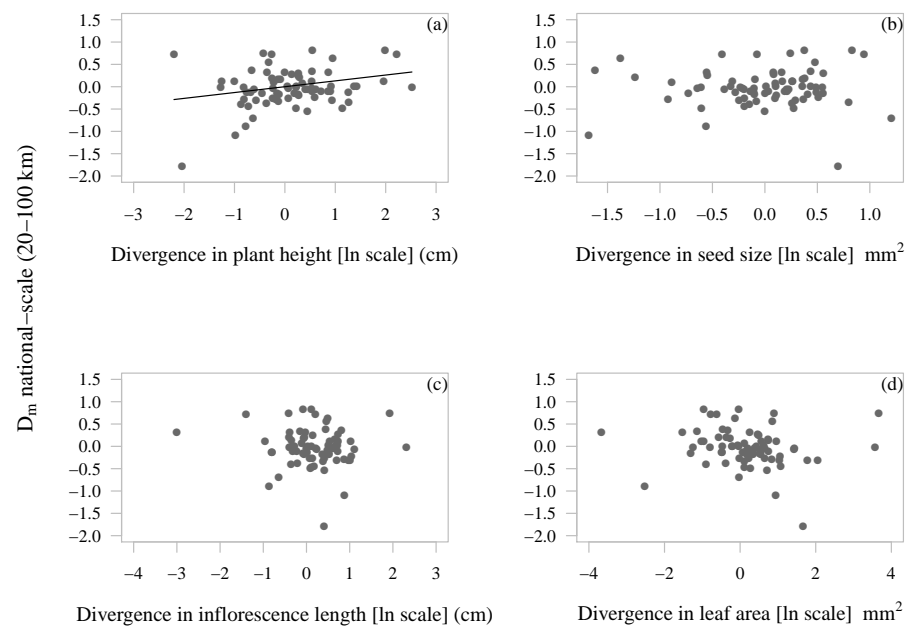
Figure 44: Scatter plots of variables used as predictors of distribution patterns (mass fractal dimension) at national-scale for PIC analyses.

4.4 DISCUSSION

Although life-history traits had low predictive ability in explaining distribution patterns in the *Veronica* complex, the results in this study showed relationships between some life-history traits and species aggregation. The specific nature of these relationships changed with the resolution at which aggregation (D_m) was measured (scale of analysis: local, regional and national scale).

4.4.1 *Life-history traits and aggregation*

Several significant relationships between aggregation and life-history traits were identified, for example, the positive relationship between leaf area and point-like distributions and the negative relationship between SLA and aggregation at local plot-scale. One possible explanation is that species with smaller and thicker leaves could be better equipped physically against predators and pathogens (Moles et al. 2011). This might result in more aggregated distributions if they were less susceptible to the type of density-dependent mortality suggested by Janzen and Connell (Janzen 1970; Connell 1971). At the same local-scale, but at a coarser resolution (60 m - 1 km) the presence of wings on the seed was associated with less aggregated distribution patterns. This result is in agreement with previous studies that have explored this relationship between dispersal traits and distribution patterns (e.g., Maurer et al. 2003; Kolb & Diekmann 2005; Tremlova & Munzbergova 2007). The fact that presence of wings was not associated with the finest scale D_m (2.5 - 60 m) suggests that wings on *Veronica* seeds make a less noticeable difference to distribution patterns at these finest scales compared to distances of 60 m - 1 km.

At regional scales, other plant traits different from local and national scale, were related to aggregation patterns. The relationship between breeding system and aggregation was in the expected direction: cosexual species had more point-like distribution. Species which can self-fertilise, like the cosexual ones, can form and persist in more scattered ranges as an isolated individual can potentially establish a population after a single, long distance dispersal event (Baker's Law, Baker 1955; Barrett et al. 1996). Similarly, I found that species with longer inflorescences were associated with isolated distributions; the length of the inflorescences could be related to increasing dispersal capacity. Lee et al. (2001) suggested that species with longer inflor-

escences can help in dispersing seeds further away from the mother plant to occupy new habitats, leading to more isolated distributions (low D_m values). Alternatively, inflorescence length could be related to the size, specificity and dispersal ability of the pollinators. Large, more specialist pollinators may be expected to affect pollen transfer over greater distances than small generalist pollinators. The observed pattern suggests that traits related to reproduction and potentially dispersal are influencing aggregation patterns at regional-scale (2.5-20 km) on the *Veronica* complex. These results were consistent in cross-species and PIC analyses.

Contrary to what I hypothesised, seed mass had no relationship to aggregation at any of the scales analysed. Seed mass has been considered one of the most important traits related to the distance that seeds can travel, thus influencing aggregation patterns (Venable et al. 2008). Nevertheless, recent studies have also found no relationship between seed mass and aggregation (Pocock et al. 2006) or to dispersal distance (Muller Landau et al. 2008; Thomson et al. in press). These results may indicate that it is not seed mass *per se*, but rather other strategies (e.g., traits related to establishment, longevity or actual dispersal) that are influencing plant distribution patterns. The test of these other plant life-history traits would enhance our understanding about the traits that are shaping distribution patterns.

An important finding from this study is that the relationship of plant life-history traits to distribution patterns (e.g., aggregation) can completely change when moving to a different scale of analysis or occasionally when using different methods of analysis (cross-species, PIC, or multivariate analysis). This was the case of the relationship between aggregation and plant height trait; in cross-species analysis this relationship was non-significant, and changed to significant in PIC analyses. *Veronica* species that comprise plants that grow taller had more space-filling or continuous distributions. Other studies (Kelly & Woodward 1996; Mathews & Bonser 2005) have shown that the height of plant species can result in increasing dispersal capacity into previously unoccupied habitats (Thomson et al. in press). Also height is associated with woody growth (Niklas 1997), and woody plants could persist and have long span life compared to herbs (Pocock et al. 2006).

The results of relationships between distribution patterns and life-history traits were consistent with cross-species analyses and PIC analyses with the exception of plant height at national scale. PIC and

PIC *phylogenetic independent contrast*

cross-species analyses give different information (see Westoby et al. 1995 for a discussion). The properties of cross-species analyses are of interest to us in their own right, not only as a source of evidence about evolutionary history but as to the selective forces that maintain present-day life history-traits.

Due to the strong correlation between leaf size and inflorescences length it is difficult to assess which of these traits is the main driver behind distribution patterns at the regional scale. Inflorescence length was the most significant trait in the cross-species univariate analysis, but dropped out of the multivariate minimum adequate model which recorded significant effects of leaf size and breeding system. In the phylogenetically informed analysis both leaf size and inflorescence length were again highlighted as significant predictors of regional scale distribution pattern. We may be approaching the limits of what can be inferred from a macroecological analysis of these data, nonetheless the results suggest areas ripe for more detailed ecological study of seed dispersal, pollination processes and population dynamics in New Zealand *Veronica* species. Results from other studies across a wider range of taxa suggest that growth rate, dispersal capacity and reproductive traits (Lloyd et al. 2002; Lloyd et al. 2003; Pocock et al. 2006) may be important, although many of these studies did not specifically include leaf properties or inflorescence length as potential explanatory factors.

4.4.2 *Scaling properties and the mass fractal dimension*

I found that distribution patterns had predictable scaling properties. Regional-scale distribution $(2.5 - 20$ km) can predict national-scale distribution $(20 - 100$ km; Figure 41a). The mass fractal dimension is useful in identifying and quantifying distribution patterns across scales (Nikora et al. 1999), which is important for ecological modelling and data interpretation. Furthermore, fractal geometry has been suggested to distinguish between expanding and contracting populations. Pocock et al. (2006) studied the distribution patterns of 391 British plants and found that species that were expanding range sizes had relatively more cohesive distributions (high D values), and those species that are declining were associated with scattered distributions (low D values). The length of this study is not long enough to detect temporal changes in tussock grassland plant communities; it is

necessary to consider the temporal scale in order to find out which species are expanding or declining. This is out of the scope of the present Chapter but it is interesting to note that the lowest values of the mass fractal dimension at regional and national-scale correspond to *Veronica* species that are in the threatened category (Nationally Endangered e.g., *V. arganthera*, *V. armstrongii*; or Nationally Vulnerable e.g., *V. baylyi*, *V. rivalis*; de Lange et al. 2009; Figure 35) of the New Zealand Threat Classification System (Townsend et al. 2008). The actual Threat Classification System is based on measures of abundance and occupancy of species data. Combining the actual knowledge of species distribution with measures of fractal dimension at different scales could improve and help us identify species that are declining, and also naturalised plants that could become invasive and are expanding range sizes. The analysis of distribution patterns therefore has the potential to inform future conservation efforts. Nevertheless, caution is needed when applying fractals in ecology (see Halley et al. 2004 for a discussion).

In this study, I have explored the effect of plant life-history traits on aggregation patterns across multiple scales. In conclusion, this study shows that when several scales of analysis are considered, we can improve our understanding about the factors that are related to species distribution patterns. Factors influencing distribution patterns at fine scales may be different from those influencing them at regional land national scales. In the case of New Zealand hebes the plant traits relating to ecological processes did not explain national-scale distribution patterns, but they did correlate with the degree of aggregation observed at local and regional scales. Moreover, the combination of methods of analysis (cross-species and phylogenetic independent contrast analysis) of these relationships, should be promoted in order to robust results from possible artefacts in the relationships between distribution patterns and plant traits, when species phylogeny is not taken into account in the analysis. The combination of these approaches investigated in this Chapter, is a promising way to disentangle the factors that influence distribution patterns.

4.5 REFERENCES

- Albach D.C. & Meudt H.M. (2010). Phylogeny of Veronica in the Southern and Northern Hemispheres based on plastid, nuclear ribosomal and nuclear low-copy DNA. Molecular phylogenetics and evolution, 54, 457-471.

— Baddeley A. & Turner R. (2005). Spatstat: an R package for analyzing spatial point patterns. Journal of Statistical Software, 12, 1-42.

— Baker H.G. (1955). Self-compatibility and establishment after'long-distance'dispersal. Evolution, 9, 347-349.

— Barrett S.C.H., Harder L.D. & Worley A.C. (1996). The comparative biology of pollination and mating in flowering plants. Philosophical Transactions: Biological Sciences, 1271-1280.

— Bayly M.J. & Kellow A.V. (2006). An illustrated guide to New Zealand hebes. Te Papa press, Wellington, NZ.

— Begon M., Townsend C.R. & Harper J.L. (2006). Ecology: from individuals to ecosystems. Wiley-Blackwell.

— Bivand R.S., Pebesma E.J. & Gómez-Rubio V. (2008). Applied spatial data analysis with R. Springer, New York, USA.

— Burnham K.P. & Anderson D.R. (2002). Model selection and multimodel inference: a practical information-theoretic approach. Springer Verlag, Colorado, USA.

— Burns K. (2004). Patterns in specific leaf area and the structure of a temperate heath community. Diversity and Distributions, 10, 105-112.

— Burns K.C. & Beaumont S. (2009). Scale-dependent trait correlations in a temperate tree community. Austral Ecology, 34, 670-677.

— Cieraad E. & Lee D.E. (2006). The New Zealand fossil record of ferns for the past 85 million years. New Zealand Journal of Botany, 44, 143-170.

— Condit R., Ashton P.S., Baker P., Bunyavejchewin S., Gunatilleke S., Gunatilleke N., Hubbell S.P., Foster R.B., Itoh A., LaFrankie J.V., Lee H.S., Losos E., Manokaran N., Sukumar R. & Yamakura T. (2000). Spatial patterns in the distribution of tropical tree species. Science, 288, 1414-1418.

— Clissold F.J., Sanson G.D., Read J. & Simpson S.J. (2009). Gross vs. net income: How plant toughness affects performance of an insect herbivore. Ecology, 90, 3393-3405.

— Connell J. (1971). On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees. Dynamics of populations, 298, 312.

— Cornelissen J.H.C., Lavorel S., Garnier E., Diaz S., Buchmann N., Gurvich D.E., Reich P.B., ter Steege H., Morgan H.D., van der Heijden M.G.A., Pausas J.G. & Poorter H. (2003). A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. Australian Journal of Botany, 51, 335-380.

— Cox C. (1998). From generalized tracks to ocean basins—how useful is Panbiogeography? Journal of Biogeography, 25, 813-828.

— Coulter J.D. (1969). The climate of Molesworth. Meteorological Service Miscellaneous Publication, New Zealand.

— Dale M.R.T. (2000). Spatial pattern analysis in plant ecology. Ecology, 88, 366-370.

— de Lange P., Norton D., Courtney S., Heenan P., Barkla J., Cameron E., Hitchmough R. & Townsend A. (2009). Threatened and uncommon plants of New Zealand(2008 revision). New Zealand Journal of Botany, 47, 61-96.

— Edwards W. & Westoby M. (1996). Reserve mass and dispersal investment in relation to geographic range of plant species: Phylogenetically independent contrasts. Journal of Biogeography, 23, 329-338.

— Eriksson O. & Jakobsson A. (1998). Abundance, distribution and life histories of grassland plants: a comparative study of 81 species. Journal of Ecology, 86, 922-933.

— Felsenstein J. (1985). Phylogenies and the comparative method. American Naturalist, 125, 1-15.

— Fortin M.J. & Dale M.R.T. (2005). Spatial Analysis: A guide for ecologist. Cambridge University Press., New York, United States of America.

— Garnock-Jones P., Albach D. & Briggs B.G. (2007). Botanical names in Southern Hemisphere Veronica (Plantaginaceae): sect. Detzneria, sect. Hebe, and sect. Labiatoides. Taxon, 56, 571-582.

— Grubb P.J. (2002). Leaf form and function–towards a radical new approach. New Phytologist, 155, 317-320.

— Guisan A. & Thuiller W. (2005). Predicting species distribution: offering more than simple habitat models. Ecology Letters, 8, 993-1009.

— Guo Q.F., Brown J.H., Valone T.J. & Kachman S.D. (2000). Constraints of seed size on plant distribution and abundance. Ecology, 81, 2149-2155.

— Haase P. (1995). Spatial pattern analysis in ecology based on Ripley's K-function: introduction and methods of edge correction. Journal of Vegetation Science, 6, 575-582.

— Hackathon R., Bolker B., Butler M., Cowan P., Vienne D.d., Eddelbuettel D., Holder M., Jombart T., Kembel S., Michonneau F., Orme D., O'Meara B., Paradis E., Regetz J. & Zwickl D. (2010). phylobase: Base package for phylogenetic structures and comparative data. URL http://CRAN.R-project.org/package=phylobase

— Halley J.M., Hartley S., Kallimanis A.S., Kunin W.E., Lennon J.J. & Sgardelis S.P. (2004). Uses and abuses of fractal methodology in ecology. Ecology Letters, 7, 254-271.

— Hanley M.E., Lamont B.B., Fairbanks M.M. & Rafferty C.M. (2007). Plant structural traits and their role in anti-herbivore defence. Perspectives in Plant Ecology, Evolution and Systematics, 8, 157-178.

— Hartley S. & Kunin W.E. (2003). Scale dependency of rarity, extinction risk, and conservation priority. Conservation Biology, 17, 1559-1570.

— He H.S., DeZonia B.E. & Mladenoff D.J. (2000). An aggregation index (AI) to quantify spatial patterns of landscapes. Landscape Ecology, 15, 591-601.

— Holt R.D., Keitt T.H., Lewis M.A., Maurer B.A. & Taper M.L. (2005). Theoretical models of species' borders: single species approaches. Oikos, 108, 18-27.

— Janzen D.H. (1970). Herbivores and the number of tree species in tropical forests. American Naturalist, 501-528.

— Jombart T., Balloux F. & Dray S.(2010). adephylo: new tools for investigating the phylogenetic signal in biological traits. Bioinformatics, 26, 1907-1909.

— Kelly C.K. (1996). Identifying plant functional types using floristic data bases: Ecological correlates of plant range size. Journal of Vegetation Science, 7, 417-424.

— Kelly C.K. & Woodward F.I. (1996). Ecological correlates of plant range size: Taxonomies and phylogenies in the study of plant commonness and rarity in Great Britain. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 351, 1261-1269.

— Kembel S.W., Cowan P.D., Helmus M.R., Cornwell W.K., Morlon H., Ackerly D.D., Blomberg S.P. & Webb C.O. (2010). Picante: R tools for integrating phylogenies and ecology. Bioinformatics, 26, 1463-1464.

— Kolb A. & Diekmann M. (2005). Effects of Life History Traits on Responses of Plant Species to Forest Fragmentation. Conservation Biology, 19, 929-938.

— Larcher W. (1975). Physiological plant ecology. Springer, Verlag, Berlin.

— Lee W.G., Macmillan B., Partridge T., Lister R. & Lloyd K. (2001). Fruit features in relation to the ecology and distribution of Acaena (Rosaceae) species in New Zealand. New Zealand Journal of Ecology, 25, 17-27.

— Legendre P. & Legendre L. (1998). Numerical ecology. 2nd edn. Elsevier Science B. V., Amsterdam, The Netherlands.

— Legendre P. & Desdevises Y. (2009). Independent contrasts and regression through the origin. Journal of Theoretical Biology, 259, 727-743.

— Lloyd D.G. (1985). Progress in understanding the natural history of New Zealand plants. New Zealand Journal of Botany, 23, 707-722.

— Lloyd K.M., Lee W.G. & Wilson J.B. (2002). Growth and reproduction of New Zealand *Acaena* (Rosaceae) species in relation to rarity and commonness. New Zealand Journal of Ecology, 26, 149-160.

— Lloyd K.M., Wilson J.B. & Lee W.G. (2003). Correlates of geographic range size in New Zealand *Chionochloa* (Poaceae) species. Journal of Biogeography, 30, 1751-1761.

— Low M.L.E. (2005). Evolution of gender dimorphism in Hebe (Plantaginaceae). Ph.D. Thesis. In: School of Biological Sciences. Victoria University of Wellington Wellington, New Zealand.

— Mathews S. & Bonser S.P. (2005). Life histories, ecological tolerance limits, and the evolution of geographic range size in Eucalyptus (Myrtaceae). Australian Journal of Botany, 53, 501-508.

— Maurer K., Durka W. & Stöcklin J. (2003). Frequency of plant species in remnants of calcareous grassland and their dispersal and persistence characteristics. Basic and Applied Ecology, 4, 307-316.

— McGlone M.S., Duncan R.P. & Heenan P.B. (2001). Endemism, species selection and the origin and distribution of the vascular plant flora of New Zealand. Journal of Biogeography, 28, 199-216.

— Moles A.T., Warton D.I. & Westoby M. (2003). Do small-seeded species have higher survival through seed predation than large-seeded species? Ecology, 84, 3148-3161.

— Moles A.T. & Westoby M. (2000). Do small leaves expand faster than large leaves, and do shorter expansion times reduce herbivore damage? Oikos, 90, 517-524.

– Moles A.T. & Westoby M. (2003). Latitude, seed predation and seed mass. Journal of Biogeography, 30, 105-128.

– Newstrom L. & Robertson A. (2005). Progress in understanding pollination systems in New Zealand. New Zealand Journal of Botany, 43, 1-59.

– Niklas K.J. (1997). The evolutionary biology of plants. University of Chicago Press.

– Nikora V.I., Pearson C.P. & Shankar U. (1999). Scaling properties in landscape patterns: New Zealand experience. Landscape Ecology, 14, 17-33.

– Paradis E., Claude J. & Strimmer K. (2004). APE: analyses of phylogenetics and evolution in R language. Bioinformatics, 20, 289.

– Pocock M.J.O., Hartley S., Telfer M.G., Preston C.D. & Kunin W.E. (2006). Ecological correlates of range structure in rare and scarce British plants. Journal of Ecology, 94, 581-596.

– Pole M. (1994). The New Zealand flora-entirely long-distance dispersal? Journal of Biogeography, 625-635.

– R Development Core Team (2011). R: A Language and Environment for Statistical Computing, Vienna, Austria. URL http://www.R-project.org

– Raven P.H. (1973). Evolution of subalpine and alpine plant groups in New Zealand. New Zealand Journal of Botany, 11, 177-200.

– Raven P.H., Raven T.E. & West K.R. (1976). The genus Epilobium (Onagraceae) in Australasia: a systematic and evolutionary study. New Zealand Dept. of Scientific and Industrial Research.

– Seuront L. (2009). Fractals and multifractals in ecology and aquatic science. CRC Press, New York, USA.

– Stephenson A. (1981). Flower and fruit abortion: proximate causes and ultimate functions. Annual Review of Ecology and Systematics, 12, 253-279.

– Venables W.N. & Ripley B.D. (2002). Modern Applied Statistics with S. Fourth edn. Springer, New York.

– Wagstaff S.J., Bayly M.J., Garnock-Jones P.J. & Albach D.C. (2002). Classification, origin, and diversification of the New Zealand hebes (Scrophulariaceae). Annals of the Missouri Botanical Garden, 89, 38-63.

– Wagstaff S.J. & Garnock-Jones P.J. (1998). Evolution and biogeography of the Hebe complex (Scrophulariaceae) inferred from ITS sequences. New Zealand Journal of Botany, 36, 425-437.

– Webb C.O. (2000). Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. American Naturalist, 145-155.

– Webb C.O., Ackerly D.D. & Kembel S.W. (2008). Phylocom: software for the analysis of phylogenetic community structure and trait evolution. Bioinformatics, 24, 2098-2100.

– Westoby M., Falster D.S., Moles A.T., Vesk P.A. & Wright I.J. (2002). Plant ecological strategies: some leading dimensions of variation between species. Annual Review of Ecology and Systematics, 33, 125-159.

– Westoby M., Leishman M.R. & Lord J.M. (1995). On misinterpreting the phylogenetic correction. Journal of Ecology, 83, 531-534.

– Winkworth R.C., Robertson A.W., Ehrendorfer F. & Lockhart P.J. (1999). The importance of dispersal and recent speciation in the flora of New Zealand. Journal of Biogeography, 26, 1323-1325.

– Winkworth R.C., Wagstaff S.J., Glenny D. & Lockhart P.J. (2002). Plant dispersal NEWS from New Zealand. Trends in Ecology & Evolution, 17, 514-520.

– Winkworth R.C., Wagstaff S.J., Glenny D. & Lockhart P.J. (2005). Evolution of the New Zealand mountain flora: Origins, diversification and dispersal. Organisms Diversity & Evolution, 5, 237-247.

– With K.A. & King A.W. (1999). Extinction thresholds for species in fractal landscapes. Conservation Biology, 13, 314-326.

– Wright I.J., Reich P.B., Westoby M., Ackerly D.D., Baruch Z., Bongers F., Cavender-Bares J., Chapin T., Cornelissen J.H.C. & Diemer M. (2004). The worldwide leaf economics spectrum. Nature, 428, 821-827.

– Wu J., Jones K.B., Li H. & Loucks O.L. (2006). Scaling and uncertainty analysis in ecology: methods and applications. Springer Verlag.

– Young A.G. & Brown A.H.D. (1998). Comparative analysis of the mating system of the rare woodland shrub Daviesia suaveolens and its common congener D-mimosoides. Heredity, 80, 374-381.

Part III

GENERAL DISCUSSION

GENERAL DISCUSSION

5.1 THESIS SUMMARY

In this thesis, I have discussed the importance of identifying and quantifying patterns and processes as well as the scale at which they are operating. In order to understand and manage ecological processes, first we need to identify and quantify ecological patterns (Fortin and Dale, 2005). In Chapter 2, five occupancy-area models were compared to choose the one best able to describe distribution patterns. The occupancy-area models have two important properties; one is that some parameters of the model can be used as descriptors of the patchiness of species across scales, the advantage of having a continuous measure of aggregation pattern is that we can test hypotheses of what process are generating these patterns; the other is the scaling property, predictions of occupancy can be made at a scale that is different from the scale at which data were acquired. For example, for predicting patterns at fine scale from coarse scale data or *vice versa* (He and Condit, 2007). As discussed in Chapter 2, although there was no overall superior occupancy-area model, we can use these models to deduce important information.

In Chapter 3, I utilised the aggregation parameters from two models: one from the occupancy-area model NBD2 investigated in Chapter 2 and the other from the mean and variance from the NBD1. This was done to test if those aggregation patterns were correlated with plant life-history traits related to ecological processes such as dispersal, competition and reproduction in a multi-scale approach. As scale changes, new patterns and processes may emerge. This is why a multi-scale approach has the advantage of detecting scale-specific associations between pattern and process that may not always be apparent at a single scale of study. For example, observations at fine scales may overlook important patterns and processes operating at coarser scales; conversely, coarse-scale observations may not have enough details necessary to understand fine-scale dynamics (Wu et al. 2006). This was the case of the results in Chapter 3, where cross-species analyses at local-scale identified specific leaf area as the life-history trait positively related to aggregation; while at regional scale the height of the plants

was positively related to aggregation patterns. However, the relationship of SLA and height could be an artefact since the phylogenetic independent contrast (PIC) analyses showed important divergences (big changes of values of traits in the phylogenetic tree at a particular node) in several life-history traits (Figure 45). Also, multivariate analyses showed that the combination of plant traits can have an additive effect in the relationship with aggregation patterns. The results of Chapter 4 confirmed this variation in the relationship between aggregation pattern and plant life-history traits, as well as the variation in the results when changing the scale of analysis. In this particular chapter I focused on one subgroup of plants of the Plantaginaceae family distributed nationally in New Zealand. Aggregation patterns were measured as the mass fractal dimension (D_m). This method works well with more-or-less continuous point patterns, while the method for measuring aggregation from Chapter 3 works conceptually from multi-scale measures of occupancy. Cross-species and PIC analyses showed similar results. Plant life-history traits like the inflorescence length, leaf area and breeding system were related to distribution patterns at a regional-scale (2.5 - 20 km). Multivariate analysis also confirmed these results, showing that leaf area, breeding system and seed size had an additive effect on the relationship between plant traits and aggregation. The main conclusion of this research is that traits that are directly or indirectly related to dispersal, competition and reproduction are likely to be influencing species distribution patterns but the relationships can be dependent on the scale of analysis (Figure 45).

5.2 IMPLICATIONS

Understanding species' distribution patterns and the processes that are generating these patterns is of high importance for planning biological conservation priorities (Myers et al. 2000), also for assessing the effect of expanding naturalised plants on the distribution on native plants (e.g., Bradshaw et al. 2008), and for monitoring and predicting species' distribution where data is not available or hard to acquire (Hurlbert and Haskell, 2003, He et al. 2003, Lennon et al. 2000, Hwang and He, in press). Therefore, spatial patterns at a variety of scales should be considered in most plant ecological studies and not simply be ignored or treated as a statistical nuisance. Here, I mention areas in which
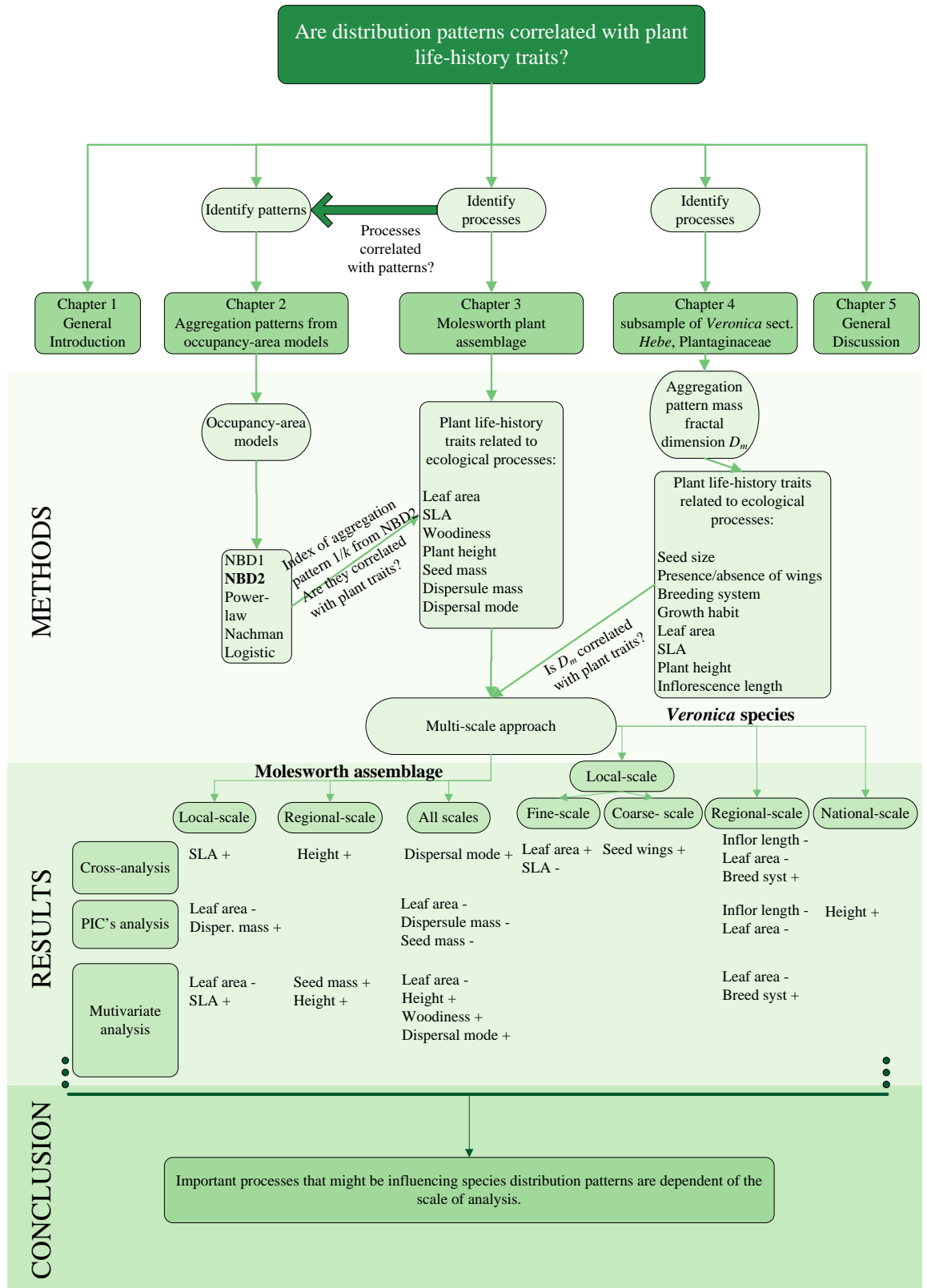
Figure 45: Diagram of thesis summary. The '+' and '−' signs indicate the direction of significant relationships in the results section.

knowledge on distribution patterns and plant life-history traits could improve assessment.

5.2.1 *Conservation*

The IUCN uses several measures in order to develop red books of threatened species (World Conservation Union, 2001). One of these measures is the rate of decline. The rate of decline can be obtained from the species occupancy-area relationships at two different times. The scale (e.g., area or resolution) at which the line shows the greatest vertical separation (on log-log axes) will be the scale at which the decline in percentage occupancy is the greatest (see example in Hartley and Kunin, 2003). In this thesis, I have used occupancy-area curves to quantify aggregation. These aggregation patterns in combination with the possible causes that are generating them (e.g., environmental factors, disturbance, and sets of life-history traits) could improve and detect more rapidly species that are declining. Decline is important to monitor since it could be a sign of a future dangerous state of small population sizes (Caughley, 1994) .

5.2.2 *Management and early identification of invasive species*

Exotic plants that naturalise and become invasive can generate dramatic ecological changes that include reductions in native species richness and altered ecosystem function (Hejda et al. 2009, Robson et al. 2009, D'Antonio and Vitousek, 1992). In this thesis, I have investigated if plant life-history traits correlate with distribution patterns at multiple scales. The same approach could be used to test if species' life-history traits correlate to the probability of becoming invasive (e.g., Phillips et al. 2010, Bradshaw et al. 2008). A multi-scale measurement would provide important information on the spatial characteristics of the species' invasiveness as well (Veldtman et al. 2010).

5.2.3 *Survey and Monitoring*

An important application of the occupancy-area curves is to estimate abundance from occupancy (He and Condit, 2007, He and Gaston, 2000a). The results from Chapter 2 suggested that the accuracy in

predicting occupancy at fine scales was improved when more resolutions were used. Surveys that need abundance data could be obtained from coarse occupancy-area estimations. These estimations can be conducted at a scale so fine that each occupied cell would contain only a single individual, thus allowing the estimation of abundance from occupancy.

5.2.4 *Ecology framework*

One of the main goals in ecology is to explain the distribution and abundance of species, however this have been proved to be challenging (Kunin et al. 2000). One of the key problems is that there are many types of abundance or rarity as scale changes (Rabinowitz, 1981). This was discussed in Chapter 3. Insights on explaining the distribution and abundance of species could improve if a multi-scale approach is taken. Processes that could explain distribution patterns should be measured at different scales so that we can disentangle which processes are influencing at specific scales. A multi-scale approach was considered throughout this thesis. Although I identified some processes that were related to distributions patterns, further work is clearly needed to test more of these hypotheses.

5.3 FUTURE DIRECTIONS

The understanding of how plant species are distributed on earth and how plant life-history traits influence plant distribution has increased markedly in recent years (Pocock et al. 2006, Van der Veken et al. 2007, Murray et al. 2002, Ackerly et al. 2002, Tremlova and Munzbergova, 2007). In this regard, I discuss some areas that still need more accuracy when measured or others that need to be considered. In the following section I make a set of suggestions that could be considered in order to fill in the gaps left that my research did not cover:

1. Predictability: The five most common occupancy-area models have been tested in a limited number of studies (e.g., He and Condit, 2007). There is still some degree of under and over-estimation of occupancy in the models. When scaling across resolutions (e.g., area, scale) of hundred or thousands of square

kilometres the overestimation or underestimation could be seriously affected. A major challenge is to develop models that would prove to be more accurate. One possible solution is to combine coarse-scale (hundreds of km²) and fine-scale (hundreds of m²) occupancy data (Gaston and He, 2010), although this is not always possible (but see Hartley et al. 2004, Veldtman et al. 2010).

2. Temporal scale: the contents of this thesis and of most occupancy-area studies have focused mainly on a short period of time, probably because of budget constraints or because the length of a PhD research is not long enough to detect temporal dynamics. This study did not consider the temporal dynamics of occupancy. Natural dynamics and changes in the sup-alpine plant community can take around 20 years to be noticed (Husheer comm. pers.). Nevertheless, permanent quadrats have been set up in the study area for future assessment. Also, each corner of the quadrats has been georeferenciated with a precision of ±5 cm. Future research might consider the temporal scale so that changes in distribution patterns as a consequence of other mechanisms such as colonisation/extinction dynamics can be identified (e.g., Wilson et al. 2004). These mechanisms could play a key role in determining occupancy (Gaston and He, 2010). Also changes over a period of time in occupancy due to anthropogenic pressures (e.g., climate change, livestock grazing), over the long term could be assessed.

3. Life-history traits: Although in this thesis I used eight plant life-history traits to test the prediction of aggregation, it would be desirable to test other traits related to vegetative expansion and multiplication, seed bank longevity and seed production. The building of open databases of plant life-history traits per region or country is very much needed in order to synthesise and gather the scattered information over many sources that are not easy to access. Furthermore, the use of these traits over a large number of species would increase the robustness in the analyses that investigate the influence of those traits in species distribution. An example of this type of data bases that are open-access and accessible on line are the LEDA plant Traitbase (Kleyer et al. 2008) based on the Northwest European flora and

the TRY global-scale database of vascular plant traits (Kattge et al. in press.)

4. Mechanisms: some life-history traits related to ecological mechanisms such as dispersal, competition and reproduction were related to distribution patterns at specific scales. Nevertheless other mechanisms might influence the present-day species' distribution. Future research is needed to determine which other mechanisms are related to distribution patterns (e.g., disturbance regimes, micro-environmental factors, and physiological traits).

This thesis also suggested further questions, particularly the role of invertebrates as pollinators and the role of native skinks and other animals as seed dispersers. These organisms might influence the present-day distribution pattern of sub-alpine plant assemblage.

5.4 CONCLUDING REMARKS

Here, I have demonstrated how processes related to dispersal, competition and reproduction may influence the present-day distribution of plant species. My main conclusions are:

– Patterns and processes are influenced by the spatial scale at which they are measured and different processes are important at different scales. A multi-scale approach of species' distribution should be promoted so that more accurate assessments can be generated.

– Occupancy-area models captured an important aspect of spatial patterns (e.g., aggregation).

– The aggregation pattern k from the negative binomial, calculated from nonlinear regression using occupancy data, could replace the k from the standard single-scale negative binomial when count data is not available.

– Distribution patterns had predictable scaling properties. However, the range of scales at which the predictions can be made are limited, most likely due to non-overlapping processes at finer and coarser scales (Hartley et al. 2004, Wiens, 1989).

– Some types of rarity are more common than others. Species with low densities and restricted range (occupancy) were more common (*R* type of rarity).

– In general, plant life-history traits had relatively low predictive ability in explaining distribution patterns.

– Leaf area was negatively related to distribution patterns, except at a very fine scale (2 - 60 m) where this relationship changed to positive. Leaf area was correlated in four out of the seven cases examined.

– Life-history traits showed low predictive ability in explaining aggregation throughout this thesis, nevertheless there was a general pattern about which processes and traits were important at different scales. At local scales traits related to dispersal and competition such as SLA, leaf area, dispersule mass and the presence of structures in seeds for dispersal, were important; while at regional scales traits related to reproduction such as breeding system, inflorescence length and traits related to dispersal (seed mass) were significant. At national scales only plant height was important in predicting aggregation.

5.5 REFERENCES

– Ackerly, D. D., Knight, C. A., Weiss, S. B., Barton, K. & Starmer, K. P. (2002) Leaf size, specific leaf area and microhabitat distribution of chaparral woody plants: contrasting patterns in species level and community level analyses. Oecologia, 130, 449-457.

– Bradshaw, C. J. A., Giam, X. L., Tan, H. T. W., Brook, B. W. & Sodhi, N. S. (2008) Threat or invasive status in legumes is related to opposite extremes of the same ecological and life-history attributes. Journal of Ecology, 96, 869-883.

– Caughley, G. (1994) Directions in conservation biology. Journal of Animal Ecology, 63, 215-244.

– D'Antonio, C. M. & Vitousek, P. M. (1992) Biological invasions by exotic grasses, the grass/fire cycle, and global change. Annual Review of Ecology and Systematics, 23, 63-87.

– Fortin, M. J. & Dale, M. R. T. (2005) Spatial Analysis: A guide for ecologist. Cambridge University Press., New York, United States of America.

– Gaston, K. J. & He, F. (2010) Species occurrence and occupancy. Biological diversity: Frontiers in measurement and assessment (eds A. E. Magurran & B. McGill), pp. 141-151. Oxford University Press, Oxford, UK.

– Hartley, S. & Kunin, W. E. (2003) Scale dependency of rarity, extinction risk, and conservation priority. Conservation Biology, 17, 1559-1570.

— Hartley, S., Kunin, W. E., Lennon, J. J. & Pocock, M. J. O. (2004) Coherence and discontinuity in the scaling of specie's distribution patterns. Proceedings of the Royal Society of London. Series B: Biological Sciences, 271, 81.

— He, F. & Condit, R. (2007) The distribution of species: occupancy, scale, and rarity. Scaling Biodiversity (eds D. Storch, P. A. Marquet & J. H. Brown). Cambridge University Press, Cambridge.

— He, F., Zhou, J. & Zhu, H. (2003) Autologistic regression model for the distribution of vegetation. Journal of agricultural, biological, and environmental statistics, 8, 205-222.

— He, F. L. & Gaston, K. J. (2000a) Estimating species abundance from occurrence. American Naturalist, 156, 553-559.

— Hejda, M., Pyšek, P. & Jarošík, V. (2009) Impact of invasive plants on the species richness, diversity and composition of invaded communities. Journal of Ecology, 97, 393-403.

— Hurlbert, A. H. & Haskell, J. P. (2003) The effect of energy and seasonality on avian species richness and community composition. The American Naturalist, 161, 83-97.

— Hwang, W. H. & He, F. (in press) Estimating abundance from presence/absence maps. Methods in Ecology and Evolution.

— Kattge, J., Diaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Bonisch, G., Garnier, E., Westoby, M., Reich, P. B. & Wright, I. A. N. J. (in press) TRY–a global database of plant traits. Global Change Biology.

— Kleyer, M., Bekker, R., Knevel, I., Bakker, J., Thompson, K., Sonnenschein, M., Poschlod, P., Van Groenendael, J., Klimeš, L. & Klimešová, J. (2008) The LEDA Traitbase: a database of life history traits of the Northwest European flora. Journal of Ecology, 96, 1266-1274.

— Kunin, W. E., Hartley, S. & Lennon, J. J. (2000) Scaling down: On the challenge of estimating abundance from occurrence patterns. American Naturalist, 156, 560-566.

— Lennon, J., Greenwood, J. & Turner, J. (2000) Bird diversity and environmental gradients in Britain: a test of the species–energy hypothesis. Journal of Animal Ecology, 69, 581-598.

— Murray, B. R., Thrall, P. H., Gill, A. M. & Nicotra, A. B. (2002) How plant life-history and ecological traits relate to species rarity and commonness at varying spatial scales. Austral Ecology, 27, 291-310.

— Phillips, M. L., Murray, B. R., Leishman, M. R. & Ingram, R. (2010) The naturalization to invasion transition: Are there introduction history correlates of invasiveness in exotic plants of Australia? Austral Ecology, 35, 695-703.

— Pocock, M. J. O., Hartley, S., Telfer, M. G., Preston, C. D. & Kunin, W. E. (2006) Ecological correlates of range structure in rare and scarce British plants. Journal of Ecology, 94, 581-596.

— Rabinowitz, D. (1981) Seven forms of rarity. The Biological Aspects of Rare Plant Conservation (ed H. Synge), pp. 205-217. Wiley, Great Britain.

— Robson, T. C., Baker, A. C. & Murray, B. R. (2009) Differences in leaf litter invertebrate assemblages between radiata pine plantations and neighbouring native eucalypt woodland. Austral Ecology, 34, 368-376.

– Tremlova, K. & Munzbergova, Z. (2007) Importance of species traits for species distribution in fragmented landscapes. Ecology, 88, 965-977.

– Van der Veken, S., Bellemare, J., Verheyen, K. & Hermy, M. (2007) Life-history traits are correlated with geographical distribution patterns of western European forest herb species. Journal of Biogeography, 34, 1723-1735.

– Veldtman, R., Chown, S. L. & McGeoch, M. A. (2010) Using scale–area curves to quantify the distribution, abundance and range expansion potential of an invasive species. Diversity and Distributions, 16, 159-169.

– Wiens, J. (1989) Spatial scaling in ecology. Functional Ecology, 3, 385-397.

– Wilson, R. J., Thomas, C. D., Fox, R., Roy, D. B. & Kunin, W. E. (2004) Spatial patterns in species distributions reveal biodiversity change. Nature, 432, 393-396.

– World Conservation Union (2001) IUCN Red list Categories and Criteria (version 3.1). World Conservation Union, Gland, Switzerland and Cambridge, United Kindom.

– Wu, J., Jones, K. B., Li, H. & Loucks, O. L. (2006) Scaling and uncertainty analysis in ecology: methods and applications. Springer Verlag.

Part IV

APPENDICES

FIVE OCCUPANCY-AREA MODELS R SCRIPT

The function `calc.SOR` fits five different occupancy-area models (negative binomial 1 and 2, power-law, Nachman and logistic) using `nls`.

INPUTS

- Vector 1. = `SU.area` = "sample unit area" or *a*: the area of a single sample unit.

- Vector 2. = `Occ` = "occupancy" or "A*a*" : the number of sample units or their combined area, at a particular size of sample unit).

- Scalar 1. = `TotalArea` : total area of all the sample units under consideration (e.g., for tussock data, the area of subquadrats in a transect).

- `occ.measure` = "number" or "area". Are the above two inputs measuring the total number of (occupied) quadrats or their combined area (default)?

- `graph` = `TRUE` or `FALSE`. If `TRUE`, plots a graph of the data points and the fitted line to the current plotting device.

NOTES

- if using "proportional occupancy" set occ.measure = "area", and Total.Area = 1

- Nachman and Logistic can only fit occupancies in the range 0-1, hence Area.Occ/A

```
calc.SOR = function(SU.area, occ, graph = "TRUE", NBD.start = c
    (1.05,0.05), ...) {
        if(length(SU.area) != length(occ))        print("Warning,
            input vectors of unequal length")
        if(length(SU.area) < 2)                   print("Warning,
            input vectors must contain data from 2 or more scales
            ")
        input.df <- data.frame(SU.area, occ)
        colnames(input.df) <- c("SU.area", "occ")
        A <- rep(1, length(SU.area))
```

```r
    if(graph == TRUE) plot(input.df, xlab = (expression(paste
        ("Size of quadrat", (m^2)))), ylab = "Proportion
        occupied", cex = 1.3,pch=16, font.main=3,cex.main=1)
    if(graph == TRUE) newdat <- seq(min(input.df[,1]), max(
        input.df[,1], length = 100))


# NBD1 nls
    model.fit <- a2 <- b2 <- sigma2 <- NLL2 <- NA
    df2 <- c(NA, NA)
    try(model.fit <- nls(occ/A ~ (1-(1+((SU.area)/(k)))^(-k))
        ,data=input.df, start=list(k=0.1), algorithm = "port
        ", lower = c(0), trace = FALSE), silent = TRUE)
    try((a2    <- coef(model.fit)[1]),    silent = TRUE)
    try((sigma2 <- summary(model.fit)$sigma),silent = TRUE)
    try((NLL2   <- logLik(model.fit)),silent = TRUE)
    try((df2    <- summary(model.fit)$df[1:2]),silent = TRUE)
    try((AIC        <- AIC(model.fit)),silent = TRUE)
    output.uNB  <- c(a2, b2, sigma2, NLL2, df2[1], df2[2],
        AIC)
    if(graph == TRUE) try(lines(newdat, predict(model.fit,
        list(SU.area = newdat)), col="#7fc97f",lwd=2), silent
         = TRUE)

# NBD2 nls
    model.fit <- a2 <- b2 <- sigma2 <- NLL2<- NA
    df2 <- c(NA, NA)
    try(model.fit <- nls(occ ~ A*(1-(1+((N*SU.area)/(A*k)))
        ^(-k)),data=input.df, start=list(N=NBD.start[1], k=
        NBD.start[2]), algorithm = "port", lower = c(0,0),
        trace = FALSE), silent = TRUE)
    try((a2    <- coef(model.fit)[1]),silent = TRUE)
    try((b2    <- coef(model.fit)[2]),silent = TRUE)
    try((sigma2 <- summary(model.fit)$sigma),silent = TRUE)
    try((NLL2   <- logLik(model.fit)),silent = TRUE)
    try((df2    <- summary(model.fit)$df[1:2]),silent = TRUE)
    try((AIC        <- AIC(model.fit)),silent = TRUE)
    output.cNB  <- c(a2, b2, sigma2, NLL2, df2[1], df2[2],
        AIC)
    if(graph == TRUE) try(lines(newdat, predict(model.fit,
        list(SU.area = newdat)), col="#beaed4",lwd=2), silent
         = TRUE)

# power law nls
    model.fit <- a2 <- b2 <- sigma2 <- NLL2<- NA
    df2 <- c(NA, NA)
    try(model.fit <- nls(occ ~ a*(SU.area^b),data=input.df,
        start=list(a=1.01, b=1.01), algorithm = "port", lower
         = c(-10000,0), trace = FALSE), silent = TRUE)
    try((a2    <- coef(model.fit)[1]),silent = TRUE)
```

```
        try((b2     <- coef(model.fit)[2]),silent = TRUE)
        try((sigma2 <- summary(model.fit)$sigma),silent = TRUE)
        try((NLL2   <- logLik(model.fit)),silent = TRUE)
        try((df2    <- summary(model.fit)$df[1:2]),silent = TRUE)
        try((AIC        <- AIC(model.fit)),silent = TRUE)
        output.power.nls <- c(a2, b2, sigma2, NLL2, df2[1], df
            2[2], AIC)
        if(graph == TRUE) try(lines(newdat, predict(model.fit,
            list(SU.area = newdat)), col="#fdc086",lwd=2), silent
             = TRUE)

# Nachman nls
        model.fit <- a2 <- b2 <- sigma2 <- NLL2<- NA
        df2 <- c(NA, NA)
        try(model.fit <- nls(occ ~ 1 - exp(-c*(SU.area^z)),data=
            input.df, start=list(c=0.1, z=0.5), algorithm =    "
            port", lower = c(0,0), trace = FALSE), silent = TRUE)
        try((a2     <- coef(model.fit)[1]),silent = TRUE)
        try((b2     <- coef(model.fit)[2]),silent = TRUE)
        try((sigma2 <- summary(model.fit)$sigma),silent = TRUE)
        try((NLL2   <- logLik(model.fit)),silent = TRUE)
        try((df2    <- summary(model.fit)$df[1:2]),silent = TRUE)
        try((AIC        <- AIC(model.fit)),silent = TRUE)
        output.nachman <- c(a2, b2, sigma2, NLL2, df2[1], df2[2],
            AIC)
        if(graph == TRUE) try(lines(newdat, predict(model.fit,
            list(SU.area = newdat)), col="#bd0026",lwd=2.5),
            silent = TRUE)

# Logistic nls
        model.fit <- a2 <- b2 <- sigma2 <- NLL2<- NA
        df2 <- c(NA, NA)
        try(model.fit <- nls(occ/A ~ (c*(SU.area^z))/(1+(c*(SU.
            area^z))),data=input.df, start=list(c=0.1, z=0.5),
            algorithm = "port", lower = c(0,0), trace = FALSE),
            silent = TRUE)
        try((a2     <- coef(model.fit)[1]),silent = TRUE)
        try((b2     <- coef(model.fit)[2]),silent = TRUE)
        try((sigma2 <- summary(model.fit)$sigma),silent = TRUE)
        try((NLL2   <- logLik(model.fit)),silent = TRUE)
        try((df2    <- summary(model.fit)$df[1:2]),silent = TRUE)
        try((AIC        <- AIC(model.fit)),silent = TRUE)
        output.logistic <- c(a2, b2, sigma2, NLL2, df2[1], df
            2[2],AIC)
        if(graph == TRUE) try(lines(newdat, predict(model.fit,
            list(SU.area = newdat)), col="#386cb0",lwd=2), silent
             = TRUE)

        output.m <- rbind(output.uNB, output.cNB, output.power.
            nls, output.nachman, output.logistic)
        output.v <- data.frame(a2, b2, sigma2, NLL2, df2[1], df
            2[2],AIC)
```

```
        colnames(output.m) <- c("a", "b", "sigma", "NLL", "p", "n
            -p","AIC")
        rownames(output.m) <- c("NBD1", "NBD2", "power.nls", "
            nachman", "logistic")
        return(output.m)
}
```

PLANT SPECIES SURVEYED

B.1 SPECIES CODES, SCIENTIFIC NAMES AND FAMILIES

Table 20: Species list surveyed in the project. Species are sorted as follows: mosses, gymnosperms, dicotyledonous and monocotyledinous

SPECIES CODE	SPECIES NAME	FAMILY
RAClan	*Racomitrium lanuginosum*	Grimmiaceae
LIBbid	*Libocedrus bidwillii*	Cupressaceae
PINcon	*Pinus contorta*	Pinaceae
PSEmen	*Pseudotsuga menziesii*	Pinaceae
PHYalp	*Phyllocladus alpinus*	Podocarpaceae
PODniv	*Podocarpus nivalis*	Podocarpaceae
ACIaur	*Aciphylla aurea*	Apiaceae
ACIgla	*Aciphylla glaucescens*	Apiaceae
ACImon	*Aciphylla monroi*	Apiaceae
BRAcass	*Brachyglottis cassinioides*	Asteraceae
CELdis	*Celmisia discolor*	Asteraceae
CELdur	*Celmisia du-rietzii*	Asteraceae
CELinc	*Celmisia incana*	Asteraceae
CELmon	*Celmisia monroi*	Asteraceae
CELsem	*Celmisia semicordata*	Asteraceae
CELspe	*Celmisia spectabilis*	Asteraceae
HELpar	*Helichrysum parvifolium*	Asteraceae
HIEpil	*Hieracium pilosella*	Asteraceae
OLEnum	*Olearia nummulariifolia*	Asteraceae
OZOvau	*Ozothamnus vauvilliersii*	Asteraceae
RAOgran	*Raoulia grandiflora*	Asteraceae
STEgrm	*Stellaria graminea*	Caryophyllaceae
CARcor	*Carex coriacea*	Cyperaceae
SCHpau	*Schoenus pauciflorus*	Cyperaceae
ACRcol	*Acrothamnus colensoi*	Ericaceae
DRApro	*Dracophyllum pronum*	Ericaceae
DRAros	*Dracophyllum rosmarinifolium*	Ericaceae

continued on next page...

SPECIES CODE	SPECIES NAME	FAMILY
DRAsp	*Dracophyllum* sp.	Ericaceae
GAUcra	*Gaultheria crassa*	Ericaceae
GAUdep	*Gaultheria depressa*	Ericaceae
LEUfra	*Leucopogon fraseri*	Ericaceae
GENcor	*Gentianella corymbifera*	Gentianaceae
CARarb	*Carmichaelia arborea*	Leguminosae
LEPsco	*Leptospermum scoparium*	Myrtaceae
VERbra	*Veronica brachysiphon*	Plantaginaceae
VERcan	*Veronica canterburiensis*	Plantaginaceae
VERcry	*Veronica cryptomorpha*	Plantaginaceae
VERdec	*Veronica decumbens*	Plantaginaceae
VERlyc	*Veronica lycopodioides*	Plantaginaceae
VERodo	*Veronica odora*	Plantaginaceae
VERpim	*Veronica pimeleoides*	Plantaginaceae
VERping	*Veronica pinguifolia*	Plantaginaceae
VERrak	*Veronica rakaiensis*	Plantaginaceae
DIStou	*Discaria toumatou*	Rhamnaceae
ACAine	*Acaena inermis*	Rosaceae
ROSrub	*Rosa rubiginosa*	Rosaceae
PIMser	*Pimelea sericeovillosa*	Thymelaeaceae
MELalp	*Melicytus alpinus*	Violaceae
ASTner	*Astelia nervosa*	Asteliaceae
JUNeff	*Juncus effusus*	Juncaceae
ANTodo	*Anthoxanthum odoratum*	Poaceae
CHIaus	*Chionochloa australis*	Poaceae
CHIpal	*Chionochloa pallens*	Poaceae
CHIrub	*Chionochloa rubra*	Poaceae
CHIsp	*Chionochloa* sp.	Poaceae
DACglo	*Dactylis glomerata*	Poaceae
RYTset	*Rytidosperma setifolium*	Poaceae

End
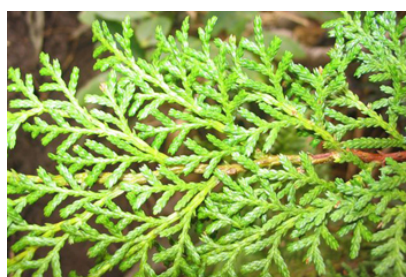
(a) *Racomitrium lanuginosum*, Grimmiaceae

(b) *Pinus contorta*, Pinaceae*

(c) *Pseudotsuga menziesii*, Pinaceae*. Photographer: Colin Ogle

(d) *Libocedrus bidwillii*, Cupressaceae. Photographer: John Barkla

(e) *Phyllocladus alpinus*, Podocarpaceae

(f) *Podacarpus nivalis*, Podocarpaceae

Figure 46: Plant species surveyed, mosses and gymnosperms. *Introduced

(a) *Astelia nervosa*, Asteliaceae

(b) *Juncus effusus*, Juncaceae*

(c) *Anthoxanthum odoratum*, Poaceae*

(d) *Chionochloa australis*, Poaceae

(e) *Chionochloa pallens*, Poaceae

(f) *Chionochloa rubra*, Poaceae

(g) *Rytidosperma setifolium*, Poaceae

Figure 47: Plant species surveyed, monocotyledinous. *Introduced

(a) *Aciphylla aurea*, Apiaceae

(b) *Aciphylla glaucescens*, Apiaceae

(c) *Aciphylla monroi*, Apiaceae

(d) *Acrothamnus colensoi*, Ericaceae

(e) *Brachyglottis cassinioides*, Asteraceae

(f) *Celmisia du-rietzii*, Asteraceae

(g) *Celmisia incana*, Asteraceae

(h) *Celmisia monroi*, Asteraceae

(i) *Celmisia semicordata*, Asteraceae

(j) *Celmisia spectabilis*, Asteraceae

(k) *Helichrysum parvifolium*, Asteraceae

(l) *Hieracium pilosella*, Asteraceae*

(m) *Olearia nummulariifolia*, Asteraceae
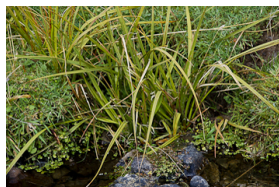
(n) *Ozothamnus vauvilliersii*, Asteraceae

(o) *Raoulia grandiflora*, Asteraceae

(p) *Stellaria graminea*, Caryophyllaceae*

(q) *Carex coriacea*, Cyperaceae

(r) *Schoenus pauciflorus*, Cyperaceae

(s) *Dracophyllum pronum*, Ericaceae

(t) *Dracophyllum sp.*, Ericaceae

(u) *Dracophyllum rosmarinifolium*, Ericaceae

Figure 48: Plant species surveyed, dicotyledonous I. *Introduced

(a) *Gaultheria crassa,* Ericaceae

(b) *Gaultheria depressa,* Ericaceae

(c) *Leucopogon fraseri,* Ericaceae

(d) *Gentianella corymbifera,* Gentianaceae

(e) *Carmichaelia arborea,* Leguminosae

(f) *Leptospermum scoparium,* Myrtaceae

(g) *Veronica canterburiensis,* Plantaginaceae

(h) *Veronica cryptomorpha,* Plantaginaceae

(i) *Veronica decumbens,* Plantaginaceae

(j) *Veronica lycopodioides,* Plantaginaceae

(k) *Veronica pimeleoides,* Plantaginaceae

(l) *Veronica pinguifolia,* Plantaginaceae

(m) *Veronica rakaiensis,* Plantaginaceae

(n) *Discaria toumatou,* Rhamnaceae
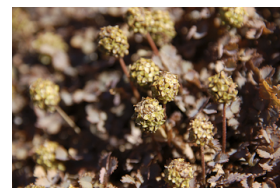
(o) *Acaena inermis,* Rosaceae

(p) *Rosa rubiginosa,* Rosaceae*
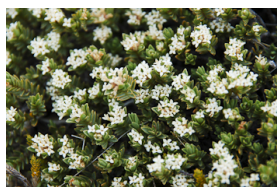
(q) *Pimelea sericeovillosa,* Thymelaeaceae

(r) *Melicytus alpinus,* Violaceae

Figure 49: Plant species surveyed, dicotyledonous II. *Introduced

B.2.1 *Fruits*

(a) *Aciphylla aurea*, Apiaceae

(b) *Acrothamnus colensoi*, Ericaceae

(c) *Gaultheria depressa*, Ericaceae

(d) *Leucopogon fraseri*, Ericaceae

(e) *Veronica pimeleoides*, Plantaginaceae

(f) *Rosa rubiginosa*, Rosaceae*

(g) *Discaria toumatou*, Rhamnaceae

(h) *Pimelea sericeovillosa*, Thymelaeaceae

Figure 50: Fruits

B.2.2 *Seeds*
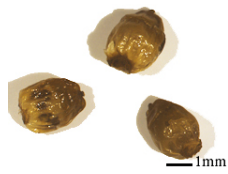
(a) *Phyllocladus alpinus*, Podocarpaceae

(b) *Podocarpus nivalis*, Podocarpaceae

(c) *Aciphylla aurea*, Apiaceae

(d) *Brachyglottis cassinioides*, Asteraceae

(e) *Celmisia discolor*, Asteraceae

(f) *Celmisia spectabilis*, Asteraceae
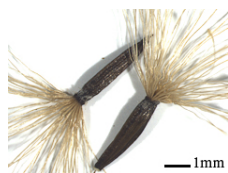
(g) *Olearia nummulariifolia*, Asteraceae

(h) *Ozothamnus vauvilliersii*, Asteraceae

(i) *Acrothamnus colensoi*, Ericaceae

(j) *Gaultheria depressa*, Ericaceae

(k) *Leucopogon fraseri*, Ericaceae

(l) *Gentianella corymbifera*, Gentianaceae

(m) *Rosa rubiginosa*, Rosaceae

(n) *Acaena inermis*, Rosaceae

(o) *Veronica decumbens*, Plantaginaceae

(p) *Veronica rakaiensis*, Plantaginaceae
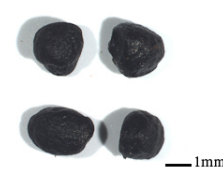
(q) *Melicytus alpinus*, Violoaceae

Figure 51: Seeds

INSTRUCTIONS FOR CALCULATING
PHYLOGENETIC INDEPENDENT CONTRAST IN R

After formatting the set of species you are working on in newick system, open the R software and install the adephylo, ape, phylobase and picante packages. Require the packages in R

```
library(adephylo)
library(ape)
library(phylobase)
library(picante)
```

Once in R, copy and paste the tree in an object. Be careful to split the phylogenetic tree in several parts or it won't work. Also you need to write "" in the beginning and at the end of each part of the tree

```
w <- NULL
w[1]<-"((((Aciphylla_aurea:114.0,(Brachyglottis_cassinioides
    :76.0,(Helichrysum_selago:38.0,Ozothamnus_vauvilliersii:38.0)
    :38.0,Olearia_nummulariifolia:76.0)asteraceae:38.0):9.0,(
    Gentianella_corymbifera:108.0,(Veronica_brachysiphon:3.9,
    Veronica_canterburiensis:3.9,Veronica_cryptomorpha:3.9,
    Veronica_lycopodioides:3.9,Veronica_odora:3.9,Veronica_
    rakaiensi:3.9,(Veronica_decumbens:1.95,Veronica_pimeleoides
    :1.95,Veronica_pinguifolia:1.95):1.95)veronica:104.099998)
    :15.0):2.0,"
w[2]<-"((Discaria_toumatou:53.5,Rosa_rubiginosa:53.5)rosales
    :53.5,Leptospermum_scoparium:107.0):18.0):11.399994,(
    Chionochloa_pallens:28.0,Chionochloa_rubra:28.0, Chionochloa_
    sp:28.0)chionochloa:108.399994)euphyllophyte:11.4;"
```

This code makes a tree with class "phylo"

```
par(mfrow = c(1,1))
w.tree<-read.tree(text= w)
plot(w.tree, cex=1.0)
```

This code converts the tree into a "phylo4" class, this makes it easier to manipulate the tree

```
tree.phylo4 <- as(w.tree, "phylo4")
tree.phylo4 <- as(w.tree, "phylo4")
tipLabels(tree.phylo4)
```

In Excel, enter trait data and save it in .csv format. Each trait must be in single column, and the values of the traits must be in the same order that in the phylogenetic tree.

Species	Trait 1	Trait 2
Aciphylla_aurea	0.7	0.5
Brachyglottis_cassinioides	0.3	2.3
Helichrysum_selago	0.1	1.5
...	...	...

The following code reads the traits data base

```
traits<-read.csv("w_tree_data.csv", sep=",", header=T)
```

This code adds the trait data base to the tree

```
tree.traits<-phylo4d(tree.phylo4,traits, missing.data="warn")
summary(tree.traits)
```

This code reads the Trait1 and Trait2 values. Trait1 is in column 2 and Trait2 is in column 3 from the .csv file.

```
Trait1<-(traits[, 2])
Trait2<-(traits[, 3])
```

This function collapse and resolve multichotomes in phylogenetic trees, ape package need it for the analysis

```
w.tree2<-multi2di(w.tree)
```

Phylogenetic independent contrast analyses for Trait1 and Trait2. The regression should pass through the origin as recommended by Legendre & Desdevises (2009). This the reason on the "-1" in the lm line code

```
Trait1.PIC <- pic(Trait1, w.tree2)
Trait2.PIC <- pic(Trait2, w.tree2)
lm1<-lm(Trait2.PIC ~ Trait1.PIC - 1)
plot(Trait2.PIC ~ Trait1.PIC)
```

Complete summary of the PIC analysis. The second line returns the correlation coefficient between Tait1 and Trait2

```
summary(lm1)
summary(lm1)[[8]]
```

D

The contribution index (Moles et al. 2005) calculates the contribution of each divergence in a phylogeny to the overall variance in trait values across species.

Table 21: Contribution Index for plant life-history traits from the the 46 spp Molesworth assemblage, Chapter 3.

NODE NUMBER IN THE PHYLOGENETIC TREE (Figure 17)	ESTIMATED AGE	NO. TAXA	NO. NODES	CONTRIBUTION INDEX (TRAIT)					
				k from NBD2	Leaf area	SLA	Dispesule mass	Seed mass	Plant height
1	325.0	44	2	0.00	0.02	0.10	0.00	0.35	0.31
2	147.8	42	2	0.01	0.01	0.00	0.01	0.02	0.00
3	125.0	37	2	0.02	0.03	0.00	0.10	0.03	0.07
4	107.0	6	2	0.00	0.00	0.02	0.06	0.05	0.00
5	98.0	4	2	0.00	0.00	0.03	0.06	0.00	0.03
6	76.0	3	2	0.00	0.00	0.03	0.16	0.02	0.00
7	38.0	2	2	0.00	0.00	0.01	0.51	0.10	0.10
8	54.5	2	2	0.01	0.00	0.00	0.00	0.00	0.07
9	122.0	31	2	0.38	0.01	0.01	0.00	0.02	0.00
10	118.0	30	2	0.00	0.06	0.02	0.00	0.07	0.00
11	90.0	23	2	0.00	0.05	0.00	0.00	0.00	0.01
12	115.0	13	2	0.01	0.43	0.02	0.00	0.01	0.01
13	58.0	2	2	0.00	0.50	0.00	0.00	0.01	0.00
14	123.0	11	2	0.00	0.00	0.03	0.00	0.00	0.02
15	114.0	10	2	0.00	0.00	0.01	0.00	0.00	0.13
16	76.0	9	3	0.00	0.00	0.03	0.00	0.00	0.02
17	39.0	5	5	0.16	0.02	0.01	0.00	0.00	0.01
18	38.0	3	3	0.01	0.00	0.01	0.00	0.00	0.00
19	108.0	10	2	0.01	0.02	0.03	0.00	0.00	0.04
20	3.9	9	7	0.10	0.00	0.00	0.00	0.01	0.04
21	1.9	3	3	0.15	0.00	0.00	0.00	0.00	0.00
22	80.0	7	4	0.02	0.00	0.25	0.00	0.18	0.05
23	41.0	3	3	0.05	0.00	0.00	0.00	0.00	0.02
24	41.0	2	2	0.00	0.00	0.37	0.00	0.00	0.00
25	40.0	5	3	0.13	0.00	0.03	0.00	0.00	0.03
26	28.0	3	3	0.00	0.00	0.02	0.00	0.00	0.04
27	298.0	2	2	0.00	0.00	0.00	0.00	0.01	0.08

CALCULATIONS OF MASS FRACTAL DIMENSION (D_M) IN R

E.1 IMPORTING *.SHP FILES INTO R

ArcGis® 9.3 software was used to transform World Geodetic System (WGS84) coordinates to New Zealand map grid (NZMG) and then converted to a shape file (Figure 30). The shape files (*.shp) were imported to R (R Development Core Team 2011) where the maptools, spatstat (Baddeley & Turner 2005) and the sp (Bivand et al. 2008) packages were utilised to calculate Ripley's *K* statistic (Haase 1995). First the spatstat, maptools and sp libraries are required. Then, bring the window or area in where your point patters are, in this case New Zealand polygon was used as a window. Next assign this window to the point pattern, in this example the point patter of *Veronica stricta* was used and was imported from ArcGIS® 9.3 as a shape file (*.shp)

```
library(spatstat)
library(maptools)
require(sp)

#Brings the polygon to R
SW<-readShapePoly("nzcoast_window_Area.shp")
SW2<-as(SW, "SpatialPolygons")
W<-as(SW2, "owin")

#Assing a window (area) and units
S<-readShapePoints("Veronica_stricta.shp")
SP<-as(S, "SpatialPoints")
P<-as(SP,"ppp")
Veronica_stricta<-unique(P)
Veronica_stricta<-Veronica_stricta[W]
unitname(Veronica_stricta) <- c("metre", "metres")
plot(Veronica_stricta, main="Veronica stricta")
```

E.2 OMEGA FUNCTION

Departure from the Ripley's *K* statistic, aggregation indices of omega (Ω) and the mass fractal dimension (D_m) were computed. The aggregation index $\Omega_{x_1-x_2}$ captures the relative neighbourhood density in

ring-shaped neighbourhoods around the point pattern. To compute $\Omega_{x_1-x_2}$, for each point it is counted how many other points have a distance between x_1 and x_2 to that point. The obtained number is then set into relation with the number of points that could be expected in that distance to the point under the assumption that all points were distributed randomly across the study area. $\Omega_{x_1-x_2}$ values below 1 indicate that the points are less aggregated than random, values above 1 indicate that the points are more aggregated. The smallest possible $\Omega_{x_1-x_2}$ value is 0 which is obtained if no point is in the ring-neighbourhood of any other point. The maximum $\Omega_{x_1-x_2}$ value, for given x_1 and x_2 parameters, is reached if every point is in the neighbourhood of every other point (Condit et al. 2000). The $\Omega_{x_1-x_2}$ aggregation index is similar to Ripley's *K* statistic, with the difference that the *K* statistic use circles instead of rings as neighbourhoods, it has only one parameter which is analogous to the outer radius x_2 of the $\Omega_{x_1-x_2}$ index while the inner radius is set to 0. Also, Ripley's *K* is not normalised with respect to the size of the neighbourhood. Therefore, $\Omega_{x_1-x_2}$ allows investigations of the aggregation on larger distances without confounding effects of the intermediate neighbourhood. In the Chapter 4, an aggregation index closely related to Ω was used, the mass fractal dimension (D_m). The mass fractal dimension can be applied over a specified range of scales and could describe major aspects of the range structure over several scales. A detailed explanation of the mass fractal dimension is given in the next subsection.

The next two functions were calculated by Stephen Hartley.

The object `dist.breaks` is the distance classes from 2.5 to 100 km. Red lines are derived from Ripley's K with the trans edge correction and the blue lines are the expectation for a completely random spatial point process. `outer.dist` is the outer radius of the annulus (`inner.dist` is the value one row above, or for row 1 = the very first value of `dist.breaks`).

The output is four plots and a data frame: plot 1 (`cum.densT`) is the cumulative number of points expected around a randomly chosen point in the pattern between within a radius of `outer.dist` (assuming a trans edge correction); plot 2 is the log-log of plot 1. CSR has a slope of 2, and slope of the cum.dens line is the correlation or mass fractal dimension; plot 3 (`omega`) equals omega as described by Condit: the typical number of points within the annulus centred around a random point. Outer radius of annulus = `outer.dist`; Plot 4 (`pcf_annu`) is the pair correlation function: how much more dense the number of

point is within an annulus, relative to a random pattern (i.e. omega divided by the expectation for CSR , so that a value of 1 =CSR , > 1 implies clustering, < 1 implies evenness).

```
dist.breaks <- c(0,2500,5000,7500, 10000, 15000, 20000, 25000,
    30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000)
calc.omega = function(input.list, dist.breaks, ...){
        K <- Kest(input.list, r = dist.breaks)
        Cum.densT <- K$trans * summary(input.list)$intensity
        Cum.densP <- K$theo * summary(input.list)$intensity
        plot(K$r, Cum.densT)
        lines(K$r, Cum.densT, col = "red"); lines(K$r, Cum.densP,
            col = "blue")
        plot(K$r, Cum.densT, log = "xy", , xlab = "search radius
            ", ylab = "cumulative count", las = 1)
        lines(K$r, Cum.densT, col = "red"); lines(K$r, Cum.densP,
            col = "blue")
        no.dists <- length(Cum.densT)
        outer.dist <- K$r[2:no.dists]
        cum.dens <- Cum.densT[2:no.dists]
        cum.densP <- Cum.densP[2:no.dists]
        Omega <- rep(NA, no.dists -1)
        Omega <- Cum.densT[2:no.dists]-Cum.densT[1:(no.dists-1)]
        OmegaP <- Cum.densP[2:no.dists]-Cum.densP[1:(no.dists-1)]
        pcf_annu <- Omega/OmegaP
        plot(outer.dist, Omega, las = 1)
        lines(outer.dist, Omega, col = "red"); lines(outer.dist,
            OmegaP, col = "blue")
        plot(outer.dist, pcf_annu, las = 1)
        lines(outer.dist, pcf_annu, col = "red"); lines(outer.
            dist, rep(1, no.dists-1), col = "blue")
        output.df <- data.frame(outer.dist, cum.dens, cum.densP,
            Omega, OmegaP, pcf_annu)
        return(output.df)
        }
```

E.3   MASS FRACTAL DIMENSION

Fit linear models to determine the slopes (mass fractal dimension) of different segments of plot 2 of previously code . Regional-scale are the first six distance classes (2.5-20 km) and national-scale = mass fractal dimension from 20-100 km.

```
mass.frac = function(output.df, min = 0, max = 100000){
        output.lm <- lm(log(cum.dens) ~ log(outer.dist), data =
            output.df[output.df$outer.dist >= min & output.df$
            outer.dist <= max,])
        mf.output.df <- data.frame(output.lm$coeff[2], summary(
            output.lm)$r.squared)
        colnames(mf.output.df) <- c("mass.FD", "R.sqd")
```

```
        rownames(mf.output.df) <- paste(min, ":", max, sep = "")
        return(mf.output.df)
        }

par(mfrow = c(2,2), mar = c(4.1,4.1,1,1), mgp = c(2.5, 0.9, 0),
    family="serif",pch=16, cex=1.2)
```

E.4 EXAMPLE

The next example invokes the omega and the mass fractal dimension function previous explained for *Veronica stricta.* The object massVeronica_stricta. 100000$coeff[2] is the mass fractal dimension from 0-100000 metres

```
###---------- Veronica stricta -----------##
dist.breaks <- c(0, 2500,5000,7500, 10000, 15000, 20000, 25000,
    30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000)

calc.omega(Veronica_stricta, dist.breaks)

output.df <- calc.omega(input.list = Veronica_stricta, dist.
    breaks)
massVeronica_stricta<- lm(log(cum.dens) ~ log(outer.dist), data =
    output.df)
massVeronica_stricta.100000 <- lm(log(cum.dens) ~ log(outer.dist)
    , data = output.df[output.df$outer.dist < 100000.01,])
random.lm <- lm(log(cum.densP) ~ log(outer.dist), data = output.
    df)

massVeronica_stricta$coeff[2]
massVeronica_stricta.100000$coeff[2]
mass.frac(output.df)
mass.frac(output.df, min = 0, max = 100000)
mass.frac(output.df, max = 20000)
random.lm$coeff[2]
```