

# **THE SPATIAL ECOLOGY OF AN ENDEMIC DESERT SHRUB**

**Lauren Gough, BSc.**

**Thesis submitted to the University of Nottingham**

**for the degree of Doctor of Philosophy**

**May, 2010**

**University of Nottingham  
Hallward Library**

## ABSTRACT

---

Using spatial patterns to infer biotic and abiotic processes underlying plant population dynamics is an important technique in contemporary ecology, with particular utility when investigating arid shrub population dynamics, for which experimental and observational methodologies are rarely feasible. Using a novel one-class classification technique, the locations of over 17,000 *Spartocytisus supranubius* individuals were mapped from aerial imagery generating a spatially extensive (162 ha), yet accurate, dataset.

The recent rapid increase in studies using pattern–process inference has not been accompanied by a rigorous assessment of the behaviour of these techniques, nor an appraisal of their utility in addressing ecological research questions. The first part of the thesis addresses these concerns, investigating whether current methodologies are adequate to test hypotheses concerning spatial interactions. A literature review reveals a preponderance of studies of small, little-replicated plots. The results of the research raise concerns about the utility of spatial point pattern analyses as currently applied in the literature. To avoid inaccurate description of fine-scale spatial structures it is recommended that researchers increase plot replication. Furthermore, studies of spatial structure and population dynamics should account for spatial environmental gradients, whatever plot size is used. The second part of the thesis presents a rigorous investigation, incorporating *a priori* inference and the application of fine-scale spatial statistical and modelling techniques, of the biotic and abiotic mechanisms underlying the spatial structure and population dynamics of *S. supranubius*, a leguminous shrub species endemic to the Canary Islands. The spatial structure of *S. supranubius* populations is consistent with the operation of clonal reproduction and intra-specific competition. However, the results indicate that spatial environmental heterogeneity (from small to broad scales), in particular topography, can interact with biotic processes to generate quantitatively different *S. supranubius* patterns in different locations. Future research into the spatial and temporal dynamics of interactions between abiotic and biotic processes is recommended.



## ACKNOWLEDGEMENTS

---

This research was funded by a studentship from the School of Geography, University of Nottingham.

There are a number of people I would like to thank whose help has been invaluable to the completion of this thesis. First and foremost I would like to thank Dr Markus Eichhorn for his continual guidance and support. I would also like to thank my supervisors, Dr Richard Field, Dr Chris Lavers and Dr Doreen Boyd for all their valuable suggestions. Several people must be thanked for their assistance in the field. In no particular order these people are: Dr Richard Field, Alex Bjørn, Duncan Cheshire, Lea de Nascimento and Jose Marí Fernández-Palacios.

I would also like to thank a number of post-graduates from the School of Geography. Firstly, I would like to thank all residents of A55, both past and present. In particular, I need to thank Oli Dunnett and Chris Parker who have joined me in the day-to-day rituals of procrastination, desperation and caffeine addiction. I would also like to thank Bronwen Whitney, Suzanne McGowan, Jim Nixon, Iain ('Crob') Cross and Tom Wainwright. Cleverly disguised as trips to the pub, their moral support and encouragement has been instrumental in preserving my sanity.

My family and friends have provided much assistance over the last three and a half years. In particular I would like to thank my parents, John and Marilyn Gough, and my brother and sister, Dan and Claire, for their encouragement, proof-reading services, and, most of all, their patience. I promise I will get a 'real job' now! Finally, special thanks go to Alex for his help with field work and figure drawing (pages 40 and 221), but above all his continual support and encouragement, and for making me see that there is life after the PhD.

# CONTENTS

---

List of Figures.....	viii
List of Tables.....	xii
List of Plates.....	xiv
<b>CHAPTER 1: INTRODUCTION AND RESEARCH CONTEXT .....</b>	<b>1</b>
1.1 General introduction .....	1
1.2 Research aims and objectives.....	3
1.3 The research context.....	5
1.3.1 The processes driving arid plant population dynamics .....	5
1.3.2 Methods of investigating the dynamics of arid shrub populations.....	9
1.3.3 The importance of space and spatial pattern analysis.....	10
1.3.4 Review of the application of spatial point pattern analyses in the ecological literature .....	13
1.4 Detailed thesis structure .....	21
<b>CHAPTER 2: STUDY AREA AND FOCAL SPECIES.....</b>	<b>23</b>
2.1 The Canary Islands.....	23
2.1.1 Biogeography of the flora of Macaronesia and the Canary Islands .....	23
2.2 Tenerife.....	25
2.2.1 Biogeography of the flora of Tenerife.....	26
2.3 The Las Cañadas caldera .....	27
2.3.1 Geology of the Las Cañadas caldera.....	27
2.3.2 The five focal substrates .....	29
2.3.3 Climate of the Las Cañadas caldera .....	32
2.3.4 Flora of the Las Cañadas caldera .....	32
2.3.5 The Teide National Park.....	34
2.4 <i>Spartocytisus supranubius</i> .....	35
2.4.1 Location and protection.....	35
2.4.2 Physical description.....	35
2.4.3 <i>S. supranubius</i> ontogeny.....	39
<b>CHAPTER 3: DATA ANALYSIS METHODS.....</b>	<b>42</b>
3.1 Quantification of spatial patterns from the 'plant's-eye' .....	42
3.2 Distance-based methods for the quantification of spatial pattern .....	45
3.2.1 Definition of $K(r)$ , $L(r)$ and $g(r)$ .....	46

**CHAPTER 4: DATA COLLECTION: MAPPING THE SIZE AND LOCATION OF *S. SUPRANUBIUS*..... 55**

- 4.1 Ecological applications of remote sensing ..... 55
  - 4.1.1 Remote sensing and arid ecology ..... 57
- 4.2 Image classification ..... 60
  - 4.2.1 Support vector machine (SVM) classification..... 62
  - 4.2.2 One-class classification (OCC) by support vector data description (SVDD) 66
- 4.3 Creating the *S. supranubius* SVDD classifier ..... 68
  - 4.3.1 Stage one: collecting training and testing data ..... 70
  - 4.3.2 Stage two: classifier training and allocation ..... 78
  - 4.3.3 Stage three: accuracy assessment ..... 82
- 4.4 Classification of the *S. supranubius* imagery ..... 96
- 4.5 Conclusions ..... 106

**CHAPTER 5: THE EFFECT OF EXTENT ON PATTERN ANALYSES USING  $g(r)$  AND  $L(r)$  107**

- 5.1 Introduction ..... 107
  - 5.1.1 Aims and objectives ..... 109
- 5.2 Methods ..... 109
  - 5.2.1 Study area and data collection ..... 109
  - 5.2.2 Spatial point pattern analysis ..... 110
- 5.3 Results ..... 116
  - 5.3.1 Environmental heterogeneity ..... 116
  - 5.3.2 The effect of extent on the estimation of pattern trend ..... 118
  - 5.3.3 The effect of extent on the spatial consistency of pattern detection .... 122
  - 5.3.4 The effect of extent on pattern detection by Monte Carlo simulation envelopes ..... 131
- 5.4 Discussion ..... 134
  - 5.4.1 The effect of extent on quantitative pattern detection ..... 134
  - 5.4.2 The spatial consistency of quantitative pattern description at ecologically meaningful scales..... 136
  - 5.4.3 The effect of extent on pattern detection by Monte Carlo simulation envelopes ..... 141
- 5.5 Conclusions ..... 143

**CHAPTER 6: THE CONSEQUENCES OF POINT VERSUS REAL-SHAPE APPROXIMATION IN SPATIAL PATTERN ANALYSIS ..... 145**

6.1	Introduction .....	145
6.1.1	Previous studies .....	150
6.1.2	Aims and objectives .....	151
6.2	Methods .....	151
6.2.1	Data collection .....	151
6.2.2	Identifying the hard-core distance .....	152
6.2.3	The equivalence of $g_{rs}(r)$ and $g_{pp}(r)$ beyond the hard-core distance ....	154
6.2.4	Analyses .....	154
6.2.5	Analytical procedures .....	156
6.3	Results .....	158
6.3.1	The effect of data representation on the type of pattern detected .....	158
6.3.2	The effect of data representation on the magnitude and scale of the strongest aggregation.....	161
6.3.3	The effect of data representation on the interpretation of ecological processes .....	163
6.4	Discussion .....	165
6.4.1	The importance of free-space .....	165
6.4.2	Real-shape versus point approximation .....	166
6.4.3	The potential utility of real-shape analysis .....	171
6.5	Conclusions .....	173

## **CHAPTER 7: PATTERN AND PROCESS IN *S. SUPRANUBIUS* AND THE EFFECT OF HETEROGENEITY ..... 175**

7.1	Introduction .....	175
7.2	Aims and objectives.....	178
7.3	Methods .....	181
7.3.1	Data collection .....	181
7.3.2	Analyses .....	182
7.4	Results .....	192
7.4.1	Environmental heterogeneity .....	192
7.4.2	Size–abundance distribution .....	195
7.4.3	Structure of different size classes .....	200
7.5	Discussion .....	208
7.5.1	Environmental heterogeneity .....	208
7.5.2	Size–abundance distribution .....	209
7.5.3	Spatial pattern of different size classes.....	210
7.5.4	A signature of competitive thinning .....	215
7.6	Conclusions .....	216

**CHAPTER 8: SPATIAL VARIATION IN THE DENSITY AND LOCAL SPATIAL  
STRUCTURE OF *S. SUPRANUBIUS*: THE ROLE OF TOPOGRAPHY..... 217**

8.1	Introduction .....	217
8.1.1	A model of water redistribution on Substrate 3 .....	219
8.1.2	Aims and objectives .....	222
8.2	Methods .....	222
8.2.1	Data collection .....	222
8.2.2	Analyses .....	226
8.3	Results .....	232
8.3.1	Analysis 1a: the effect of ridges on the density of <i>S. supranubius</i> .....	232
8.3.2	Analysis 1b: the effect of slope on the density of <i>S. supranubius</i> .....	235
8.3.3	Analysis 1c: combined models .....	238
8.3.4	Analysis 2: the effect of ridge distribution and slope on the local spatial structure of <i>S. supranubius</i> .....	240
8.4	Discussion .....	244
8.4.1	The effect of topography on <i>S. supranubius</i> density .....	244
8.4.2	The effect of topography on <i>S. supranubius</i> spatial structure.....	246
8.5	Conclusions .....	249

**CHAPTER 9: DISCUSSION AND CONCLUSIONS ..... 251**

9.1	The utility of remote sensing in arid shrub ecology .....	253
9.2	Spatial point pattern analysis: reproofs and recommendations .....	255
9.2.1	Plot extent and replication .....	256
9.2.2	Assessing pattern using Monte Carlo simulation envelopes.....	258
9.2.3	Point process modelling .....	258
9.3	The biotic processes driving arid shrub population dynamics .....	259
9.3.1	Clonal reproduction .....	259
9.3.2	Intra-specific competition.....	261
9.3.3	The myth of regularity and a new method for detecting competition ...	263
9.4	Abiotic processes driving arid shrub population dynamics and their interaction with biotic processes .....	264
9.5	A critical evaluation of the thesis .....	267
9.5.1	The major assumptions .....	267
9.5.2	Other possible explanatory factors.....	270

**REFERENCES.....274**

**APPENDICES.....301**

Appendix A.....302

Appendix B.....321

Appendix C.....325

Appendix D.....326

Appendix E.....328

Appendix F.....329

## LIST OF FIGURES

Figure 1-1 The hypotheses tested in Chapters 5 – 8 and the associated methods. ....	4
Figure 1-2 Results of a review of 109 published articles using spatial pattern analyses ( $g(r)$ , $L(r)$ , $O(r)$ , or $K(r)$ ) to study the pattern of woody plants (trees and shrubs).. ....	17
Figure 1-3 The plot extent (geometric mean) used in 109 published articles using spatial pattern analyses ( $g(r)$ , $L(r)$ , $O(r)$ , or $K(r)$ ) to study the pattern of woody plants (trees and shrubs).....	19
Figure 2-1 The location of the Canary Islands.....	24
Figure 2-2 Geological map of the central Pico Teide – Pico Viejo (PT–PV) formation. ....	28
Figure 2-3 The five focal substrates used in this thesis.....	29
Figure 2-4 Aerial photographs of a section of Substrate 1 obtained in (a) 1954 and (b) 2007 .....	38
Figure 2-5 Diagrammatic representation of the ontogeny of <i>S. supranubius</i> .....	40
Figure 3-1 A limitation of Clark and Evans' (1954) index.....	45
Figure 3-2 The quantification of spatial pattern using $L(r)$ and $g(r)$ .....	49
Figure 3-3 Sample-size driven changes in the width of Monte Carlo simulation envelopes.....	52
Figure 3-4 Implementation of $g(r)$ and $L(r)$ in the grid-based software Programita. ....	54
Figure 4-1 Support vector machine (SVM) classification for linearly separable classes and non-linearly separable classes.....	64
Figure 4-2 Classifying non-linearly separable classes using kernels. ....	65
Figure 4-3 Classification using support vector data description (SVDD) and hyperspheres .....	67
Figure 4-4 The stages involved in the classification of imagery in this thesis .....	69
Figure 4-5 Illustration of the collection of target data ( <i>S. supranubius</i> ) .....	72
Figure 4-6 Intelligent training of outlier datasets for image classification .....	76
Figure 4-7 The 960 classifier models assessed. ....	81
Figure 4-8 Distribution of the target ( <i>S. supranubius</i> ) and outlier ( <i>A. viscosus</i> ) training data in (a) red–green, (b) red–blue and, (c) blue–green feature space.....	85
Figure 4-9 The change in overall accuracy (measured as a proportion) as the rejection error ( $C$ ) increases from 0.0001 to 0.1. The results are divided by kernel (Gaussian	

RBF, Exponential RBF and polynomial) and by classifier model (INCSVDD and SVDD) .....	87
Figure 4-10 The change in overall accuracy (measured as a proportion) as the rejection error ( $C$ ) increases from 0.0001 to 0.1. The results are divided by dataset (A, B, C and D) and by classifier model (INCSVDD and SVDD).....	88
Figure 4-11 The change in overall accuracy (measured as a proportion) as the parameter value ( $p$ ) increases from 1 to 10. The results are divided by kernel (Gaussian RBF, Exponential RBF and polynomial) and by classifier model (INCSVDD and SVDD). .....	90
Figure 4-12 The change in overall accuracy (measured as a proportion) as the parameter value ( $p$ ) increases from 1 to 10. The results are divided by dataset (A, B, C and D) and by classifier model (INCSVDD and SVDD) .....	91
Figure 4-13 Comparison of the classification of a 4 ha sample area on Substrate 1 using optimum Classifiers #4 and #6. ....	95
Figure 4-14 Classified imagery of <i>S. supranubius</i> on the five focal substrates .....	99
Figure 5-1 Calculation of the average magnitude and scale of the dominant pattern from several replicate plots. ....	114
Figure 5-2 Calculation of the number of scales at which $g(r)$ exceeds the 99% CSR simulation envelopes. ....	115
Figure 5-3 Comparing the heterogeneity on Substrate 2 and Substrate 4 as assessed by (a) the homogeneous $g(r)$ and, (b) the homogeneous $L(r)$ .....	117
Figure 5-4 Unweighted mean homogeneous $g(r)$ calculated from ten replicate plots at each of the six experimental extents on Substrate 2. ....	118
Figure 5-5 Unweighted mean inhomogeneous $g(r)$ calculated from ten replicate plots at each of the six experimental extents on Substrate 4 .....	119
Figure 5-6 Unweighted mean homogeneous $L(r)$ calculated from ten replicate plots at each of the six experimental extents on Substrate 2 .....	120
Figure 5-7 Unweighted mean inhomogeneous $L(r)$ calculated from ten replicate plots at each of the six experimental extents on Substrate 4 .....	121
Figure 5-8 The unreliability of $g(r)$ estimates calculated on Substrate 2 from ten replicate plots at six extents.....	123
Figure 5-9 The unreliability of $g(r)$ estimates calculated on Substrate 4 from ten replicate plots at six extents.....	124
Figure 5-10 The unreliability of $L(r)$ estimates calculated on Substrate 2 from ten replicate plots at six extents.....	126



Figure 5-11 The unreliability of $L(r)$ estimates calculated on Substrate 4 from ten replicate plots at six extents.....	127
Figure 5-12 The $L(r)$ functions produced in each of ten replicate plots of 1 ha and 6.25 ha on Substrate 2 and 4.....	128
Figure 5-13 Number of scales at which a significant pattern is detected by $g(r)$ and $L(r)$ on Substrate 4 at extents of 1 ha and greater, as measured by 99% Monte Carlo simulation envelopes.....	132
Figure 6-1 The effects of using 'real shape' analysis compared with point approximation.....	149
Figure 6-2 Calculation of the hard-core distance from 1) the maximum observed canopy diameter ( $\hat{HCD}$ ) and, 2) from the scale of convergence of the real-shape and point $g(r)$ ( $\hat{HCD}_g$ ).....	153
Figure 6-3 The calculation of plant–plant distance between one individual and a neighbouring individual using point analysis and real-shape analysis .....	154
Figure 6-4 The pair-correlation function (black lines) generated for the pattern of small, medium-sized and large individuals using point analysis, real-shape analysis and, free-space analysis. ....	160
Figure 6-5 Comparing the pattern detected in each size class when using real-shape, free-space and, point analysis. ....	164
Figure 6-6 Diagram explaining the potential use of real-shape analysis to examine clonal reproduction.....	171
Figure 7-1 Size abundance distributions as conceptualised by Muller-Landau et al. (2006) and with the addition of clonal reproduction. ....	185
Figure 7-2 Assessing the presence of environmental heterogeneity on the five substrates.....	194
Figure 7-3 The observed and best-fitting canopy area distribution of <i>S. supranubius</i> individuals on Substrates 1 to 5. ....	198
Figure 7-4 The spatial pattern of different size classes of <i>S. supranubius</i> on five substrates of contrasting spatial environmental heterogeneity.....	202
Figure 7-5 A signature of competitive thinning derived from the changes in the difference between the hard-core distance and the maximum scale of aggregation as cohorts age. ....	216
Figure 8-1 Conceptual model of the spatial variation in deep water availability on Substrate 3.....	221

Figure 8-2 The focal site for Chapter 8 .....	223
Figure 8-3 Raster image (resolution 1 m) showing the distance to the nearest ridge top (m). .....	224
Figure 8-4 Raster image showing the slope (degrees) of the terrain .....	225
Figure 8-5 The AIC of models using the binary ridge distribution covariates to explain the density of small, medium-sized and, large <i>S. supranubius</i> individuals.....	234
Figure 8-6 The AIC values of models using the binary slope covariates to explain the density of small, medium-sized and, large <i>S. supranubius</i> individuals.....	237
Figure 8-7 The AIC values of models containing the optimum Cartesian, ridge and slope effects (and combinations thereof) when fit to the distribution of small individuals, medium-sized individuals and large individuals.....	239
Figure 8-8 Relationship between the sum of the canopy areas of the focal shrub and the five nearest neighbours and the sum of the distances to the five nearest neighbours using quantile regression. ....	241
Figure 8-9 Spatial variation in the effect of shrub distribution on shrub size .....	243
Figure 9-1 Summary of the key findings of the research, methodological recommendations, wider implications and resulting research questions.....	252

LIST OF TABLES

---

Table 1-1 The variables reviewed in a review of articles using second-order spatial statistics to describe the univariate patterns of woody vegetation.....	15
Table 2-1 Summary information on the five focal substrates. ....	30
Table 3-1 Description of some fundamental spatial point pattern concepts. ....	44
Table 4-1 Summary statistics for the target ( <i>S. supranubius</i> ) and outlier (not <i>S. supranubius</i> ) spectral data.....	71
Tables 4-2 Comparison of substrate-specific spectral responses of <i>S. supranubius</i> on the (a) red, (b) green and, (c) blue wavebands respectively .....	75
Table 4-3 The spectral characteristics of the outlier training data in each of the four datasets.....	77
Table 4-4 The six optimum support vector data description (SVDD) classifier models. ....	92
Table 4-5 Comparison of six optimum support vector data description (SVDD) classifier models.....	94
Table 4-6 Summary information of the classified images .....	97
Table 4-7 Error matrices for the classification of <i>S. supranubius</i> individuals. ....	105
Table 5-1 Average number of shrubs per replicate plot at each extent.....	110
Table 5-2 The average magnitude and scale of the dominant pattern detected across ten replicates at each of six experimental extents using $g(r)$ .....	130
Table 5-3 The average magnitude and scale of the dominant pattern detected across ten replicates at each of six experimental extents using $L(r)$ .....	131
Table 5-4 Summary of the main findings in Chapter 5 .....	133
Table 6-1 Comparing the magnitude and scale of the strongest aggregation detected for small, medium-sized and large individuals when using point, real-shape, and free-space analysis.....	162
Table 7-1 Hypothesised processes influencing the spatial structure of <i>S. supranubius</i> and the related spatial pattern predictions. ....	179
Table 7-2 Dimensions of the sample windows used in Chapter 7 .....	182
Table 7-3 The number of <i>S. supranubius</i> individuals within each size class in each of the five sample windows used in Chapter 7 .....	188

Table 7-4 Size-structure statistics for the *S. supranubius* population on each of the five substrates..... 196

Table 7-5 Effect sizes of pair-wise comparisons of the mean *S. supranubius* canopy area on different substrates, measured using Cohen's *d* ..... 196

Table 7-6 Parameter estimates and Akaike Information Criterion (AIC) values for four distribution functions fit to the size–abundance distribution of canopy areas of all *S. supranubius* individuals on each of the five substrates ..... 199

Table 7-7 The hard-core distance, maximum magnitude of  $g(r)$  ( $g_{max}(r)$ ) and scale of  $g_{max}(r)$  for each *S. supranubius* size class on each of the five substrates. .... 201

Table 8-1 The four point process models fit to the observed point pattern of *S. supranubius*..... 227

Table 8-2 The AIC values for each of the four point process models described in Table 8-1 (using the ridge distribution covariate) when fitted to the pattern of *S. supranubius* individuals in each of the three size classes..... 232

Table 8-3 The AIC values for each of the four point process models described in Table 8-1 (using the slope covariate) when fitted to the point pattern of *S. supranubius* in each of the three size classes..... 235

LIST OF PLATES

---

Plate 2.1 The five focal substrates.....31

Plate 2.2 The ontogeny of *S. supranubius*.....41

### 1.1 GENERAL INTRODUCTION

Understanding the forces that generate spatial patterns in natural communities is one of the main goals of ecology (Levin, 1992; Tuda, 2007). Observational and experimental techniques provide some of the most direct ways to measure the presence, strength and influence of biotic and abiotic processes on population spatial structures. However, in some situations the observation and measurement of processes is not feasible. For example, the slow and event-driven demographics of arid shrubs typically operate over longer timescales than the duration of most experimental and observational studies (Wiegand and Jeltsch, 2000). This thesis focuses on an increasingly popular approach used to investigate the processes and dynamics structuring plant populations: spatial point pattern analysis.

The vegetation of arid and semi-arid systems is usually dominated by shrub and tree species. These species are important elements of the semi-arid landscape as they regulate many community and ecosystem processes. In contrast, perennial herbaceous plants are generally too sparse to determine ecosystem and landscape properties, and annuals are too temporally transient to have long-lasting effects (Whitford, 2002). Understanding the dynamics of dominant arid shrub species may be an important first step in understanding the dynamics of the ecosystem as a whole, yet little is known about their population processes and the factors underlying their dynamics (Jiménez-Lobato and Valverde, 2006; Kyncl et al., 2006). There is a long history of studies investigating the dynamics of arid shrubs at the patch-scale, initiated by the identification of spatially periodic arid vegetation patterns in the 1950s (MacFadyen, 1950; Clos-Arceduc, 1956; cited in Couteron, 2002). To obtain a greater understanding of arid shrub population dynamics, however, we need to investigate the biotic and abiotic processes operating at the scale of the individual. Pattern–process inference may allow information about the dynamics and long-term demographics of arid shrub populations to be extracted in situations where experimental and observational techniques are not feasible.

Pattern–process inference relies upon the spatio-temporal theory that the biotic and abiotic processes underlying a population's dynamics will give rise to a non-random pattern of individuals. Consequently, the pattern of individuals can be interpreted as a spatial signature of the processes structuring a population (Law et al., 2009). Despite the increasing popularity of the technique, it continues to be criticised (Mahdi and Law, 1987; Cale et al., 1989; Turner et al., 2001; Moravie and Robert, 2003). In many cases, efforts to deduce processes from patterns have been rejected because the inferential link between pattern and process is believed to be too weak (McIntire and Fajardo, 2009). It has, however, been proposed that these concerns are rooted in the analytical and methodological procedures used, rather than the biological justification of pattern–process inference (McIntire and Fajardo, 2009). Despite its increasing utilisation in the ecological literature, applications of pattern–process inference have not been accompanied by an appraisal of the methodological application of the technique, or its ability to address ecological research questions.

This thesis investigates the application of pattern–process inference in the contemporary ecological literature and asks whether the methodological approaches commonly used hinder our ability to infer processes from patterns. Subsequently, fine-scale spatial statistical and modelling techniques are used to investigate the spatial structure and the inferred population dynamics of *Spartocytisus supranubius*, a narrow-ranged endemic dominating the high-altitude desert of Tenerife.

## 1.2 RESEARCH AIMS AND OBJECTIVES

The overall aim of the research is to:

*Investigate the potential methodological constraints of spatial point pattern analysis, as currently applied in the literature, and how, with the support of remotely sensed data, their application could be improved to help understand the biotic and abiotic processes structuring populations of *Spartocytisus supranubius*.*

To address the central aim of the research, four hypotheses were developed to be tested in subsequent chapters. These hypotheses, the associated methods and the corresponding chapters in which the results are reported and discussed are outlined in Figure 1-1. The research context underlying the hypotheses is discussed in Section 1.3 which examines the current state of knowledge regarding the biotic and abiotic processes driving arid vegetation and shrub population dynamics, the discipline of spatial ecology, and the inference of process from pattern. The final part of this section reviews the methodological application of spatial point pattern analyses in the ecological literature. A detailed thesis structure is provided in Section 1.4.



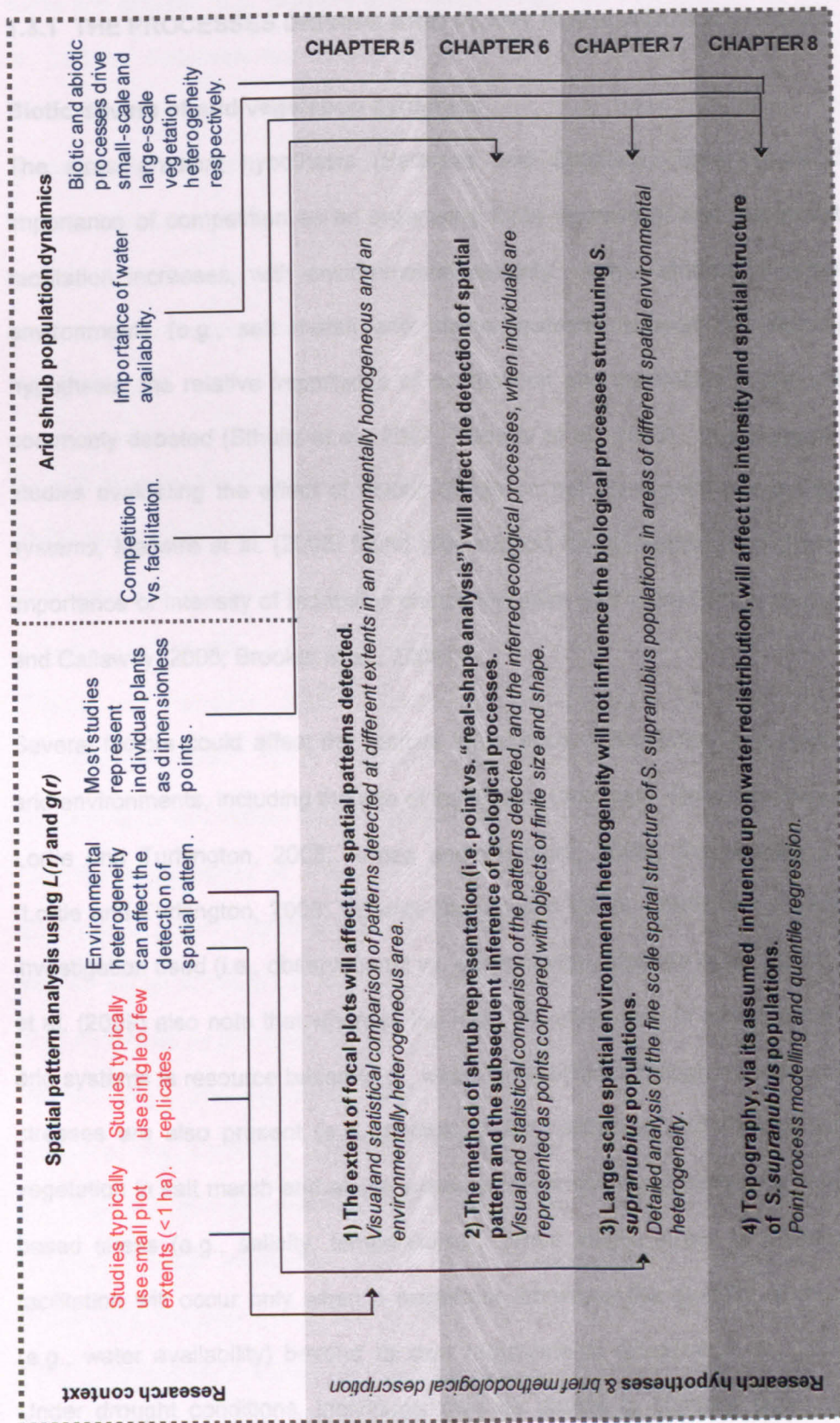


Figure 1-1 The hypotheses tested in Chapters 5 – 8 and a brief description of the associated methods. The research context corresponds to section 1.2. Blue text shows knowledge taken from the literature; red text shows new knowledge generated in this chapter (see section 1.2.5).

## **1.3 THE RESEARCH CONTEXT**

### **1.3.1 THE PROCESSES DRIVING ARID PLANT POPULATION DYNAMICS**

#### **Biotic drivers of arid vegetation dynamics**

The stress-gradient hypothesis (Bertness and Callaway, 1994) states that the importance of competition as an organising force decreases, and the importance of facilitation increases, with environmental severity. While studies in some severe environments (e.g., salt marsh and alpine systems) support the stress-gradient hypothesis, the relative importance of competition and facilitation in arid systems is commonly debated (Sthultz et al., 2007; Maestre et al., 2009). In a meta-analysis of studies evaluating the effect of abiotic stress on net plant–plant interactions in arid systems, Maestre et al. (2005) found little support for theoretical predictions that the importance or intensity of facilitation should increase with abiotic stress (but see Lortie and Callaway, 2006; Brooker et al., 2008).

Several factors could affect the relative importance of facilitation and competition in arid environments, including the age of individuals (ontogeny, Reisman-Berman, 2007; Lortie and Turkington, 2008; Armas and Pugnaire, 2009), the species concerned (Lortie and Turkington, 2008; Valiente-Banuet and Verdú, 2008) and the methods of investigation used (i.e., observational vs. experimental, Maestre et al., 2005). Maestre et al. (2009) also note that whereas the most important type of stress experienced in arid systems is resource based (e.g., water availability), although non-resource based stresses are also present (e.g., photoinhibition, Jefferson and Pennacchio, 2005), vegetation in salt marsh and alpine systems predominantly experiences non-resource based stress (e.g., salinity, temperature). When abiotic stress is resource-based, facilitation will occur only when a neighbour increases the quantity of the resource (e.g., water availability) beyond its own requirements (Maestre et al., 2005, 2009). Under drought conditions, individuals may be unable to increase water availability beyond their own requirements and competition for water may over-ride any amelioration of non-resource based stresses (Maestre and Cortina, 2004).

Almost all studies of biotic interactions in arid systems have focused on inter-specific interactions, typically between individuals of different functional types (trees, shrubs, herbs and grasses [e.g., Maestre et al., 2003; Armas and Pugnaire, 2005; Holzapfel et al., 2006]). Relatively few studies have considered the importance of competition and/or facilitation in structuring populations and how these interactions may vary spatially and temporally. This thesis seeks to address this shortcoming. Interestingly, a recent study of intra-specific processes, a simulation model by Malkinson and Jeltsch (2007), concluded that the dominant force structuring a shrub population in xeric sites was neither competition nor facilitation, but random mortality because of drought stress.

### **Abiotic drivers of arid vegetation dynamics**

Water availability is believed to be the most important abiotic factor driving biological processes in arid ecosystems. Precipitation in arid systems typically occurs in short pulses. Consequently, arid areas exhibit great temporal variability in water availability (Snyder and Tartowski, 2006). Introduced over 30 years ago, the pulse-reserve paradigm (Noy-Meir, 1973) assumes that the dynamics of arid vegetation are predominantly determined by their reaction to the environment, in particular the highly intermittent availability of water. In other words, rainfall events trigger a pulse of activity (e.g., growth) some of which is stored in a reserve, such as seeds. The pulse-reserve model continues to form the basis of contemporary arid vegetation research, often motivated by the need to understand the consequences of changing precipitation regimes (Ogle and Reynolds, 2004; Schwinning et al., 2004). Climate change scenarios are predicting significant alterations to the timing and magnitude of precipitation in arid and semi-arid areas (Robertson et al., 2009). Understanding how variation in water availability affects biological processes is essential if we are to predict and manage the effects of future climate change (Snyder and Tartowski, 2006). Recent studies have investigated (empirically and theoretically) the effects of variation in precipitation timing (Snyder et al., 2004; West et al., 2007), magnitude (Huxman et al., 2004) and frequency (Heisler-White et al., 2009) on plant physiological

activity. However, recent research is suggesting that arid vegetation is more responsive to soil water availability than precipitation *per se* (Robertson et al., 2009)

The translation of precipitation into soil moisture, and the resultant spatial and temporal heterogeneity in water availability, is complex (Loik et al., 2004). The main focus of studies to date has been on the role of vertical heterogeneity in soil-water availability (Loik et al., 2004; e.g., Rye et al., 2004, 2008). Whereas water in shallow soil layers is quickly lost to evaporation, water that infiltrates to deep soil layers is conserved for a longer time (Chesson et al., 2004). Because the root systems of woody plants typically penetrate to deeper soil layers than herbaceous species, a larger amount of deep water is believed to benefit shrubs and trees (the two-layer hypothesis - Walter, 1971). However, there are several situations where a relationship between woody biomass and deep soil moisture availability has not been found (Breshears et al., 2009). There is growing evidence that horizontal variation in soil moisture may be as substantial as vertical heterogeneity (Breshears et al., 2009), although this is rarely considered in models of arid ecosystem and vegetation dynamics (Loik et al., 2004).

To date, horizontal water heterogeneity has been most frequently documented at the scale of canopy patches (Loik et al., 2004; Breshears et al., 2009). However, horizontal water heterogeneity is also influenced by geomorphological and topographical characteristics (Loik et al., 2004; Zou et al., 2010). For example, Fravolini et al. (2005) found that the response of a woody legume to precipitation pulses of different magnitude was largely dependent upon soil texture (Fravolini et al., 2005). Monger and Bestelmeyer (2006) propose that arid vegetation dynamics are influenced by the 'soil-geomorphic template'. Their conceptual model describes the combined effects of the soil, topography and parent material on vegetation patterns and dynamics. All three factors are believed to influence water availability, primarily through their control on the water-holding capacity of the soil and the lateral redistribution of water. However, the potential effect of geomorphologically driven,

horizontal variation in soil water availability on the spatial structure and dynamics of arid shrub populations has been relatively understudied.

### **Understanding arid vegetation heterogeneity: the interaction of abiotic and biotic factors**

Understanding the independent and interactive effect of biotic and abiotic processes on the dynamics of populations remains a fundamental aim of ecology (Dahlgren and Ehrlén, 2009). Following a review of research into Monte Desert ecosystems, Bisigato et al. (2009) concluded that coarse-scale vegetation heterogeneity (i.e., at the landscape and community scale) was determined by abiotic factors, whereas biotic interactions determined fine-scale vegetation patterns (i.e., at the patch and intra-patch scale). Theoretical evidence, however, suggests that interactions and feedbacks between biotic and abiotic processes across a range of spatial and temporal scales may be important drivers of vegetation and population dynamics. Indeed, Agrawal et al. (2007) identified a need in population and community ecology studies to understand how biotic interactions vary with abiotic context, and to understand how biotic and abiotic factors interact over time and space. One conceptual model, the storage effect (Chesson, 2000a, b), proposes that the dynamics of arid vegetation communities are in part driven by an interaction between temporal environmental variation and biological processes, specifically competition. If competing species experience fitness advantages at different times, and are able to store the gains made using favourable periods, then coexistence will be enhanced (Adler et al., 2009). In order for the storage effect to operate, competition must vary with temporal environmental variation such that intra-specific competition is strongest and limits growth in favourable periods (Verhulst et al., 2008). Support for the covariance of competition and temporal environmental variation is, however, limited (Adler et al., 2009).

Research on how spatial environmental variation influences biotic interactions in arid vegetation has predominantly focused on comparing the spatial structure and dynamics of populations under different abiotic scenarios (Schenk et al., 2003;

Malkinson and Kadmon, 2007; Biganzoli et al., 2009). There is surprisingly little understanding of whether and how continuous spatial environmental heterogeneity interacts with biological processes to determine population dynamics (Wagner and Fortin, 2005; Murrell, 2009). Understanding the spatial dynamics of biotic interactions and their relationship with abiotic factors is important if we are to gain an understanding of the dynamics and processes that organise arid shrub populations.

### **1.3.2 METHODS OF INVESTIGATING THE DYNAMICS OF ARID SHRUB POPULATIONS**

The population dynamics of arid, perennial shrubs are slow, with infrequent establishment of new individuals, low growth rates and extended longevity (Cody, 2000; Bowers, 2005). Vegetative responses to abiotic and biotic pressures may operate over much longer time scales than variability in the pressures themselves. This results in a temporal mismatch between the typical duration of observational and experimental methodologies (years) and the time scales of vegetation change (decades) which has made it difficult to investigate the long-term dynamics of arid shrub populations (Wiegand and Jeltsch, 2000). One approach to investigating arid vegetation dynamics over medium to long time scales has been to mark and repeatedly sample individuals in permanent plots (e.g., Shreve and Hinckley, 1937; Goldberg and Turner, 1986; Turner, 1990; Tielbörger and Kadmon, 1997; Pierson and Turner, 1998; Cody, 2000; Bowers et al., 2004; Bowers, 2005; Kraaij and Milton, 2006). There is no doubt that long-term, systematic studies of arid vegetation have advanced our understanding of the processes influencing their population dynamics, especially in relation to grazing pressures (e.g., Ward et al., 2000; Angell and McClaran, 2001) and precipitation regimes (e.g., Milton and Dean, 2000; Ward et al., 2000; McClaran and Angell, 2006). However, most long-term observational studies are limited to monitoring a small subset of the target population from which generalisations about the demography of a species or population are made (e.g., Henschel and Seely, 2000).



Because of the time and financial limitations of long-term studies, most studies of arid vegetation dynamics are conducted over relatively short time periods (i.e., seasons or years). The slow dynamics of arid perennials means that many studies base their interpretations upon physiological responses at the individual level. Studies often compare the performance of target individuals growing in the vicinity of a neighbour with the performance of individuals growing in open areas (either naturally devoid of vegetation or where vegetation has been experimentally removed). Common measures of performance include growth (biomass, height, diameter or the number of leaves) and fecundity (number/weight of flowers/fruit/seeds; Maestre et al., 2005). However, the ability of individual-level responses to impart structure at the population level has been questioned (Freestone, 2006). To understand population dynamics, population-level responses should be measured (Goldberg et al., 1999). The discipline of spatial ecology may provide suitable techniques. In this approach the spatial structure of individuals within a population is used to infer its dynamics and the influence of abiotic and biotic factors. Coupled with the increasing availability of high-resolution remotely sensed data and Geographical Information Systems (GIS) software, spatial ecology may provide an important opportunity to investigate the population dynamics of arid shrubs over large spatial extents.

### **1.3.3 THE IMPORTANCE OF SPACE AND SPATIAL PATTERN ANALYSIS**

*[W]e must find ways to quantify patterns of variability in space and time, to understand how patterns change with scale..., and to understand the causes and consequences of pattern... (Levin, 1992 p. 1961)*

#### **Why space matters**

Over recent decades ecologists have become increasingly aware of the importance of the spatial dimensions of the phenomena they study. Organisms, both motile and sessile, are discrete entities that interact with their biotic and abiotic neighbourhood. Spatial confinement is strongest in sessile organisms such as terrestrial plants, marine macrophytes, corals and other species that are attached to surfaces (Tilman et al., 1997). Spatial ecology is a specialisation of geography and ecology that aims to

understand the spatial dimensions of the processes driving the dynamics and spatial structure of populations and communities (Murrell et al., 2001).

Some have attempted to investigate the spatial mechanisms underlying arid vegetation patterns with experimentation (e.g., Buonopane et al., 2005; Sthultz et al., 2007; Weedon and Facelli, 2008) or modelling biological and abiotic processes within a spatial domain (e.g., Meyer et al., 2007; Barbier et al., 2008; Moustakas et al., 2009; Popp et al., 2009). However, converting ecological theories into mathematical formulae can over-simplify the processes operating in natural communities, and experimental techniques can be time consuming, financially demanding and unethical (McIntire and Fajardo, 2009). Furthermore, depending upon the process being investigated, experimental techniques can be impractical. This is especially so when studying the dynamics of arid shrubs which operate over extremely long time scales. Therefore, this thesis is concerned with our ability to infer biological and abiotic processes from detailed analyses of observed spatial patterns.

### **Inferring process from pattern**

The theory of spatial point processes can be used to extract information from the spatial pattern of plants (Law et al., 2009). A spatial point process is a stochastic model that generates a set of countable points in a two-dimensional plane. The simplest point process is the Poisson point process which describes complete spatial randomness (i.e. a random number of individuals are located independently following a uniform distribution in region A [Law et al., 2009]). More complicated point processes introduce interactions between neighbouring points (e.g. Neyman-Scott processes).

The theory of spatial point processes can be applied to plant ecology by envisaging individual plants as points, with their locations represented by Cartesian coordinates. The abiotic and biotic processes driving a species' spatial structure (i.e., the spatial pattern presented by individuals) operate at discrete scales. These processes give rise to non-random patterns of individuals at the population level. Therefore, the pattern displayed by individuals can, with appropriate caution, be interpreted as a



spatial reflection of the mechanisms underlying population dynamics. Point process theory uses 'space as a surrogate' for unmeasured spatio-temporal processes (McIntire and Fajardo et al., 2009: 46).

The ability to infer biologically important process from observed spatial patterns was first recognised by Watt (1947). Nowadays there is a great impetus to study the spatial structure of plant populations and relate their characteristics to underlying biotic and abiotic processes. The spatial patterns of individuals within a population have been used to investigate and infer processes such as seed dispersal (Plotkin et al., 2002; Strand et al., 2007; Cousens et al., 2008; Wiegand et al., 2009), competition (Stoll and Bergius, 2005; Meyer et al., 2008), facilitation (Montesinos et al., 2007), herbivory (Zavala-Hurtado et al., 2000) and predation (Rossi et al., 2009).

Over recent decades, advances in computation have allowed ecologists to become increasingly sophisticated in their ability to quantify spatial patterns. However, despite the increasing ease with which patterns can be quantified, the inference of processes from observed patterns still remains theoretically challenging. Inferring processes from spatial patterns requires a substitution of space for time, and is consequently controversial. Many authors have questioned the extent to which processes can be reliably inferred from spatial patterns (Mahdi and Law, 1987; Cale et al., 1989; Moravie and Robert, 2003). In a study of the spatial organisation of limestone grassland species Mahdi and Law (1987) stated that:

*"...a spatial analysis of a plant community does not, on its own, give insights into the processes operating in a community." (Mahdi and Law, 1987:474)*

However, despite these uncertainties it can be argued that non-random processes will typically result in highly structured, distinctive patterns (McIntire and Fajardo, 2009). Biological organisation exists and, although the link between pattern and process may be imperfect, patterns of ecological phenomena continue to provide important opportunities for enhancing our understanding of population dynamics and spatial structure. Inferring processes will, however, be more difficult in complex communities

as many processes and factors will be operating simultaneously (Felinks and Wiegand, 2008). Studies of relatively simple systems, therefore, should improve our understanding of the theoretical association between pattern and process. In such systems, conceptual models of anticipated processes can be translated into expected spatial signatures. The relative abiotic and biotic simplicity of arid systems (Holzapfel and Mahall, 1999) aids the application and utility of spatial pattern analyses.

McIntire and Fajardo (2009) describe the inference of process from pattern as a multistage procedure requiring the precise implementation of ecological theory and knowledge, *a priori* inference of the anticipated processes and their spatial signatures, and the precise application of spatial analytical tools. Numerous authors repeatedly disapprove of the lack of a fourth stage: experimental verification of the operation of processes inferred from observed patterns (Steinberg and Kareiva, 1997; Murrell et al., 2001; Perry et al., 2006). However, because of the slow demographics and dynamics of arid shrubs and legislation protecting the focal species (*S. supranubius*; see Chapter 2) this fourth stage is not addressed in the current thesis. Chapters 5 and 6 of this thesis are specifically concerned with the third stage of process inference: the precise application of spatial analytical tools. Chapters 7 and 8 use all three stages in an investigation of the processes underlying the spatial structure of *S. supranubius* populations.

Methods of spatial pattern analysis are discussed in Chapter 3. Specifically, Chapter 3 provides a quantitative description of the pattern analysis techniques with which this thesis is concerned: the  $L(r)$ -function (a derivative of Ripley's  $K(r)$ -function) and the pair-correlation function ( $g(r)$ ). Although the details of these techniques are not discussed until Chapter 3, the following section reviews the application of these techniques in the ecological literature, in order to inform the research context.

### **1.3.4 REVIEW OF THE APPLICATION OF SPATIAL POINT PATTERN ANALYSES IN THE ECOLOGICAL LITERATURE**

Spatial pattern analysis is a subject of considerable current statistical research (Law et al., 2009). However, much of this work is technical and there is relatively little

discourse between mathematicians and ecologists (Law et al., 2009). Furthermore, with the ever-increasing power of desktop computers, the widening availability of GIS software, and the publication of independent spatial analysis programs over the internet, it seems that undertaking statistical analysis of spatial point patterns is no longer limited to either mathematicians or those who understand the techniques well enough to apply them with care. This has resulted in a rapid increase in the application of pattern analyses to plant distribution data over recent years. This increase has not, however, been paralleled by rigorous assessment of the behaviour of these techniques. This section reviews the ecological literature using second-order spatial statistics (i.e., statistics that quantify the pattern of points relative to one another, see Table 3-1) to describe the spatial patterns of woody plants (trees and shrubs), to identify where research is required. To maintain comparability with the analyses performed in this thesis, the review was limited to those studies using  $g(r)$  and  $L(r)$  (and the related functions  $K(r)$  and  $O(r)$ ). Only those studies applying spatial pattern analysis techniques to real (i.e., not simulated) data were included in the review. Methodological studies were not considered. While the review was not limited to studies of single species patterns, studies solely investigating the pattern of seedlings were not included, nor were articles that investigated the spatial patterns of herbs or grasses. Articles were selected by searching ISI Web of Knowledge using various combinations of the following search terms: 'spatial pattern', 'spatial point pattern analysis', 'ecology', 'Ripley's', 'pair correlation function', 'tree' and 'shrub'. No restriction was placed on the year of publication. The aspects considered in the review are detailed in Table 1-1. Some information could not be retrieved from some articles. A total of 109 articles were reviewed (Appendix A). The notable features of the review are discussed below.

Table 1-1 The variables reviewed in a review of articles using second-order spatial statistics to describe the univariate patterns of woody vegetation.

Variable	Description	
General	Focal species/community	The species/community that is the focus of the pattern analysis.
	Life form	Tree or shrub.
	Arid environment	Whether or not the research was conducted in an arid ecosystem.
Data collection	Data collection technique	Whether the data were collected manually or computationally (e.g., remote sensing).
	Data representation	Whether the size and shape of each individual is preserved or the position is reduced to a dimensionless point.
	Extent of plot	The extent of the plot(s) within which point data were collected. Both the range in plot extent and the geometric mean plot extent were calculated.
	Environmental contexts	The number of spatial environmental contexts studied.
Pattern analysis	Number of replicates	The number of replicate plots within each spatial environmental context (e.g., the number of replicate plots used per site when comparing the patterns observed in two sites of differing fire regime).
	Sample size	The number of points used to quantify univariate spatial pattern. Both the range in sample size and the geometric mean sample size per univariate analysis were calculated.
	Function	The function used to quantify spatial pattern (Ripley's $K(r)$ , the pair-correlation function ( $g(r)$ ) or one of their derivatives).
	Interpretative technique	How the presence of notable/significant spatial pattern was detected: comparison of function to Monte Carlo simulation envelopes or the height of the empirical function.

### ***Spatial pattern analysis is increasingly popular***

The review confirmed an increase in the ecological application of pattern analyses over the last 15 years (Figure 1-2a). There is a notable lack of articles published in 2001 and 2002. Interestingly, articles produced in these years are noted for attempting to develop new indices of spatial pattern (Dale and Powell, 2001; Fehmi and Bartolome, 2001), apply alternative pattern indices (Dovčiak et al., 2001; Cressie and Collins, 2001) or to deal with the problems facing point pattern analyses, such as heterogeneity (Pélissier and Goreaud, 2001) and missing data (Freeman and Ford, 2002). The majority of the articles investigated the spatial patterns of temperate or tropical forest trees with only 18 articles investigating species in arid or semi-arid systems (Appendix A).

### ***Cumulative measures favoured***

Despite the complexities of interpreting cumulative measures (see Chapter 3), the majority of articles (c. 76%) analysed spatial patterns using only  $L(r)$  or  $K(r)$  (Figure 1-2b). Of the 25 articles that used a discrete measure (i.e.,  $g(r)$  or  $O(r)$ ), nine used the techniques in combination with  $L(r)$ .

### ***Low number of replicates used***

The majority of articles ( $n = 75$ ) based their interpretations upon analyses performed in a single replicate plot (Figure 1-2c). Of these studies, 46% used a single plot of 1 ha or less in extent and 20% used a plot of  $\leq 0.25$  ha. Fifteen articles used two replicates per spatial environmental context. Seven of these studies used plots of  $\leq 0.33$  ha in extent. Nineteen articles used more than two replicate plots. In general these studies used the smallest plot extents with over half ( $n = 10$ ) of the studies using plots of  $\leq 0.25$  ha.

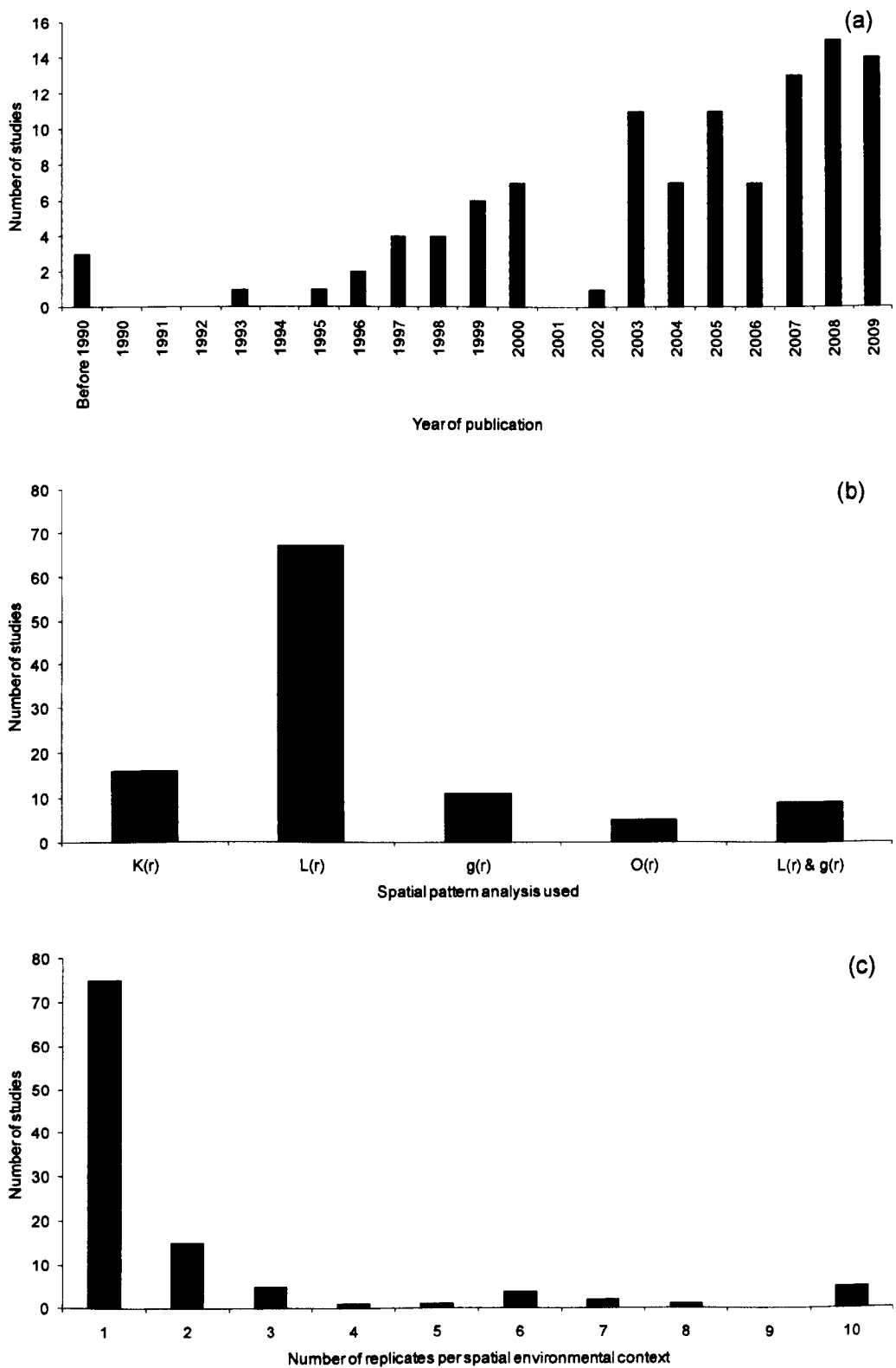


Figure 1-2 Results of a review of 109 published articles using spatial pattern analyses ( $g(r)$ ,  $L(r)$ ,  $O(r)$ , or  $K(r)$ ) to study the pattern of woody plants (trees and shrubs). Graphs show (a) the year of publication (2010 data not included [ $n=107$ ]), (b) the spatial pattern analysis technique(s) used, and (c) the number of replicate plots used per environmental context (e.g., the number of replicate plots used per site when comparing the patterns observed in two sites of differing fire regime). NB: in graph (b) the studies using both  $L(r)$  and  $g(r)$  (i.e., the final column) are not included in the count of studies using only one of the techniques (i.e., the 2<sup>nd</sup> and 3<sup>rd</sup> columns).

### ***Point analysis vs. real shape***

Only one study (Barbeito et al., 2008) represented the location of individual plants as objects as well as dimensionless points with co-ordinates x,y.

### ***No consensus on sample size requirements***

In standard statistical methods it is often accepted that a minimum sample size of 30 is adequate for simple comparisons and correlations. There is no such established consensus in spatial point analyses, with articles imposing independent, unfounded sample size restrictions. Malkinson et al. (2003), Meyer et al. (2008) and Linares-Palomino and Ponce-Alvarez (2009) restricted their analyses to plots containing 30 or more individuals, whereas Eccles et al. (1999), Mast and Veblen (1999) and Zhu et al. (2010) required 40 individuals and Wiegand et al. (2007a) required 70 individuals. Baddeley and Turner (2005) claim that  $K(r)$  is biased if fewer than 15 points are used (Rossi et al., 2009). Other studies have set limits of 10 individuals (Arévalo and Fernández-Palacios, 2003; Aldrich et al., 2003) and 20 individuals (Fulé and Covington, 1998), whereas Jacquemyn et al. (2009: p. 211) considered a sample size of 80 individuals to be 'relatively low'. However, the majority of studies do not impose a lower sample size limit, with some articles interpreting the spatial pattern detected from as few as six individuals (Fajardo et al., 2006). Precise data on sample sizes could only be extracted from 51 of the 109 studies reviewed. Of those studies, 34 performed one or more analyses on data with fewer than 70 individuals. Twenty-four of these articles performed one or more analyses on fewer than 30 individuals.

### ***Plots mostly < 1 ha in extent***

About 79% of the articles (for which information on plot extent was available,  $n = 106$ ) used plots of less than 5 ha in extent ( $n = 84$ ; Figure 1-3a). Over half of the articles ( $n = 56$ ) used plots of 1 ha or less in extent. Of those studies using plots of less than 1 ha, there was a general trend for the smallest extents to be the most popular (Figure 1-3b).

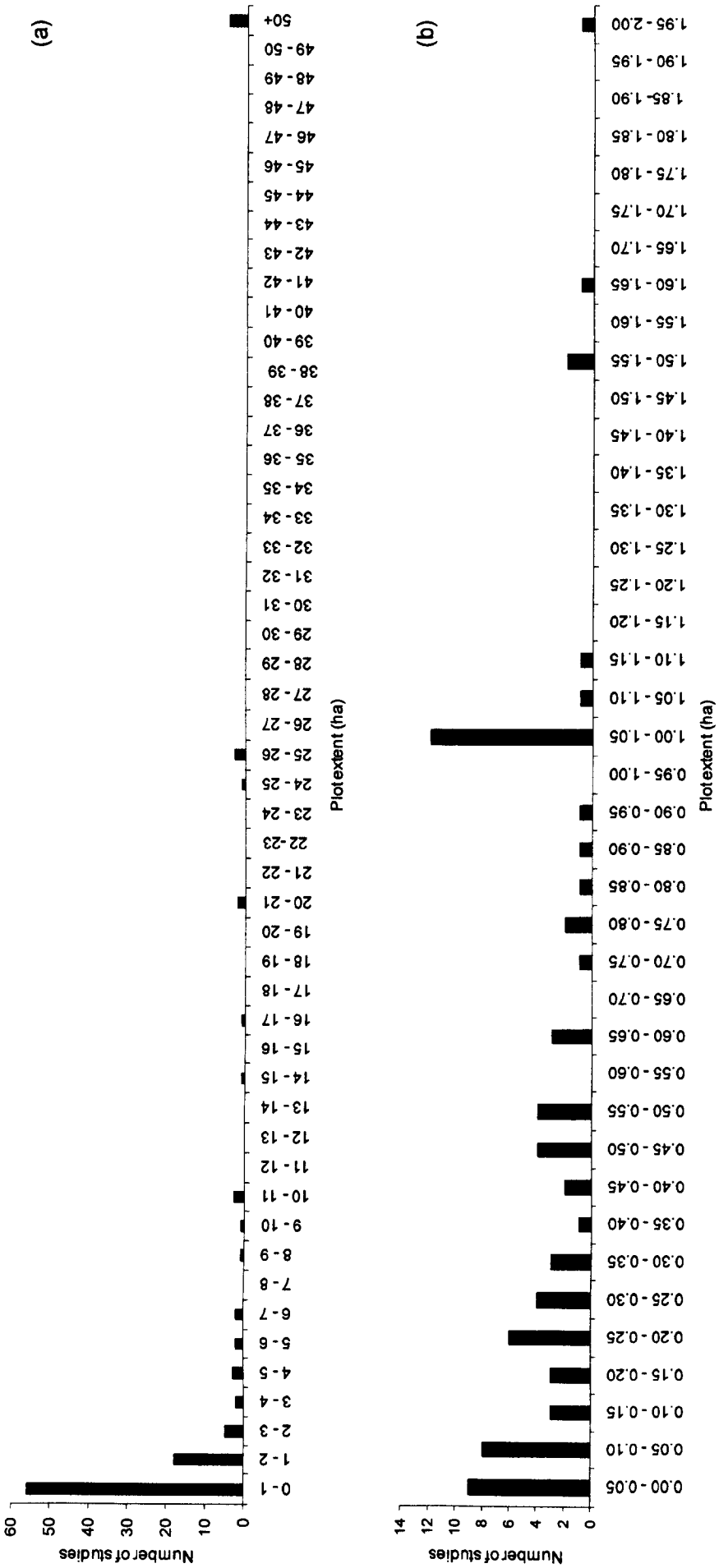


Figure 1-3 The plot extent (geometric mean) used in 109 published articles using spatial pattern analyses ( $g(r)$ ,  $L(r)$ ,  $O(r)$ , or  $K(r)$ ) to study the pattern of woody plants (trees and shrubs). Graph (a) shows all data whereas (b) is a subset of (a) and shows only those studies using plot extents of < 2 ha.



### ***Data collection***

Only eight (c. 7%) of the studies used computational techniques to generate data on the distribution of individual shrubs and trees. One study used ground-based, hemispherical photography (Montes et al., 2008), whereas the remaining seven studies used remotely sensed data. All seven studies used aerial photography, although Moustakas et al. (2008) and Koukoulas and Blackburn (2005) combined this with Ikonos imagery and LiDAR data respectively. Of the seven studies using airborne sensors, four were performed in arid areas.

### ***Interpretation***

The detection of significant spatial pattern from Monte Carlo simulation envelopes has been criticised because building envelopes from the result of many simulated patterns underestimates the Type I error rate (Loosmore and Ford, 2006, Chapter 3). Furthermore, the width of simulation envelopes is in part determined by sample size (Figure 3-3). Despite this, almost all of the articles ( $n = 104$ ) assessed spatial pattern by the scales at which the empirical function fell outside Monte Carlo simulation envelopes. Only eight studies (Peterson and Squiers, 1995; Pélissier, 1998; McDonald et al., 2003; Fang, 2005; Seidler and Plotkin, 2006; Getzin et al., 2008; Barbeito et al., 2009; LeMay et al., 2009) used the height of the empirical functions above the value expected from a completely spatially random pattern (CSR;  $L(r) = 0$ ,  $g(r) = 1$ ) to assess and compare patterns (see Table 3-1).

### ***Summarising remarks***

The results of the literature review presented above reveals a clear preponderance of methodological procedures that rely upon small plot extents with little replication. Very few studies use remote sensing technologies, which could enable the extent of the study area to be increased with relative ease. The majority of studies investigate spatial pattern using cumulative statistics ( $L(r)$  or  $K(r)$ ), and almost all assess the presence of pattern from Monte Carlo simulation envelopes.

## 1.4 DETAILED THESIS STRUCTURE

Chapter 2 introduces the study site and the focal species, *Spartocytisus supranubius*. Chapter 3 describes some of the methods available to quantify the pattern of individuals within a population. Particular attention is paid to the methods used in the subsequent results chapters (Chapters 5 – 8), the  $L(r)$ -function and the pair-correlation function ( $g(r)$ ). Specific methodologies and analytical techniques are discussed in further detail as appropriate in the subsequent chapters.

To address the stated research objectives, this thesis required the mapped location and size of individual *S. supranubius* shrubs across extensive areas. For this purpose data were collected via the classification of aerial photographs (Objective 1). Image classification was performed by one-class classification using support vector data description. Chapter 4 details the performance of 960 alternative classifier models. Following extensive cross-validation, a high-accuracy classifier is developed. This chapter details the application of this classifier to 162 ha of the study area for this research, and presents the resulting maps.

Chapter 5 investigates the effect of varying plot extent on the detection of spatial pattern using  $L(r)$  and  $g(r)$  (Objective 2). Extensive analyses at six different extents are performed to examine how changing plot extent influences the accuracy and reliability of pattern detection. Particular interest is directed towards pattern detection within small sample windows ( $< 1$  ha), which are commonly applied in the contemporary literature (see Section 1.3.4). The potential interaction between plot extent and spatial environmental heterogeneity is considered.

Chapter 6 investigates the differences in the pattern detected by analyses that approximate the location of individual shrubs as points, and those that preserve the size and shape of individuals (Objective 3). The types of pattern detected and the magnitude and scale of the strongest pattern are compared, and the consequences for the inference of ecological processes are considered.

Chapter 7 uses fine-scale, second-order spatial statistics to investigate the spatial structure of *S. supranubius* on five focal substrates (Objective 4). Deductive reasoning was adopted: *a priori* hypotheses of the likely abiotic and biotic processes driving the dynamics of *S. supranubius*, and their expected spatial signatures were formulated. These hypotheses are challenged with data. The spatial structure of *S. supranubius* on substrates experiencing different levels of spatial abiotic heterogeneity is compared to investigate whether environmental variation interacts with biological processes to determine *S. supranubius* dynamics.

Chapter 8 builds upon the findings of Chapter 7. Using data collected in the field, point process modelling techniques are used to investigate whether topography influences the spatial structure of an *S. supranubius* population (Objective 5). A conceptual model of topographically driven spatial variation in water availability is developed.

Chapter 9 summarises the principal conclusions and implications of the previous chapters, and provides recommendations for further research.

### **2.1 THE CANARY ISLANDS**

Situated in the northeast Atlantic Ocean, the Canary Islands extend for more than 500 km in a WSW-ENE orientation from Cape Juby on the African Coast (Kunkel, 1976). The archipelago comprises seven volcanic islands: Lanzarote, Fuerteventura, Gran Canaria, Tenerife, La Gomera, La Palma and El Hierro (Figure 2-1).

#### **2.1.1 BIOGEOGRAPHY OF THE FLORA OF MACARONESIA AND THE CANARY ISLANDS**

The Canaries, along with the other north-west Atlantic archipelagos – the Azores, Madeira, the Salvage Islands and the Cape Verde Islands – comprise the biogeographical region of Macaronesia. Macaronesia has been considered to be a phytogeographically distinct region for more than a century (Whittaker and Fernández-Palacios, 2007). However, there is a general floristic trend that parallels the considerable latitudinal range the archipelagos span. This has led some to question the validity of a distinct biogeographical Macaronesian region (Whittaker and Fernández-Palacios, 2007).

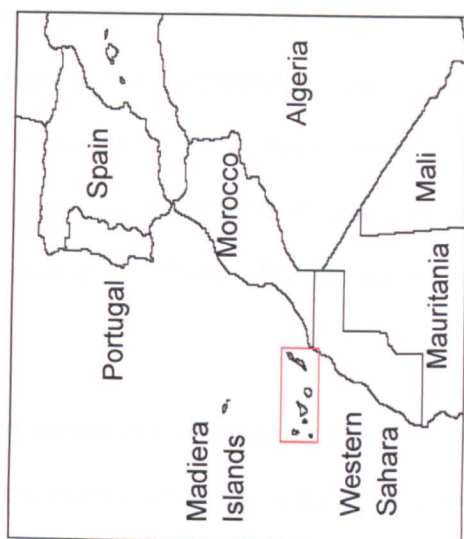
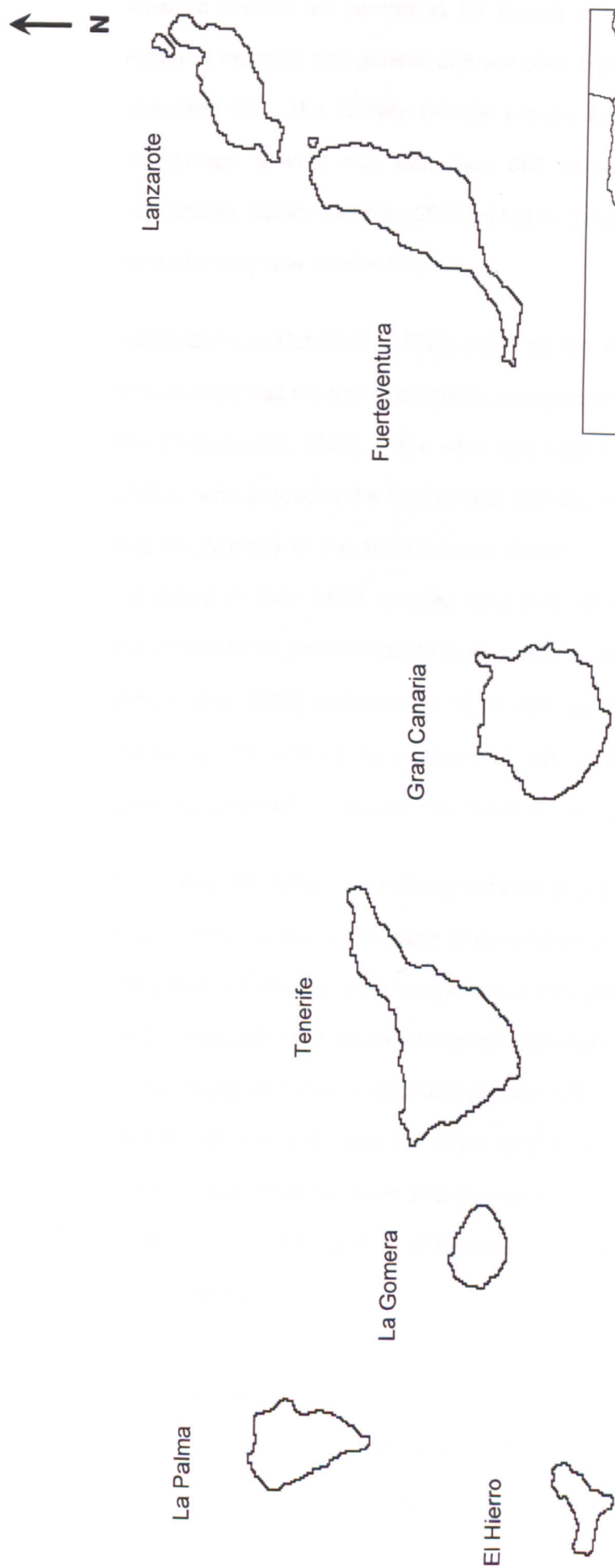


Figure 2-1 The location of the Canary Islands. Image adapted from Google Maps (<http://maps.google.co.uk>).

Oceanic islands are renowned for having a high proportion of endemic species. Adaptive radiation and genetic drift are often cited as the mechanisms underlying this phenomenon. The Canary Islands are no exception. The endemic flora of the archipelago is very rich with over 680 endemic taxa recognised (species and subspecies, Santos-Guerra, (2001), cited in Carine et al., 2009). This corresponds to an endemism rate of over 50%.

Alexander von Humboldt (1799), aided by the earlier manuscripts of Bonpland and Broussonet, was the first to attempt a description of the archipelago's flora (Ministry of the Environment, 2006). This work was later completed by Bory de Saint-Vincent (1802), who produced the first printed account of the flora and fauna of the Canary Islands (Ministry of the Environment, 2006). The floral richness of the islands is estimated at over 2000 species (Izquierdo et al., 2004). Generic endemism is common with 15 genera thought to exist only on the Canarian archipelago (Francisco-Ortega et al., 2009) and a further 12 genera specific to the Macaronesian region. The remaining non-endemic flora consists mostly of Mediterranean species as well as a large proportion of introduced, non-native species (Bramwell, 1976).

The Canary Islands have a strong network of protected areas (146 in total) covering around 40% of the archipelago (Reyes-Betancort et al., 2008). There are eight categories of protected area recognised in the Canary Islands (Reyes-Betancort et al., 2008). Populations of most endangered plant species can be found in these protected areas. National parks in particular contain >30% of the archipelago's endemic flora despite only covering about 4% of the land area (Marrero-Gómez et al., 2003). This thesis focuses on populations of *S. supranubius* in the Las Cañadas caldera, which is within the Teide National Park of Tenerife. The *Spartocytisus* genus is endemic to the Canary Islands.

## 2.2 TENERIFE

At 2,058 km<sup>2</sup> in area and reaching 3,718 m in altitude, Tenerife is the largest of the Canary Islands and the highest peak in Spain. Tenerife resulted from the fusion of

three palaeo-islands (Teno, Anaga and Adeje) following volcanic activity c. 3 Ma (Carracedo and Day, 2002). The landscape of Tenerife is dominated by the Cañadas volcanic series, rising to about 2,000 m, and the Pico del Teide (3,718 m.a.s.l.) and Pico Viejo (3,103 m.a.s.l.), which rise from the floor of the Las Cañadas caldera. The volcanic history of Tenerife prior to the formation of the Cañadas caldera (c. 0.17 Ma) is beyond the scope of this thesis: see Gill et al. (1994) and Guillou et al. (2004). In the following sections the floral biogeography of Tenerife is discussed in the context of the other Canary Islands before the focal site of this research – the Cañadas caldera – is introduced.

### **2.2.1 BIOGEOGRAPHY OF THE FLORA OF TENERIFE**

As mentioned above, the floral richness and endemism of the Canary Islands are very high. In a recent study, Reyes-Betancort et al. (2008) concluded that the endemic flora had a highly heterogeneous distribution both within and between the separate islands. They found that high rates of plant endemism occurred in the three palaeo-islands (Teno, Anaga and Adeje) which fused to form the island of Tenerife c. 3 Ma (Carracedo and Day, 2002). The Teno, Anaga and Adeje palaeo-islands had plant endemism rates of c. 23, 18 and 16% respectively (Reyes-Betancort et al., 2008).

Following Reyes-Betancort et al.'s study, Carine et al. (2009) attempted to delimit areas of endemism within the Canarian archipelago. Of the 17 areas of endemism they recognised, six occurred within Tenerife, more than in any of the other islands. The palaeo-islands of Anaga and Teno were both resolved as areas of endemism in their own right. According to Carine et al.'s (2009) study, the area of highest endemism was the high mountain area (i.e., Las Cañadas), together with the Güimar valley and the Tamadaya ravines of the south. This area also includes the third palaeo-island, Adeje. Thus the geological history of the island appears to contribute highly to its endemism. However, studies by both Reyes-Betancort et al. (2008) and Carine et al. (2009) uncovered a more contemporary gradient between the high species richness and high endemism of the north side of the island and the

comparatively species poor south of the island. This is believed to be driven by current climatic differences.

## **2.3 THE LAS CAÑADAS CALDERA**

### **2.3.1 GEOLOGY OF THE LAS CAÑADAS CALDERA**

The volcanic nature of the Las Cañadas caldera has created a mosaic of lava flows of differing ages and morphology (both pahoehoe and aa) and pyroclastic sediments (Figure 2-2). The Las Cañadas caldera is an elliptical depression measuring 16 km at its widest axis and 9 km at its smallest with a total perimeter of 45 km (Martí and Gudmundsson, 2000). Truncating the Las Cañadas edifice at an altitude of 2000 to 2200 m.a.s.l. (Galindo et al., 2005), the caldera is delimited by the Circo de las Cañadas, an elliptical wall up to 500 m high that encompasses all but the northern flank of el Teide. From the floor of the caldera rise the Pico Viejo (3,103 m.a.s.l.) and the Pico del Teide (3,718 m.a.s.l.), eruptions from which have covered much of the caldera floor with a mosaic of lava flows (both aa and pahoehoe) and pyroclastic sediments. Construction of the Teide–Pico Viejo complex began c. 2 Ma (Edgar et al., 2007). Both of the cones, and the smaller eruptive centres on their flanks (e.g., Montaña Blanca, Montaña Rajada), remain active to this day.





2.3.2 THE FIVE FOCAL SUBSTRATES

The focal site of this research contains five different substrates (Figure 2-3). These were selected to incorporate variation in the age and formation of the substrate while maintaining spatial proximity.

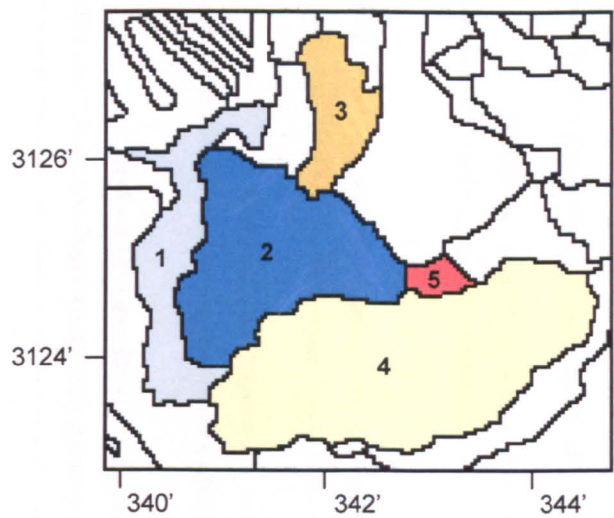


Figure 2-3 The five focal substrates used in this thesis (see Figure 2-3 and Table 2-1). The colours correspond to the colours used in Figure 2-2. Substrates that are not considered in this thesis have been left unshaded (white). Map co-ordinates in UTM ('000).

Further information and photographs of the five focal substrates are provided in Table 2-1 and Plate 2-1 respectively.

Table 2-1 Summary information on the five focal substrates.

Substrate number	Name of substrate	Formation	Age (relative)	Characteristics	Environmentally homogeneous /heterogeneous
1	Parador colluvium	Colluvium (erosional deposits)	-	Composed of erosional deposits from surrounding substrates and pumice. Some stones/rocks on surface. Topographically very flat.	Homogeneous
2	Montaña Majua	Aa lava flow	2	Lumps of solid lava protrude from mostly flat surface composed of some pumice, erosional material and eroded lava fragments.	Homogeneous
3	Arenas Blancas	Pahoehoe lava flow	1	Prominent ridge-trough topography. Intervening ridges predominantly filled with pumice and topographically flat. Ridges composed of solid lava and large rocks.	Heterogeneous
4	Montaña de la Cruz	Pahoehoe lava flow	3	Ridge-trough topography, not as prominent as on Substrate 3. Intervening ridges covered in eroded lava fragments (up to c. 30 cm diameter).	Heterogeneous
5	Montaña los Conejos	Aa lava flow	4	Lumps of solid lava protrude from flat surface composed of some pumice, erosional material and some eroded lava fragments. Surface sediments finer than on Substrate 2.	Homogeneous

Relative age is in order of 1 (youngest) to 4 (oldest). No relative age is given for Substrate 1 as it is composed of erosional deposits unlike the other four substrates which are lava flows. The calculation of environmental heterogeneity is discussed in greater detail in Chapter 7 (see Section 7.3.2).



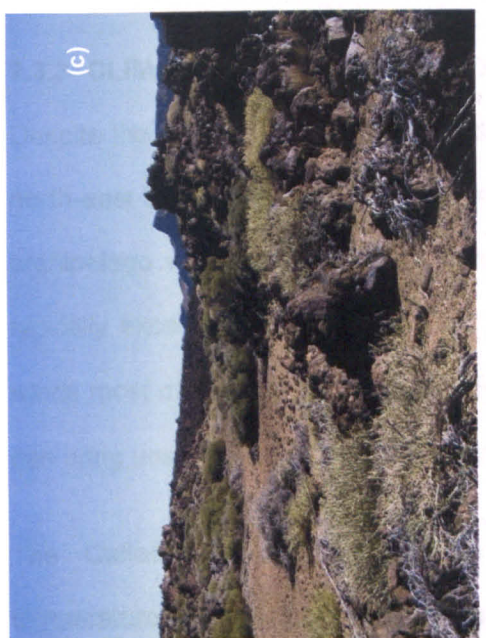
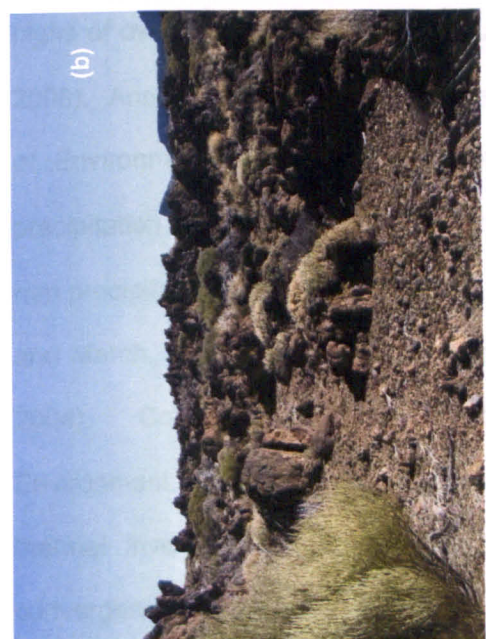


Plate 2-1 Photographs of the five focal substrates.  
Substrates 1-5 are shown in images (a) to (e) respectively.

### **2.3.3 CLIMATE OF THE LAS CAÑADAS CALDERA**

Despite the proximity of the Canary Islands to the African Sahara, the passage of the north-east trade winds over cold oceanic upwellings off the African coast provides the archipelago with a mild climate (Fernandopullé, 1976). Temperatures at sea level typically exceed 20°C throughout the year. Atmospheric stability produces constant winds most of the year, although the cyclonic influence of the Atlantic weather system can bring unstable conditions in the winter months (Fernandopullé, 1976).

The Cañadas caldera experiences large daily and seasonal oscillations in temperature. Diurnal temperature ranges of 15°C, winter lows of -16°C and summer highs of over 30°C are common (Seguela and Trujillo, 2004; Ministry of Environment, 2006). Annual insolation is high with around 3500 hours of sunlight per year (Ministry of Environment, 2006). The caldera is dry for 90% of the year with annual precipitation typically less than 300 mm (most consider a desert to have less than 250 mm precipitation a year [Ward, 2009]). Precipitation is concentrated between October and March, with the majority falling in December and January (Seguela and Trujillo, 2004). Consequently, much of the precipitation falls as snow (Ministry of Environment, 2006). The high mountain climate of Tenerife is controlled by a constant thermal inversion that occurs between 1,500 m and 2,000 m, because of the convergence of the cool and humid north-east trade winds and the hot and dry north-west winds (from the African continent). The thermal inversion and insular orography isolate the caldera from marine influences, producing climatic conditions more similar to continental areas than witnessed elsewhere on the archipelago.

### **2.3.4 FLORA OF THE LAS CAÑADAS CALDERA**

Sventenius (1946) was one of the first to attempt to catalogue the caldera's flora, gathering specimens until the 1960s (cited in Dickson et al., 1987). Estimates of species richness in the Cañadas caldera vary. In 1946, Sventenius published 76 vascular plant species comprising the flora of Las Cañadas. In 1980, Kunkel listed 94 taxa for the same area (cited in Dickson et al., 1987), whereas Dickson et al. (1987)

estimates the vascular plant richness of the Cañadas caldera at 125 species. Richness estimates are higher when the entire Teide National Park is considered with de la Torre and Osorio (2004) reporting the vascular floral diversity of the park at 168 taxa. Dickson et al. (1987) believe that at least half of the vascular plant species within the caldera can be considered aliens, most thought to have been introduced by tourism and pre-1950 pastoralism. Prior to the creation of the Teide National Park, the caldera's flora was affected by the grazing of goat herds, which migrated to the area in the spring and summer months. In the prologue of Sventenius' 1946 work, Jorge Menéndez provides the following description of the condition of the caldera's flora:

*"Debido al aislamiento de tales parajes y a la incultura de los cabreros y leñadores que a ellos acuden, se encuentra gravemente amenazada de extinción, todo esta interesantísima y bella vegetación, habiendo algunas especies descritas en épocas anteriores, que ya no es posible encontrar hoy y otras muchas en la que los ejemplares que existen son tan contados que hacen prever su próxima desaparición si no se toman urgentes medidas para su defensa"* (quoted from de la Torre and Osorio, 2004).

*(Because of the isolation of [the caldera] and ignorance of the goat-herders and woodcutters, the [flora] is seriously threatened with extinction. All this interesting and beautiful vegetation, some species of which are already extinct and others with seriously low abundances, may disappear if urgent measures for their protection are not taken)*

The Las Cañadas caldera is renowned for its distinctive flora comprising many endemic species (Bramwell and Bramwell, 2001). Many of these endemics are also strictly confined to the high altitude Cañadas area (Reyes-Betancort et al., 2008). The high proportion of endemism can be explained by both the physical isolation of Tenerife from continental landmasses, and the ecological insularity of the caldera resulting from its altitude and climate. However, endemic annuals are absent and herbs are few (Dickson et al., 1987). The two most common species (*Spartocytisus supranubius* (L.f.) Christ ex Kunk and *Adenocarpus viscosus* (Wild.) Webb and

Berthel), are both in the Fabaceae. The vegetative community is typically open, with limited vertical stratification; herbs and low shrubs tend to occur in the open spaces between taller shrubs rather than beneath their canopies (Lausi and Nimis, 1986; pers. obs.).

Today, the main threat to the caldera's flora comes from rabbits (*Orctolagus cuniculus*) and the Corsican mouflon (*Ovis gmelini musimon*) (de la Torre and Osorio, 2004). While the rabbit was introduced during the colonisation of the islands, the mouflon was deliberately introduced to the caldera in 1971 as game for hunters. The mouflon population has increased ever since. Attempts are now being made to control both herbivores by excluding them from ecologically fragile areas, in addition to long-term plans to reduce both populations to ecologically benign levels, although these plans are meeting some resistance from local hunters (De Nascimento, pers. comm.).

### **2.3.5 THE TEIDE NATIONAL PARK**

The Cañadas caldera is situated within the Teide National Park (TNP). Created by decree on the 22<sup>nd</sup> of January 1954, the TNP is the most visited protected natural area in Spain, receiving an average of 3 million visitors per year since 1996. In 1981 the TNP was reclassified under Law 5/1981 to include a peripheral protection zone (7,515 ha) to prevent external impacts on the ecology and landscape of the park. The TNP covers almost 19,000 ha, extending from 1,650 m.a.s.l. at its lowest point to 3,718 m.a.s.l. at the top of Teide. Surrounding the TNP is a buffer zone which includes the Corona National Park. In total, the Teide National Park and the surrounding buffer zone (consisting of the Corona Forestal Natural Park and the Peripheral Protection Zone) cover in excess of 54,000 ha (Ministry of the Environment, 2006).

All resource use has now become tightly controlled, regulated by the Management and Usage Administration Plan (Decree 153/2002). The greatest visual and ecological improvement has been gained by the prevention of goat herding. Certain

areas of the park are, however, still harvested for their resources but only traditional practices are authorised and are both spatially and temporally limited. These activities include the collection of coloured soil and flowers for decorative use in the celebration of the Octava del Corpus Christi, the operation of twenty apiaries in the spring and the collection of firewood and culm from the area of Llano La Rose (Montaña Limón). These practices are not believed to impact on the dynamics of the flora or fauna. Following the designation of the site as a World Heritage Site in June 2007, all terrain within both the Park and buffer zone has been classified as non-buildable land under special protection, except when it can be certified that building is in the public interest (Ministry of the Environment, 2006).

## **2.4 SPARTOCYTISUS SUPRANUBIUS**

### **2.4.1 LOCATION AND PROTECTION**

*Spartocytisus supranubius* (L.f.) Christ ex Kunk. (Fabaceae; hereafter *S. supranubius*) is endemic to the high altitude communities of Tenerife (> 1,900 m.a.s.l.) and La Palma (> 1,700 m.a.s.l.) in the Canary Islands (Bramwell and Bramwell, 2001). *S. supranubius* is protected under regional legislation (Annex II of the Flora Order 20/02/1991). Annex II of the Flora Order requires that governmental authorisation is sought before any *S. supranubius* individual, or any part of it, is up-rooted or cut down, deliberately disturbed or destroyed (including their seeds), or used for commercial purposes.

### **2.4.2 PHYSICAL DESCRIPTION**

*S. supranubius* is more commonly known as 'retama blanca' (white broom) or 'retama del Teide' (Teide broom). Individuals usually take a hemispherical shape, and can grow to a height of 3 m. The stems are thick and glaucous with a greyish-hue. The leaves of *S. spartocytisus* are small, deciduous, ephemeral, nearly sessile and trifoliate. The individual leaflets are pale green and linear, measuring less than 5 mm in length. *S. supranubius* blooms between May and July producing highly aromatic



dense racemes at the upper parts of its branches. The individual flowers are bilabiate, have short pedicels and a very short-toothed calyx with a white or pink corolla. The fruit of this shrub consists of black, 4–6-seeded, villous legumes. The seeds are dormant and may remain viable for a long time in the soil. Mechanical abrasion is thought to be necessary for germination (Kyncl et al., 2006).

Both isozyme and morphological analyses demonstrate that as well as reproducing sexually, *S. supranubius* is capable of clonal growth (Kyncl et al., 2006). *S. supranubius* reproduces asexually by the rooting of lateral branches (branch layering), unlike related species which sprout from roots (Kyncl et al., 2006; personal observations). Kyncl et al. also suggest that the successful establishment of both sexually and clonally produced juveniles, and the growth of adult individuals, is controlled by winter precipitation levels. Kyncl et al. (2006) indicate that mechanical abrasion by water may be required to break *S. supranubius* seed dormancy. They also indicate that whereas 1978–1980 had above average winter precipitation levels, the decade 1990–1999 was the driest ever recorded. The longevity of *S. supranubius* individuals is not known. However, by examining aerial photographs from 2007 and 1954 it is clear that individuals are visible on both images suggesting that *S. supranubius* individuals can live for at least 50 years (Figure 2-4).

*Spartocytisus supranubius* roots are infected by the endo-symbiotic bacteria *Bradyrhizobium canariense* bv. *genistearum* (Vinuesa et al., 2005a). This relationship enables *S. supranubius* to fix atmospheric nitrogen, sometimes creating a 16-fold increase in the nitrogen content of the soil beneath *S. supranubius* canopies compared with soils 5 m away (Wheeler and Dickson, 1990). Wheeler and Dickson (1990) cited this attribute as a potential reason for the dominance of the shrub within parts of the caldera. A previous study (Jarabo-Lorenzo et al., 2000) has shown that endemic canarian genistoid legumes are nodulated exclusively by *Bradyrhizobium* species. *B. canariense* has a known geographic distribution covering the Canary Islands, Spain, Morocco and the Americas (Vineusa et al., 2005a). Relatively little is known about *B. canariense*, but as with all rhizobia it is expected to live

saprophytically in the soil when not infecting root hairs (Salisbury and Ross, 1992), is acid-tolerant and is thought to experience optimum growth at temperatures of 28-30°C, but be inhibited at 37°C and above (Vinuesa et al., 2005b).

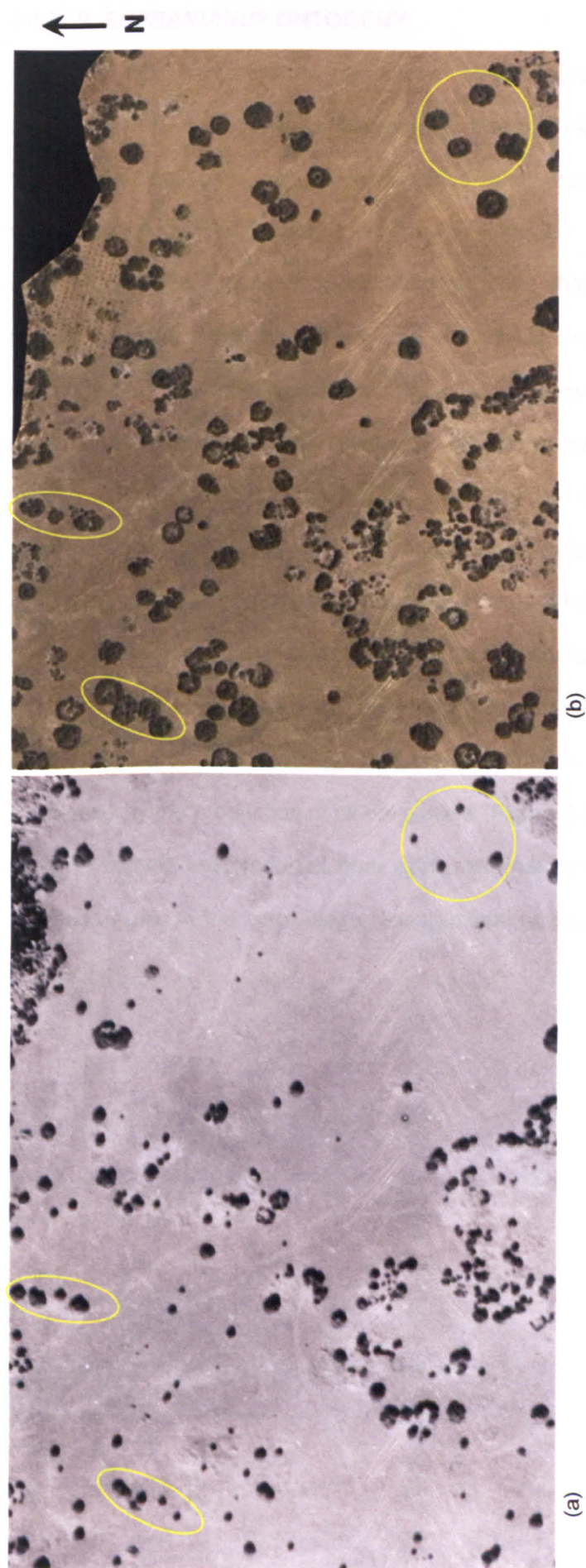


Figure 2-4 Aerial photographs of a section of Substrate 1 obtained in (a) 1954 and (b) 2007. Despite the time lapse between these images, patches of *S. supranubius* individuals are easily distinguished as being the same individuals in both images (examples highlighted in yellow). Image source: (a) provided by José María Fernández-Palacios; (b) Grafcan Plc.

### **2.4.3 S. SUPRANUBIUS ONTOGENY**

Following field observations, a seven stage morphological ontogeny for *S. supranubius* is hypothesised comprising the following stages: seedling, cone, hemisphere, collapse, ring, outliers, dead and young clone (Figure 2-5). A corresponding plate of photos is provided in Plate 2.1. Following reproduction from seed (Figure 2-5a), young *S. supranubius* are cone shaped (Figure 2-5b). Over time the shrub expands both horizontally and vertically to become hemispherical in shape (Figure 2-5c). Eventually most shrubs collapse outwards (Figure 2-5d), perhaps because of the increased strain placed on the central stem as the shrub continues to expand. Total collapse is typified by the formation of a complete or partial ring structure (Figure 2-5e) whereby the main stems run horizontal to the ground producing upright vegetation some distance from the centre. The collapse stage is critical for clonal reproduction as it positions the branches in contact with the ground. Once an individual has collapsed and formed a ring it is unclear how long it takes for clonal outliers to be produced. It is apparent, however, that in many cases the ring stage is succeeded by the production of clonal outliers (Figure 2-5f). In most cases only a few daughter ramets are produced from each maternal shrub. Although often no taller than individuals in the cone stage, young clonal ramets are hemispherical in shape (Figure 2-5g).

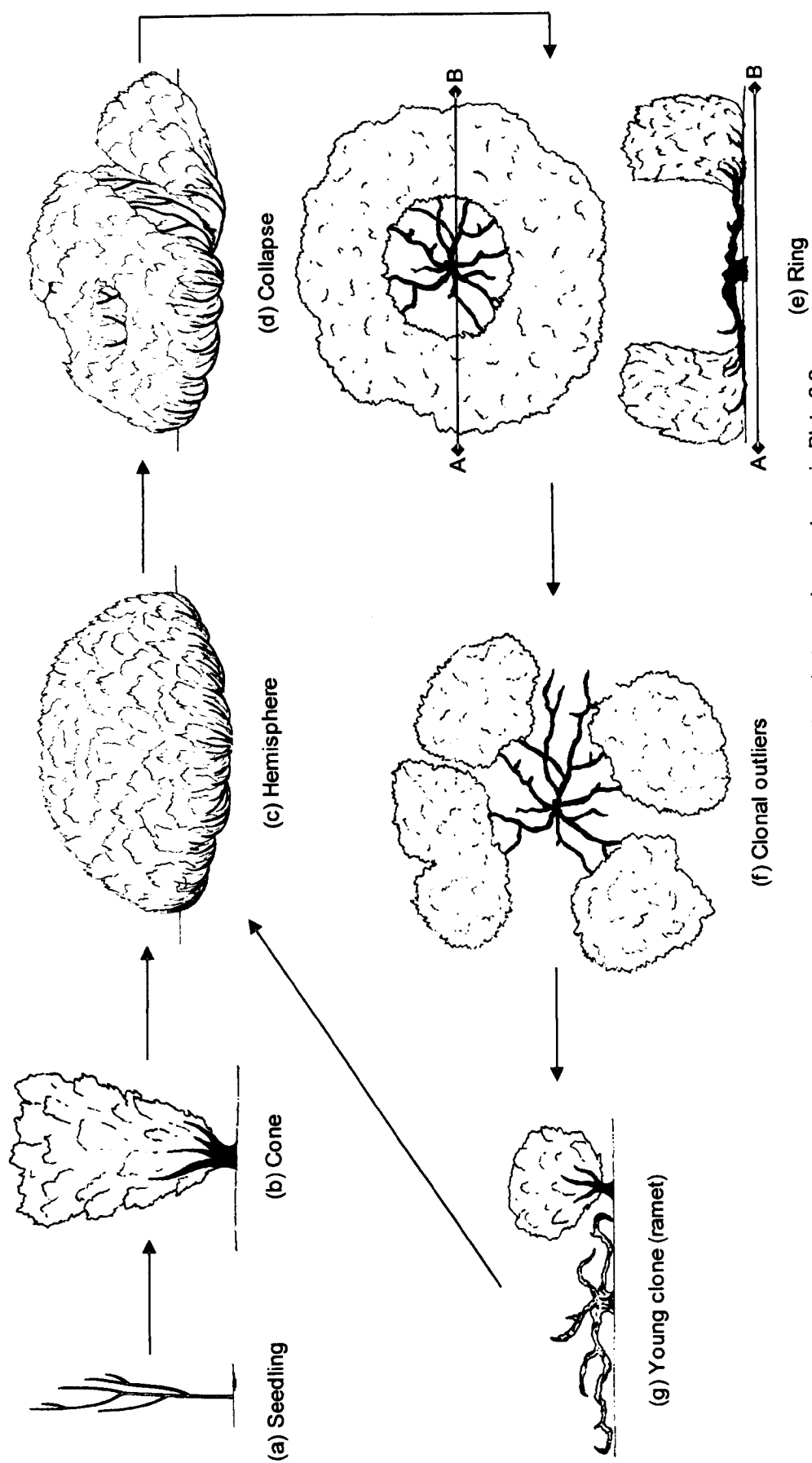


Figure 2-5 Diagrammatic representation of the ontogeny of *S. supranubius*. Corresponding photographs are shown in Plate 2.2.





Plate 2-2 Photographs of *S. supranubius* individuals at different ontogenetic stages. Photos (a) to (g) show *S. supranubius* individuals in the following stages: seedling, cone, hemisphere, collapsed, ring, outlier and young clone (see Figure 2-5)

## CHAPTER 3: DATA ANALYSIS METHODS

---

This chapter provides descriptions of the spatial point pattern analysis techniques that form the basis of this thesis, namely Ripley's  $K(r)$  (and its derivative, the  $L(r)$ -function) and the pair-correlation function ( $g(r)$ ).

### 3.1 QUANTIFICATION OF SPATIAL PATTERNS FROM THE 'PLANT'S-EYE'

It is well accepted that plants do not respond to average spatial structures, such as density per hectare, but to the biotic and abiotic composition of their immediate neighbourhood (Purves and Law, 2002). Turkington and Harper (1979) argued that analyses of vegetation are most biologically meaningful if there is no imposed anthropocentric scale. By using plants themselves as sampling locations, indices based on plant–plant distances quantify the “plant's-eye view” (Turkington and Harper, 1979), assessing the spatial distribution of individuals within the local neighbourhood of an individual plant (Murrell et al., 2001).

This section describes the techniques available to analyse the spatial pattern of individuals within a population. The study of spatial pattern has arisen more or less independently, and with different motivations, in several branches of science (e.g., geology, geography, ecology, hydrology). Consequently, the choice of analytical techniques is overwhelming (Dale et al., 2002). Although many of the techniques are computationally and mathematically similar, they are often shown to have differing powers and sensitivities and can result in contrasting interpretations of equivalent patterns (Diggle, 2003; Perry et al., 2006). Some studies have attempted to direct the selection of techniques depending upon research objectives and/or sample design (Fortin et al., 2002; Perry et al., 2006). It is often recommended that multiple techniques are employed simultaneously to avoid interpretative bias (Dale, 1999; Perry et al., 2002, 2006). In the following text attention is paid to the use of distance based indices in general and specifically the techniques with which this thesis is

concerned: Ripley's  $K(r)$ -function (and the associated  $L(r)$ -transformation) and the  $g(r)$ -function. These techniques provide formal measures of the density of individuals in the neighbourhood of the average plant, providing a quantification of the average plant's-eye view.

Table 3-1 defines some of the fundamental concepts used in spatial pattern analysis. The analysis of spatial patterns first became commonplace in ecological studies in the 1950s and 1960s (Perry et al., 2006). Early techniques were based on counts of individuals within sampling units, such as quadrats (Dale et al., 2002). These are broadly defined as area based methods. Area-based methods are, however, heavily criticised, primarily because the detection of spatial pattern is strongly influenced by the size of the sample unit used (Curtis and McIntosh, 1950; Grieg-Smith, 1983). Furthermore, area-based analyses such as block-quadrat variance and the variance-to-mean ratio could not reveal the scales at which spatial structure is most apparent (Mahdi and Law, 1987). Thus, while area-based methods can help understand a pattern's 'first-order effects' (see Table 3-1), they neglect information about the distances separating individual points. The spatial correlation structure of points describes the 'second-order effects' of a pattern (see Table 3-1). An understanding of the second-order properties of a pattern is required if inferences about the underlying mechanisms (e.g., abiotic or biotic processes) are to be made. Over recent decades, advances in computation power and the improved ability to manage spatial datasets has meant that mapped distributions of individuals can be analysed, leading to a new class of analyses using distance-based measures. The following sections describe some of the main distance-based techniques used in the analysis of ecological spatial patterns. A complete survey of all the techniques available is beyond the scope of this thesis, and several comprehensive texts already exist (e.g., Dale, 1999; Diggle, 2003; Fortin and Dale, 2005; Illian et al., 2008; and the 2002 special issue in *Ecography* [Volume 25, Issue 5]). Information on how the various methods compare and relate to one another is provided by Dale et al. (2002), Fortin et al. (2002) and Perry et al. (2006).



Table 3-1 Description of some fundamental spatial point pattern concepts.

Spatial pattern concept	Description
Spatial pattern	The arrangement of organisms, or patches of organisms, in space.
Point	An occurrence of the variable whose spatial pattern is being analysed (i.e., an occurrence of an individual shrub).
Univariate pattern	The pattern displayed by points of a single 'type' (e.g., a single species).
Bivariate pattern	The pattern of two (or more) 'types' of points (e.g., different species or different size classes).
First-order pattern	First-order spatial pattern describes the trend in the data over the entire study area (i.e., the mean number of points per unit area; Illian et al., 2008). First-order statistics describe the probability of a point being found at a given location. If point density is constant over space the point distribution is described as being homogeneous; if point density is not constant over space the point distribution is described as being heterogeneous.
Second-order pattern	Second-order pattern describes the small-scale spatial correlation structure between points (i.e., the pattern of points relative to one another).
Complete spatial randomness (CSR)	A type of second-order pattern. Defined formally, CSR has two properties (Stoyan and Penttinen, 2000, p. 72): 1) the number of points in a region $A$ (with area $ A $ ) follows a Poisson distribution with a mean intensity of $\lambda A $ and, 2) the distribution of $n$ points in $A$ follows an independent sample from the uniform distribution of $A$ . Under CSR, therefore, points occur independently of each other (i.e., they do not interact) and the intensity of points will not vary across the region studied (i.e., all sub-regions of the same size have an equal probability of containing a given number of points; Diggle, 2003; Dale, 1999). Second-order patterns that do not follow CSR are described as either aggregative or dispersed.
Aggregation	Aggregated patterns are also described as being 'clustered' or 'contagious'. In an aggregative pattern the presence of a point increases the probability of finding another point in its vicinity (Dale, 1999).
Dispersion	Dispersed patterns are also described as being 'regular' or 'segregated'. In a dispersed pattern the presence of a point reduces the probability of finding another point within its vicinity (Dale, 1999).

### 3.2 DISTANCE-BASED METHODS FOR THE QUANTIFICATION OF SPATIAL PATTERN

Second-order spatial point pattern statistics describe the correlation structure among points relative to point density (Illian et al., 2008). One of the earliest and most widely recognised techniques is Clark and Evan's (1954) nearest neighbour method. This technique measures the distance separating randomly selected plants from their nearest neighbouring plant. The mean nearest neighbour distance is compared to the mean distance expected under CSR, with the ratio of these two values indicating the presence and form of spatial pattern. Specifically, the ratio would be less than one, equal to one, or greater than one (with a maximum value of 2.1491) under conditions of aggregation, randomness or dispersion respectively. Although Clark and Evans' technique gave some indication of the scale of the observed patterns, it could only assess spatial structures occurring at the 1<sup>st</sup> spatial order, and is biased towards the detection of regularity (Figure 3-1).

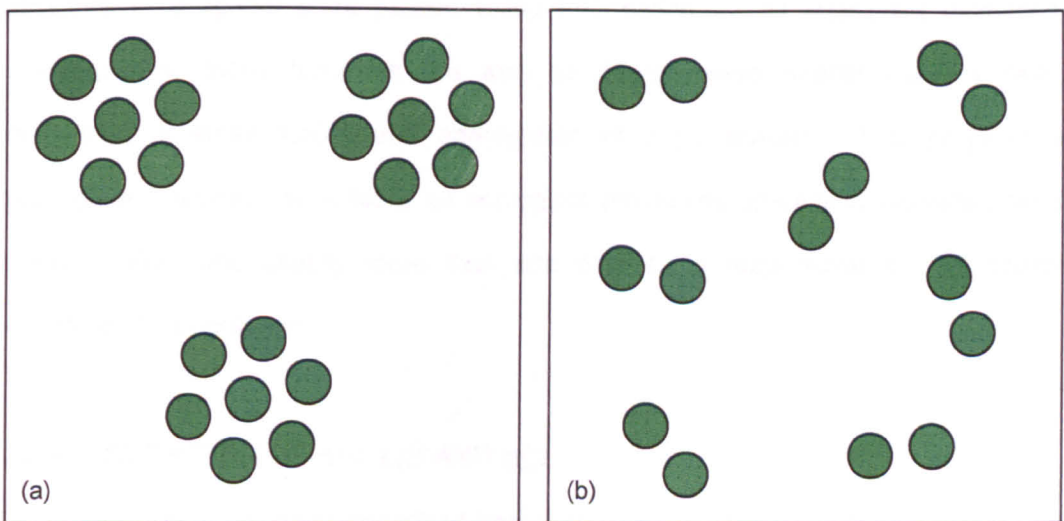


Figure 3-1 A limitation of Clark and Evans' (1954) index. In both (a) and (b) the distances between plants and their nearest neighbours are equal and will thus produce the same distribution under the nearest neighbour analysis of Clark and Evans (1954). However, (a) and (b) clearly demonstrate contrasting spatial patterns operating at different spatial scales.

The biotic and abiotic processes that drive a population's spatial structure operate at multiple spatial scales. In order to fully understand the spatial structure of a population, analytical techniques must be capable of identifying how patterns change

with scale. Many refinements of Clark and Evans' technique (1954) were recommended (e.g., Cottam et al., 1953; Thompson, 1956; Davis et al., 2000; Diggle, 2003). These mostly suggested the use of the 2<sup>nd</sup>, 3<sup>rd</sup>...n<sup>th</sup> nearest neighbour. However, these techniques only assessed pattern from a random selection of individuals. Without considering all individuals the scales of pattern identified would depend upon local neighbourhood densities.

Nowadays the most commonly employed methods of spatial pattern analysis are Ripley's  $K(r)$ -function (Ripley, 1976, 1981), the  $L(r)$ -function (a transformation of  $K(r)$ ), and the  $g(r)$ -function (also known as the 'pair-correlation function' and the 'neighbourhood density function'; Stoyan and Stoyan, 1994). Such analyses require spatially referenced maps of all individuals within a sample plot. Analysing such data requires computer software capable of its manipulation, especially if large datasets are to be analysed. Over recent decades such computational power has been developed. All three techniques are powerful tools capable of describing the second-order structure of a spatial point pattern using information on all inter-point distances. Consequently, these functions are able to detect mixed spatial patterns (e.g., dispersion at small scales and aggregation at large scales). This property is particularly important as virtually all ecological processes operate at discrete scales (Levin, 1992), and usually more than one process is responsible for the spatial structure of a population.

### **3.2.1 DEFINITION OF $K(r)$ , $L(r)$ AND $g(r)$**

All three of the techniques described below are global pattern statistics providing an indication of the density of other plants at increasing distances ( $r$ ) around an average plant (Law et al., 2009).

#### ***Ripley's $K(r)$ -function***

Ripley's  $K(r)$ -function superimposes circles of increasing radius  $r$  on each point. The maximum scale of interest ( $r_{\max}$ ) is pre-defined by the user.  $K(r)$  provides the expected number of points within distance  $r$  of an arbitrary point (without counting the focal

point; Felinks and Wiegand, 2008). The observed distribution of points is compared with the distribution of points expected for a spatially random pattern of the same intensity.  $K(r)$  is defined as:

$$K(r) = \frac{A}{n^2} \sum_{i=1}^n \sum_{j=1}^n \frac{I_r(d_{ij})}{w_{ij}}$$

Where  $A$  is the area of the plot,  $n$  is the number of points in the study region,  $d_{ij}$  measures the distance between point  $i$  and point  $j$ ,  $I_r$  is a counter variable [ $I_r(d_{ij}) = 1$  if  $d_{ij} \leq r$ , and  $I_r(d_{ij}) = 0$  otherwise], and  $w_{ij}$  is a weighting factor used to reduce the problem of edge effects. The weight,  $w_{ij}$ , for a pair of points is given by the proportion of the area of a circle centred on the  $i$ th point, with radius  $d_{ij}$ , that lies within the study region (Perry et al., 2006). If the circle is completely contained within the study area,  $w_{ij} = 1$ , otherwise it is the reciprocal of the proportion of the circle's circumference within the plot (Fortin and Dale, 2005; Haase, 1995; Goreaud and Pélissier, 1999). For a completely random (Poisson) process:

$$K_{pois}(r) = \pi r^2$$

### **The $L(r)$ transformation**

It can be difficult to interpret  $K(r)$  visually (Wiegand and Moloney, 2004). To stabilise the variance and increase the ease of interpretation a square-root transformation of  $K(r)$ -function, the  $L(r)$ -function, is often used (Besag, 1977; Wiegand et al., 2006):

$$L(r) = \sqrt{\frac{K(r)}{\pi}}$$

$L(r)$  has a value of 0 under CSR. Aggregation is indicated if  $L(r) > 0$ , whereas dispersion is indicated if  $L(r) < 0$  (although note that some authors use a slightly different form of the formula which gives the reverse interpretation [e.g., Dale, 1999; Dale et al., 2002; Stoll and Bergius, 2005]; Figure 3-2a and b).

### ***The $g(r)$ -function***

$g(r)$  is directly related to  $K(r)$ :

$$g(r) = \frac{K'(r)}{2\pi r}$$

where  $K'(r)$  is the derivative of  $K(r)$ . The  $g(r)$ -function (hereafter  $g(r)$ ) is calculated by replacing the circles used in the calculation of  $K(r)$ , with rings (Figure 3-2c). Thus, instead of counting the number of points within circles of radius  $r$ ,  $g(r)$  counts the number of points at distance  $r$  away from the focal point. As with  $K(r)$ , the observed distribution of inter-point distances is compared to the distribution of distances expected for a spatially random pattern of the same intensity. Under CSR  $g(r) = 1$ . If  $g(r) > 1$  then pairs of plants are more frequent at distance  $r$  than expected under CSR, indicating aggregation. If  $g(r) < 1$ , then pairs of plants are less frequent at distance  $r$  than the spatial average distance, indicating dispersion (Law et al., 2009; Figure 3-2c and d).

The calculation of  $g(r)$  requires a technical decision on the width of the rings. Ring widths that are too small will produce jagged plots as there will be too few points falling within each distance class (Figure 3-2e and f). This may lead to erroneous interpretations of aggregation and/or dispersion. However, increasing the ring widths too much will remove the advantage of isolating specific distance classes (Wiegand and Moloney, 2004). Wiegand (pers. comm.) recommends that several ring widths are trialled and the smallest ring width that produces a smooth plot is selected.

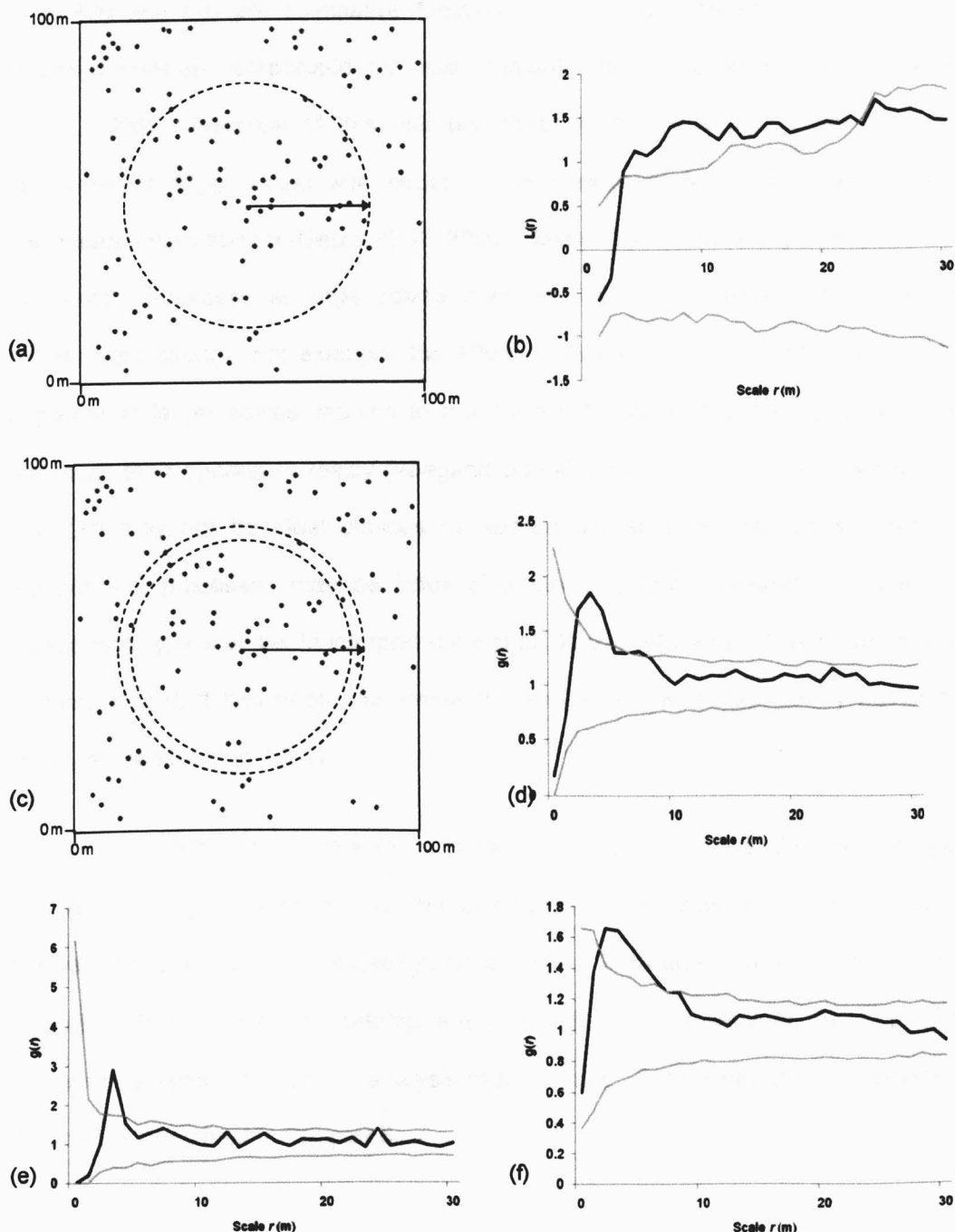


Figure 3-2 The quantification of spatial pattern using, (a) and (b)  $L(r)$  and, (c) and (d)  $g(r)$ . The effects of different ring widths on  $g(r)$  are shown in graphs (e) and (f) which use ring widths of 1 m and 4 m respectively. The black lines on graphs (b), (d), (e) and (f) show the values of the empirical functions, whereas the grey lines show the values of approximately 99% Monte Carlo simulation envelopes constructed from the 5<sup>th</sup>-highest and 5<sup>th</sup>-lowest value of 999 simulations of the CSR null model. The data underlying figures (b), (d), (e) and (f) are taken from Chapter 5. Specifically they display and analyse the point pattern of all *S. supranubius* individuals in the 5<sup>th</sup> replicate at an extent of 1 ha on Substrate 2 (see Chapter 5 for more details).

Both  $K(r)$  and  $L(r)$  are cumulative functions. Thus measurements made at small values of  $r$  will be incorporated into measurements made at larger values of  $r$  (Getzin et al., 2006). Because of this memory effect, both  $K(r)$  and  $L(r)$  can confound structures at large scales with structures at smaller scales. This can severely complicate interpretation (Getzin et al., 2006). Most notably, the ability of  $K(r)$  and  $L(r)$  to identify processes at large scales may be impeded, especially if small-scale patterns are strong. For example, the effect of small-scale aggregation may still be apparent at larger scales, leading to the misinterpretation of clustering over longer distances than operate in reality (Wiegand and Moloney, 2004). Consequently,  $K(r)$  and  $L(r)$  may not be ideal choices for exploratory analysis, especially if several independent processes may be influencing the population's spatial structure. In comparison,  $g(r)$  is easier to interpret (Law et al., 2009). However, despite the relative benefits of  $g(r)$ , it has been less frequently applied in the contemporary ecological literature (see Section 1.3.4).

The following sections describe some of the major features of spatial pattern analysis before providing a description of the calculation of the indices in the grid-based software Programita, which is used in Chapters 5 – 7 (Wiegand and Moloney, 2004). Specific details of analytical settings are provided in the subsequent chapters. The following sections focus on the analysis of point data. Techniques for the analysis of objects are described in Chapter 6.

### **Null models**

Spatial patterns can be compared to null models of spatial processes, enabling researchers to test biological hypotheses. The most common null model is CSR, although alternative null models may be defined. The selection of a suitable null model is a critical step if misinterpretations and incorrect biological conclusions are to be avoided (Diggle, 2003).

## Simulation envelopes

The 'significance' of an observed pattern is usually assessed by comparing the observed data with Monte Carlo simulations of the processes underlying the spatial null model. Each Monte Carlo simulation generates a test function ( $g(r)$  or  $L(r)$ ). Approximate  $n(n+1) \times 100\%$  simulation envelopes can be constructed from the highest and lowest values (at each scale  $r$ ) of  $n$  simulations of the null model (Bailey and Gatrell, 1995). A more accurate approach constructs simulation envelopes from the 5<sup>th</sup>-highest and 5<sup>th</sup>-lowest values of the simulations of the function (Wiegand and Moloney, 2004). Departure from the null model is indicated if the empirical function lies outside the simulation envelopes. It is stressed that this technique does not provide a formal estimate of the statistical significance of an observed pattern, as Type I error rates are underestimated (Loosmore and Ford, 2006). Nevertheless, Monte Carlo envelopes provide an indication of the likelihood of an observed pattern. Monte Carlo simulations of the null model are used in subsequent chapters, although their interpretations are reinforced with additional analyses. In the subsequent chapters Monte Carlo simulation envelopes are constructed from the 5<sup>th</sup>-highest and 5<sup>th</sup>-lowest value of 999 simulations of the null model to try to reduce the likelihood of making Type I errors. Increased sample sizes produce narrower simulation envelopes (Wiegand, 2004). Therefore, it is possible that, when comparing spatial patterns, apparent differences in pattern significance could be due to either a real difference in the magnitude of aggregation (for example), or a sample-size driven change in the width of the simulation envelope (Figure 3-3).



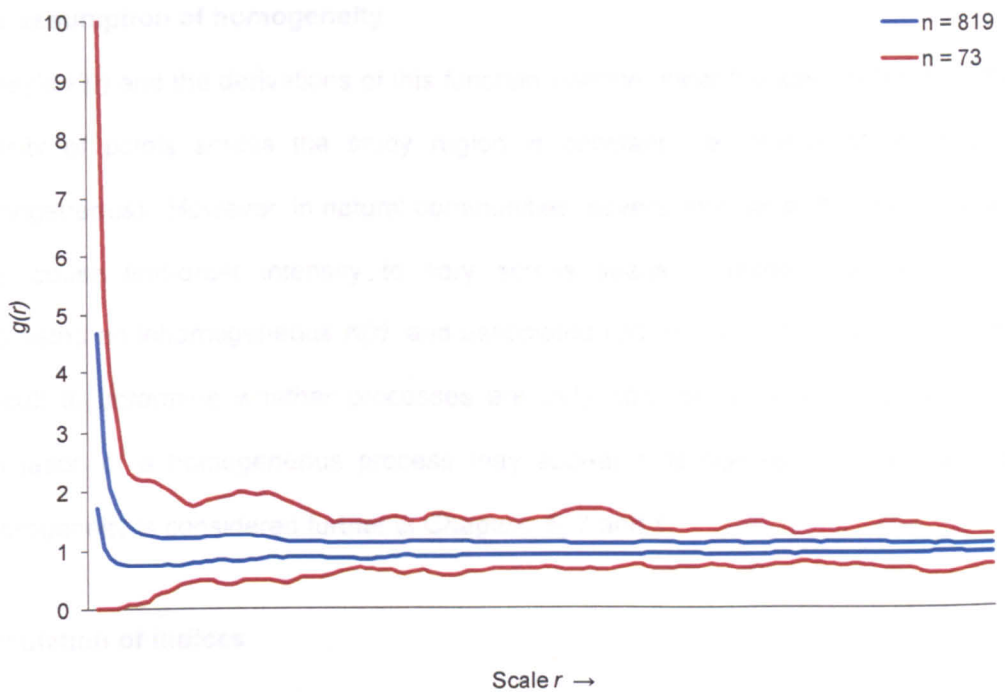


Figure 3-3 Sample-size driven changes in the width of Monte Carlo simulation envelopes. Two simulations of a Thomas cluster process ( $r\text{Thomas}(10,0.2,5)$ ) were performed in Spatstat (Baddeley and Turner, 2005), the first over a small area ( $n = 73$  points; red lines), the second over a larger area ( $n = 819$  points; blue lines). The simulations envelopes were produced from the highest and lowest value of 99 simulations of the CSR null model using each dataset in turn.

### Edge effects

Individuals located near the edge of the sample window have fewer neighbours than individuals located nearer the centre of the sample window. If  $K(r)$ ,  $L(r)$  or  $g(r)$  were calculated from an individual near the edge of the sample window, portions of the circles and/or rings of the function would fall outside the sample window and they would necessarily detect a lower density of individuals than equivalent analyses centred on points closer to the centre of the sample window. When the size of the sample window is large relative to the scales of the analysis, relatively few individuals are located near the edge of the sample window, and the average function should not be greatly affected by edge effects. However, when the sample window is small relative to the scale of the analysis a high proportion of individuals will be located near the edge of the sample window and edge effects may influence the calculation of the statistic. In these situations corrections for edge effect should be used. The calculation of  $K(r)$ ,  $L(r)$  and  $g(r)$  includes a weighting factor that helps to account for edge effects.

## **The assumption of homogeneity**

Ripley's  $K(r)$  and the derivations of this function operate under the assumption that the density of points across the study region is constant (i.e., the point process is homogeneous). However, in natural communities, several internal and external forces may cause first-order intensity to vary across space. Baddeley et al. (2000) suggested an inhomogeneous  $K(r)$ , and associated  $L(r)$  and  $g(r)$ . However, it is often difficult to determine whether processes are truly homogeneous since a particular realisation of a homogeneous process may appear heterogeneous. The issue of heterogeneity is considered further in Chapters 5, 7 and 8.

## **Calculation of indices**

In Chapters 5 to 7, both  $L(r)$  and  $g(r)$  are used to analyse the spatial structure of *S. supranubius*. Because of the aforementioned complications with the interpretation of  $L(r)$ , most emphasis is placed on  $g(r)$ . In Chapters 5 to 7 analyses using  $L(r)$  and  $g(r)$  were performed in the software Programita (Wiegand and Moloney, 2004). Because of the need for modelling capabilities, analyses in Chapter 8 were performed using the spatstat package (version 1.17-4; Baddeley and Turner, 2005) in R (version 2.10.1; R Development Core Team, 2009).

The analytical approaches described above use all pairs of points to derive  $K(r)$ ,  $L(r)$  and  $g(r)$ . These calculations are based on complex algorithms that can be computationally intensive and time-consuming to compute, especially when large quantities of data are used (Law et al., 2009). Programita uses an underlying grid to simplify the computation of second-order statistics (Figure 3-4). The calculation of point–point distances necessary for estimation of second-order statistics is then based on distances between cells, and counting points in cells (Figure 3-4). It may be argued that the use of grids will reduce the accuracy of pattern analyses at small scales as information on spatial location below the size of the cell will be lost (Wiegand and Moloney, 2004). This may introduce error into the small scale approximation of the location of points, which has been shown in previous studies to

affect the performance of  $L(r)$  (Freeman and Ford, 2002). However, if grid sizes are small relative to the size of the individuals considered, accuracy will not be lost. Analyses in Chapters 5 to 7 use a grid size of 1 m<sup>2</sup> as this is the smallest *S. supranubius* canopy area investigated (see Section 4.4).

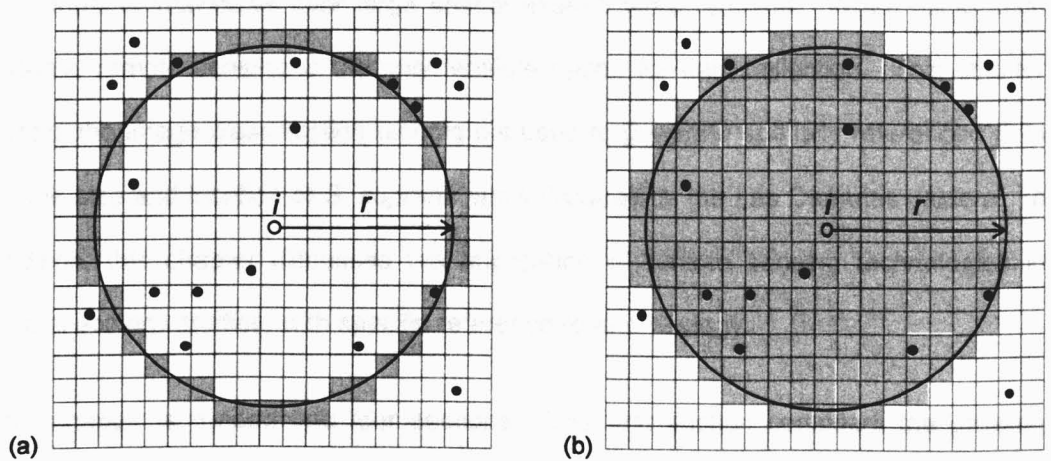


Figure 3-4 Implementation of (a)  $g(r)$  and, (b)  $L(r)$  in the grid-based software Programita. Graph (a): For implementation of  $g(r)$  Programita counts the number of points inside the ring at distance  $r$  from the focal point (i.e., the grey shaded area), and the number of cells within this area. Graph (b): For implementation of  $L(r)$  Programita counts the number of points within a circle with a radius  $r$  (i.e. the grey shaded area), and the number of cells within this region.

## CHAPTER 4: DATA COLLECTION: MAPPING THE SIZE AND LOCATION OF *S. SUPRANUBIUS*

---

To address the aims of the research (Section 1.2) data on the size and location of *S. supranubius* individuals over large spatial extents were required. When using such extents, remote sensing is the only feasible option for data collection. This chapter details the image classification techniques used to generate spatially referenced data on the size and location of *S. supranubius* individuals in the Las Cañadas caldera. In addition, this chapter discusses the application of remote sensing technologies in spatial ecology studies, with specific reference to arid ecology.

This chapter is divided into four sections. The first section considers the growing importance of remote sensing in ecological studies and the potential difficulties surrounding its use in arid environments. This section also discusses the data source used in this thesis. The second section considers methods of image classification and introduces classification by support vector machines, specifically one-class classification by support vector data description. The third section details the extensive cross-validation used to determine the optimum classifier structure for analysis of the current imagery. In the final section methods of image post-processing are detailed and the optimum classifier is used to produce one-class maps of the distribution of *S. supranubius* on five contrasting substrates in the Las Cañadas caldera.

### 4.1 ECOLOGICAL APPLICATIONS OF REMOTE SENSING

Although ecological studies at all spatial extents remain important, the ability to conduct ecological studies at broad extents has arisen in part from advances in technology (Foody, 2007), in particular the increasing availability and utility of remote sensing and geographical information systems (GIS). Through the acquisition and interpretation of aerial and satellite images, remote sensing is able to provide ecologists with information on the Earth's surface and environment, whereas GIS provides a means to store, visualise and analyse the data generated (Foody, 2007).

Remote sensing is routinely used to produce thematic maps of land cover. Depending upon the spatial resolution of the imagery, these thematic maps can provide information ranging from the location of vegetation assemblages, to the location of specific species, and sometimes individuals. There is a long history of studies using aerial photography to classify, delineate and map broad vegetation types (e.g., Wilson, 1920; Tiwari and Singh, 1984; Ibrahim and Hashimi, 1990; Turner et al., 1996; Huebner et al., 1999; Fensham et al., 2002; Müllerová, 2004). Furthermore, aerial photographs have been instrumental in discovering vegetation organisations not observable from the ground, such as the discovery of spatially periodic vegetation patterns in arid and semi-arid ecosystems (MacFadyen, 1950; Clos-Arceud, 1956; cited in Coutron, 2002).

Aerial photographs are one of the only information sources extending back into the 20<sup>th</sup> century, with many areas photographed as early as the Second World War (Verheyden et al., 2002; Okeke and Karnieli, 2006). This provides an unparalleled opportunity to study medium- to long-term temporal vegetation dynamics with minimal time and financial input (e.g., Ambrose and Bratton, 1990; Jacobson et al., 1991; Tanaka and Nakashizuka, 1997; Huebner et al., 1999; Kadmon and Harari-Kremer, 1999; Wu et al., 2000; Fujita et al., 2003). Furthermore, the availability of high-resolution imagery (particularly from airborne platforms) is increasingly allowing users to locate, identify and monitor the dynamics of individual plants and populations. Despite the potential utility of remote sensing, however, its application in ecological studies remains limited (Newton et al., 2009). This could be attributed to concerns over the accuracy with which the technique is capable of identifying vegetation cover types and individual species, financial limitations, or a lack of awareness of remote sensing techniques and their capabilities.

The surface characteristics and structure of an object determine its reflectance. Different land covers usually have distinct spectral characteristics displayed in the ratio of red, green and blue wavebands. Most historical aerial images are monochromatic and, as such, many studies have employed grey-level thresholding

to discern floristic compositions. However, by expressing all visible wavelengths in grey tones, the spectral distinctiveness of different species is reduced. Thus many studies have been limited to identifying phytosociological *groups* of species (e.g., 'tree', 'shrub' and 'herbaceous') instead of individuals (e.g., Carmel and Kadmon, 1998; Hudak and Wessman, 1998; Kadmon and Harari-Kremer, 1999; Sharp and Whittaker, 2003; Laliberte et al., 2004; Briggs et al., 2007), or studying easily separable (often dominant) species (e.g., Goslee et al., 2003; Leckie et al., 2003). Realistically, classifications based upon grey values alone are merely mapping changes in image intensity (i.e., vegetation density) and inferring changes in vegetation class, which can result in many errors of omission and commission (Goslee et al., 2003). Colour photography provides greater between-species spectral discrimination. Using various measures of pixel colour, brightness and intensity, Meyer et al. (1996) and Leckie et al. (2003) were able to use semi-automatic analysis to identify four and six tree species respectively. The accuracy of automatic identification from colour photography depends upon species' spectral uniqueness (i.e., low within-species variation and high between-species variation in spectral responses) (Müllerová, 2004). Variations in crown structure (shading effects), crown density (background materials) and differing visibility of tree components (twigs, needles, leaves, branches) can all increase the within-species spectral variability, potentially reducing the accuracy of automatic identification (Meyer et al., 1996).

#### **4.1.1 REMOTE SENSING AND ARID ECOLOGY**

Aerial photographs covering large spatial and temporal extents provide an exciting opportunity to monitor the dynamics of slow-growing arid shrubs. Several studies of arid systems have employed digital imagery analysis to investigate the dynamics of woody vegetation encroaching into savannah grasslands (Hudak and Wessman, 1998; Sharp and Whittaker, 2003; Laliberte et al., 2004; Briggs et al., 2007). The sparse and structurally simple vegetation of arid regions means that the canopies of individual plants are often clearly discernable, making the delimitation and

identification of individuals easier than in less dispersed communities. It is therefore surprising, perhaps, that more studies have not used aerial photographs to investigate the demography and dynamics of arid shrubs. However, despite offering a seemingly simple canvas, several attributes typical of arid landscapes and vegetation can limit the accuracy of image classification, even when high resolution imagery is available.

### **Substrate spectral qualities**

The accurate retrieval of pixel-based information is influenced by the nature of the substrate. With low organic matter content, desert soils tend to be bright. Therefore, the spectral qualities of the soil may interfere with the spectral contribution of the vegetation, especially when pixels are large (Okin and Roberts, 2004). This may result in the mis-classification of vegetation as substrate, especially when vegetation cover is sparse.

### **Evolutionary adaptations of desert plants**

Because of the intense radiation they experience, many desert plants have evolved several morphological and physiological adaptations that can have a marked effect on their spectral reflectance. For example, because of the high water and energy demands of producing new leaves, many desert plants reduce leaf surface area, or avoid leaves altogether, moving photosynthesis to stalks and stems (Okin and Roberts, 2004). Many desert plants have highly reflective spines or hairs encasing their stems (Sandquist and Ehleringer, 1998). These are designed to further reduce evapotranspiration by reflecting radiation and creating a still-air layer around photosynthetic organs. Furthermore, because of the high concentration of ambient photosynthetically active radiation, many desert plants maintain low chlorophyll concentrations in their leaves and stems.

Highly reflective organs and reductions in plant biomass can greatly reduce the per-pixel spectral contribution of vegetation (Ehleringer and Mooney, 1978; Okin and Roberts, 2004). Highly reflective vegetation may be difficult to distinguish from the

typically bright and reflective soils of arid regions. Further, with numerous arid species adopting these techniques, different species may be spectrally similar. For example, when using field reflectance spectra Okin et al. (2001) found that the desert shrubs *Atriplex polycarpa* (Torr.) S. Wats. (Chenopodiaceae) and *Larrea tridentata* (DC) Colville (Zygophyllaceae) were spectrally alike.

### **Intra-specific spectral variability**

The spatial and temporal heterogeneity of resources in deserts can produce parallel heterogeneities in intra-specific plant morphology (Okin and Roberts, 2004). Event-driven demographics may mean that individuals of the same species may exhibit highly variable spectral qualities across space and through time. Consequently, *a priori* knowledge of the spectral qualities of species during different phenological stages may be an important tool for arid species identification (Karnieli et al., 2002). In addition to event-driven phenological changes, arid plants are known to undergo temporally sequential morphological changes. For example, *Atriplex hymenelytra* (Torr.) S. Wats. (Chenopodiaceae), found in the hot deserts of Mexico and the southwestern USA, changes its leaf characteristics (surface area) in time with the seasons in an apparently adaptive manner (Mooney et al., 1977).

Perhaps because of the difficulties surrounding image interpretation of arid vegetation, relatively few studies have used the technique to map and investigate the dynamics of arid shrubs. However, with recent improvements in both spectral and object-orientated image classification techniques, several studies have been able to locate and quantify attributes of individual plants (Couteron, 2001; Strand et al., 2006; Malkinson and Kadmon, 2007). However, these studies typically only consider a small area of a species' range. By combining improved image classification techniques and high resolution imagery with pattern–process inference (see Section 1.3.3), ecologists are now ideally placed to investigate the dynamics of arid shrubs. In addition to this theoretical opportunity, aerial imagery analyses have practical benefits. Arid ecosystems are often some of the most inhospitable and inaccessible regions of the



world, so field-based investigation of communities can be both expensive and physically demanding.

## **4.2 IMAGE CLASSIFICATION**

A three waveband (red, green and blue) image of the Las Cañadas caldera was obtained from Grafcan Plc. This imagery was captured on the 31<sup>st</sup> of December 2006 using a Wild RC-30 camera and provides high resolution, with pixel widths of 0.26 m. Imagery was obtained for the Parador colluvium, and the Majua, Arenas Blancas, Moñtana and Conejos lava flows (Substrates 1 to 5 respectively, see Section 2.3.2). Other than geo-rectification in Erdas Imagine 9.1, no pre-processing of the image was necessary.

Deriving data from remotely sensed images requires the image to be classified. Image classification aims to categorise all the pixels in a digital image into one of a number of classes. This categorisation can then be used to produce thematic maps detailing the spatial distribution and extent of a particular class of interest. This chapter applies the most common classification technique: spectral classification.

Spectral classification can be manual or computer-aided. Although manual interpretation has been shown to produce accurate distribution maps of species (Driscoll and Coleman, 1974; Myers and Benson, 1981; Trichon, 2001; Trichon and Julien, 2006), the technique has been criticised for two main reasons. Firstly, manual photo-interpretation lacks objectivity and consistency in measurement approach. Secondly, manual interpretation is labour intensive, increasing research costs and often limiting analysis to the consideration of either coarse vegetation structures over broad areas, or detailed structure (i.e., individuals) over small extents (Kadmon and Harari-Kremer, 1999; Müllerová, 2004). Furthermore, the success of manual photo-interpretation is largely determined by the familiarity of the researcher with both the technique and the area being studied, and therefore may be highly subjective and non-transferable (Driscoll and Coleman, 1974).

In contrast, automated techniques are non-subjective, transferable, efficient and capable of producing classification accuracies in excess of 80% (Meyer et al., 1996; Carmel and Kadmon, 1998; Kadmon and Harari-Kremer, 1999; Pouliot et al., 2002; Müllerová, 2004). Automated spectral classifications either categorise pixels by their spectral similarity (unsupervised) or allocate pixels to classes based on their similarity with pre-defined spectral responses defined by the user (supervised) (Foody, 2002). Supervised spectral classification is by far the most commonly applied technique in image classification.

Supervised classification comprises three distinct stages: training, allocation and testing. In the training stage, a quantitative description of the spectral characteristics of each class of interest is generated. Using this information each pixel in the image is allocated to the class with which it has the highest spectral similarity. Finally, the accuracy of the final classification is assessed. To avoid over-estimating classification accuracy, the testing stage should use a sample of pixels not used to train the classifier (Foody and Mathur, 2004). The accuracy of the classification is dependent upon the appropriateness of training data and the precise classification algorithm selected. Consequently, much research has focused on optimising these two stages. Achieving a high performance and transferable classifier is challenging. Specifically, the selection of training data (dataset size, composition and sampling design) and classifier algorithm are usually interdependent (Mathur and Foody, 2008), and often influenced by the research questions being addressed.

One of the most commonly applied supervised techniques is the maximum likelihood (ML) classifier (Huang et al., 2002). The ML classifier is parametric, and as such requires an exhaustive quantitative description of the spectral characteristics of each class in the image. Furthermore, the classes must be spectrally discrete and mutually exclusive (Foody, 2004). Resulting classifications may be spurious if these conditions are not met. ML classification requires large training datasets to ensure any variation in the spectral response of a class is fully described. It is often suggested that 30 independent training cases per class per waveband are needed to form a

representative training dataset, with a 'the larger the better' attitude often held (Mather, 2004). All classes within the image must be included in the training stage of the analysis (Foody et al., 2006). Standard supervised classifiers, such as ML, will typically seek to optimise the classification accuracy of all classes within the image even though researchers are typically only interested in the accurate classification of a single class of interest. Thus conventional classifiers are largely inefficient, and may produce poor classification accuracies. Several classification techniques have been developed that significantly reduce the effort required during training. This thesis considers classification by support vector machines (SVMs) which have been shown to be at least as accurate as other widely used techniques, if not more so.

#### **4.2.1 SUPPORT VECTOR MACHINE (SVM) CLASSIFICATION**

The principles of SVMs were established in the 1970s by Vapnik and Chervonekis' (1971) theories of statistical machine learning. The SVM algorithm was later developed by Vapnik (1995). SVM classifiers have been used in handwritten digit recognition (Gorgevik and Cakmakov, 2005), face detection (Osuna et al., 1997), text categorisation (Manomaisupat et al., 2006), signal recognition (Fagerlund, 2007) and automated animal species identification by sound (Acevedo et al., 2009). However, the potential of SVM in remote sensing has only recently been realised (Chapelle et al., 1999; Huang et al., 2002; Zhu and Blumberg, 2002; Pal and Mather, 2005). Comparative studies have shown that classification by a SVM can produce classifications that are at least as accurate as those from techniques such as maximum likelihood, neural networks and decision trees (Huang et al., 2002; Camps-Valls et al., 2004; Melgani and Bruzzone, 2004; Pal and Mather, 2005; Muñoz-Marí et al., 2007; Sanchez-Hernandez et al., 2007; Dixon and Candade, 2008; Guo et al., 2008). SVMs were designed for the binary separation of two classes, although the technique can be extended to multi-class scenarios (Huang et al., 2002). A detailed explanation of SVM can be found in Vapnik (1995). The salient features are discussed below.

SVMs provide a non-parametric boundary classification technique. Consider a situation in which there are two spectral classes whose values do not overlap (i.e., they can be easily separated by a linear plane): the class of interest and another class representing everything else in the image that does not belong to the former class. Each class occupies a unique location in  $n$ -dimensional feature space. For each of the  $r$  training cases there is a vector  $\mathbf{x}_i$  that represents its location in feature space along with a definition of class membership,  $y_i$ . Using the training data represented by  $\{\mathbf{x}_i, y_i\}, i = 1, \dots, r, y_i \in \{1, -1\}$ , an optimal separating hyperplane (OSH) is defined that divides the two classes (Foody and Mathur, 2006). The image can then be classified using the position of each pixel in relation to the OSH to determine which class the pixels should be allocated to. Theoretically several hyperplanes could be fitted, but only one OSH exists. The OSH is expected to generalise well when applied to unseen data requiring classification (Foody and Mathur, 2006).

Optimum separation is achieved by focusing on the data points located at the boundary of each class's distribution in feature space (so-called 'support vectors'). The support vectors of the two classes lie on two hyperplanes ( $H_1$  and  $H_2$  in Figure 4-1). The OSH lies equidistant between the two hyperplanes such that all the samples of a class are on one side of it and the distance from the OSH to the training cases in both of the classes is as large as possible (Foody and Mathur, 2006; Figure 4-1a). Support vectors are the critical elements of the training dataset. If all other training points were removed or changed location, and the training was repeated, the same separating hyperplane would be created. Thus whereas other non-parametric classification techniques often require large training datasets (Hubert-Moy et al., 2001), SVM classification is able to achieve high accuracies from a very small training dataset comprising of a few support vectors. Mathur and Foody (2008) have shown that by intelligently selecting training samples predicted to lie on the edge of a class's spectral distribution, SVM classifiers can achieve high accuracy (~91%) with only very small training samples (see also Foody and Mathur, 2004).

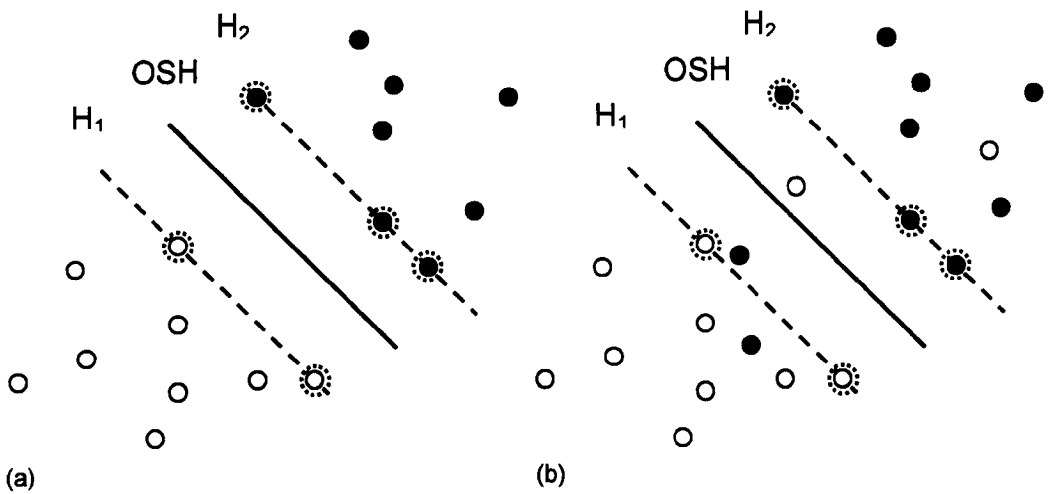


Figure 4-1 Support vector machine (SVM) classification for (a) linearly separable classes and, (b) non-linearly separable classes. The optimal separating hyperplane is shown as a solid line. The boundary hyperplanes for each class are shown as dashed lines. Circled cases are support vectors.

The above description applies when dealing with linearly separable classes (*cf.* Figure 4-1b). However, datasets are rarely linearly separable. More often classes overlap in feature space. As such both classes will have a high proportion of outliers (i.e., training samples on the 'wrong side' of the linear hyperplane). In these situations it is possible to distinguish non-linearly separable classes (Figure 4-1b) by applying kernel functions. Kernel functions map the training data into a higher dimensional space where a linear learning machine can be applied (Figure 4-2). Kernels have several parameters that must be pre-defined by the user. The accuracy with which a SVM classifies is largely dependent upon these parameters; however, there is very little guidance in the literature on the criteria to be used to select the optimum kernel and parameter values (Pal and Mathur, 2005). The most commonly used kernels are the Gaussian radial basis function and the polynomial kernel. More information on these and other kernels is provided in 4.3.2.

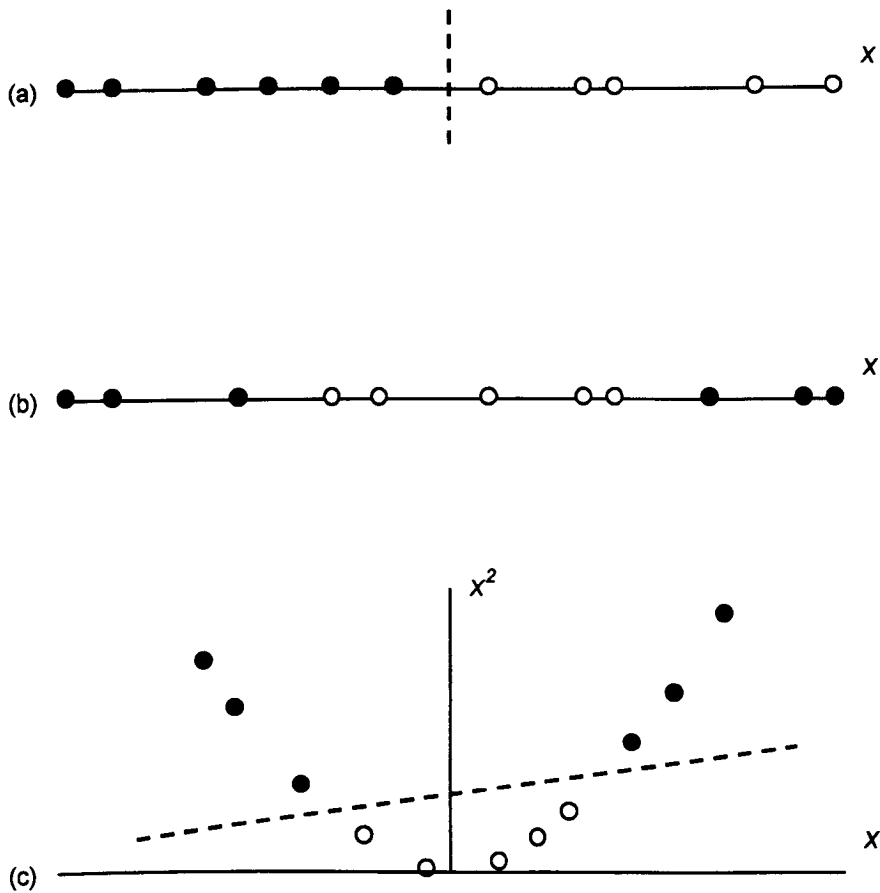


Figure 4-2 Classifying non-linearly separable classes using kernels. Re-mapping data to a higher dimension via a polynomial kernel function can make the data more easily separable. Figure (a) shows two sets of data that are easily separable by a linear classifier (dashed line) whereas in (b) they are not. If the data in (b) are re-mapped at a higher dimension (quadratic in this case) they may be more easily separable (c).

In the above examples, training data from two classes are used: the class of interest and another class representing everything that does not belong to the former class (Foody et al., 2006). However, it is possible to adapt the technique to a one-class classifier (OCC), requiring only data on the class of interest. OCCs have demonstrated considerable utility in many applications including ecological modelling (Guo et al., 2005; Drake et al., 2006; Kelly et al., 2007), document classification (Zhuang and Dai, 2006; Manevitz and Yousef, 2007), facial expression recognition (Zeng et al., 2006) and machine diagnostics (Shin et al., 2005). Although they have great potential in remote sensing, OCCs have been little used (but see Sanchez-Hernandez et al., 2007). OCCs can use a variety of analytical approaches, including

reconstruction methods, density methods and boundary methods (Sanchez-Hernandez et al., 2007). Using the principles of SVM, Tax (2001) was able to create an OCC that employs the boundary method and so requires relatively few training data. This technique, the support vector data description (SVDD), is ideally suited to remote-sensing applications.

#### **4.2.2 ONE-CLASS CLASSIFICATION (OCC) BY SUPPORT VECTOR DATA DESCRIPTION (SVDD)**

Support vector data description (hereafter referred to as SVDD) is a boundary method OCC based on the principles of the support vector machine (SVM) (Tax, 2001; Tax and Juszczak, 2003; Tax and Duin, 2004). In the following text the basic operation and main features of SVDD are discussed.

In SVDD classification a closed  $n$ -dimensional sphere (hereafter referred to as a hypersphere) is fitted around the training data, separating the class of interest from all other classes. The hypersphere can be described as having centre  $a$  and radius  $R$  (Figure 4-3). The hypersphere can be constructed using only the spectral data for the class of interest (hereafter referred to as the target data). SVDD aims to find a hypersphere of minimal volume containing all or most of the training data. Knowing the location of the centre of the hypersphere,  $a$ , and the hypersphere's radius,  $R$ , it is a relatively simple task for the classifier to test whether a new case (i.e., pixel) belongs to the class of interest (Foody et al., 2006). The distance of a new pixel,  $z$ , to the centre of the hypersphere is calculated. The pixel will be allocated to the class of interest when this distance ( $z$ ) is smaller than or equal to  $R$ . Because a closed boundary is placed around the data, only those data points that lie on the edge of the hypersphere (the support vectors) are used in defining the classifier (Figure 4-3). Thus, as with SVM classification, SVDD has the advantage of needing very few data during training and does not rely on restrictive assumptions regarding data distribution. For instance, the example data plotted in Figure 4-3 are described by only five support vectors.

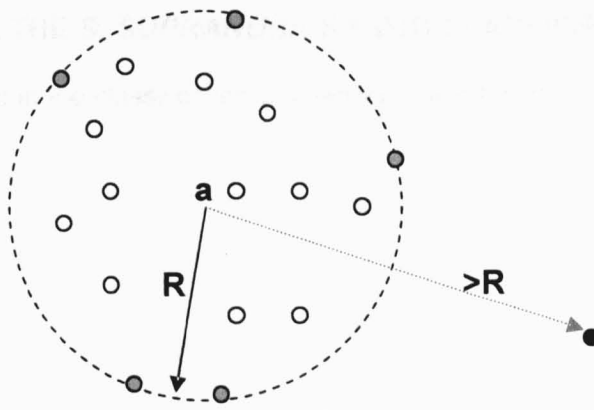


Figure 4-3 Classification using support vector data description (SVDD) and hyperspheres. A hypersphere is fitted around the class of interest; objects shaded grey are support vectors. The black object is classified as an outlier as it falls outside the hypersphere.

As with a standard SVM classifier, the SVDD can be extended to allow for non-linearly separable cases by mapping the data to a higher dimension using kernel functions. The hypersphere is a very rigid model and will generally provide a poor description of the target data (Tax and Juszczak, 2003). By applying kernel functions and varying the parameter values, it is possible to improve the fit between the hypersphere and the actual boundary of the data. The kernels and parameter values considered in this research are discussed in 4.3.2.

The SVDD requires only target training data. However, when data on outlier objects are available (i.e., pixels that do not belong to the class of interest), they can be incorporated into classifier training to improve the model description. By using two feature classes (target and outlier) the decision boundary is supported from two sides allowing a tighter boundary around target data to be calculated (Tax, 2001). As with SVM classifiers, two boundaries will be calculated to describe the distributions of the target and outlier data. The optimal hypersphere will be located equidistant between these boundaries. When large, representative samples of both the target and outlier data are available, a conventional two-class classifier may out-perform the SVDD (Tax, 2001). The SVDD is preferred when the outlier data are poorly sampled or unavailable. More detail on the incorporation of outlier data is given in 4.3.2.



### **4.3 CREATING THE S. *SUPRANUBIUS* SVDD CLASSIFIER**

The stages involved in the classification of imagery in this thesis are detailed in Figure 4-4.

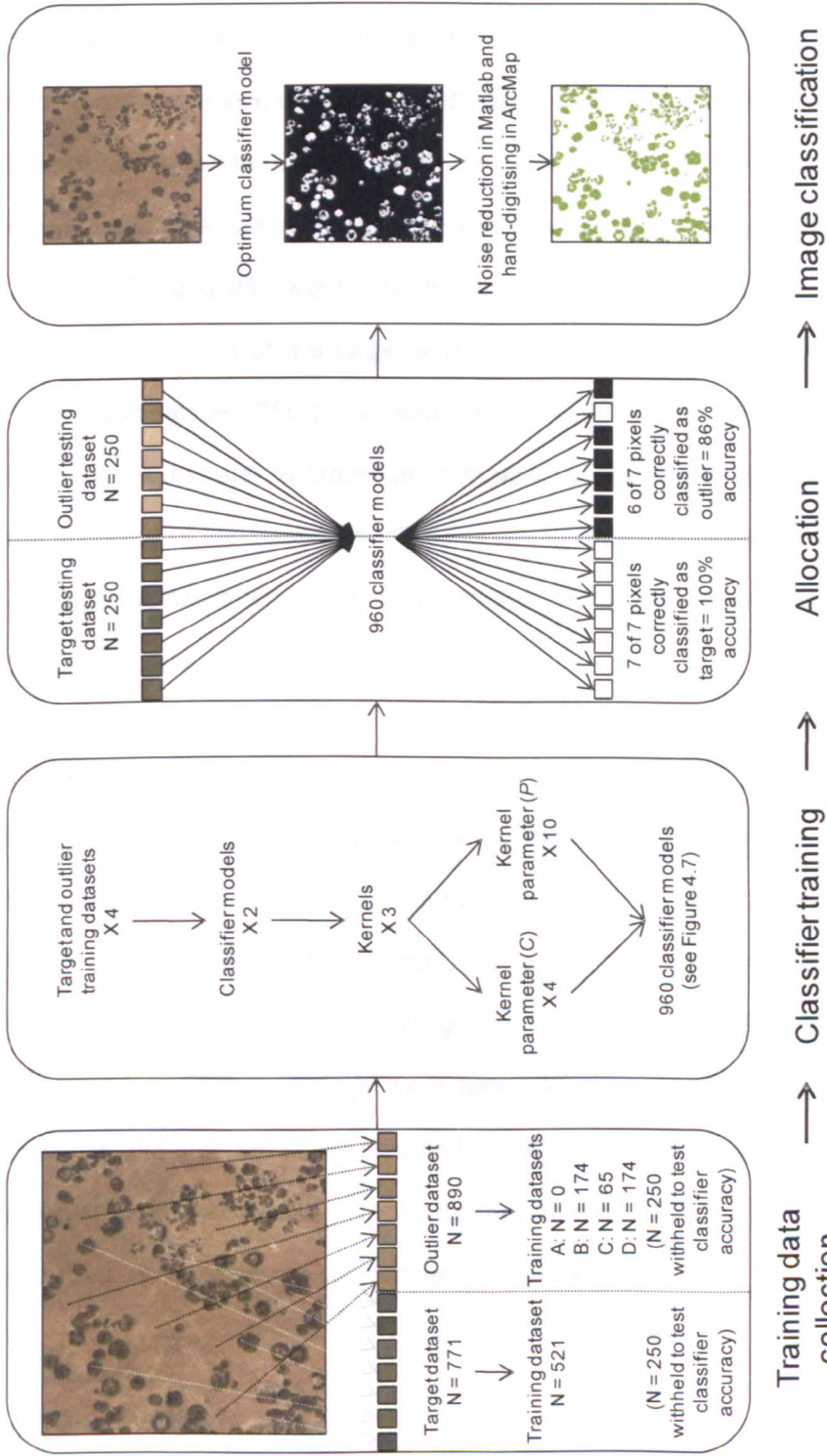


Figure 4-4 The stages involved in the classification of imagery in this thesis. Data on the class of interest is used to train the classifier. The classifier is then used to allocate a set of testing pixels to the classes with which they have most similarity. Following accuracy assessment, the optimum classifier model is used to produce the final image classification. 'N' refers to the sample size of pixels. More details on each of the four stages can be found in sections 4.3.1, 4.3.2, 4.3.3 and 4.4

#### **4.3.1 STAGE ONE: COLLECTING TRAINING AND TESTING DATA**

For accurate classification, a reliable sample of the spectral characteristics of pixels belonging to the classes of interest must be provided. Although larger sample sizes may provide a more comprehensive account of the spectral characteristics of the class, collecting such data can be time consuming. It is usually suggested that a minimum sample size of  $30p$  is collected per class, where  $p$  is the number of features (i.e., spectral bands; Mather, 2004). In the present analysis this corresponds to a minimum sample size of 90 pixels for both the target and outlier classes. Total sample sizes of 771 and 890 were collected for the target and outlier classes respectively. Summary statistics of the target and outlier data collected are shown in Table 4-1. From each dataset, 250 pixels were randomly selected (stratified: 50 pixels selected from each substrate) to create an independent testing dataset with which to assess the accuracy of the classifiers. The remaining data were used to train the classifiers, although, as explained below, only a subset of the remaining data were used to train the outlier class.

##### **Target data**

The geographical locations of 301 *S. supranubius* individuals were recorded during fieldwork in December 2007. Individuals were recorded from randomly located 60 x 6 m transects as well as opportunistically. Locations were logged on a hand-held GPS receiver (ProMark3) allowing the reading to stabilise for a minimum of 25 seconds. GPS recordings were taken as close to individual's rooting points as possible, although the sheer size and density of some individuals meant this was not always within two metres. The locations of individuals on the transects were also recorded using Cartesian co-ordinates (Figure 4-5).

Table 4-1 Summary statistics for the target (*S. supranubius*) and outlier (not *S. supranubius*) spectral data. Data were generated from digitised images of 13 transects completed in the field. N represents the number of pixels sampled,  $\bar{x}$  and s are the mean and the standard deviation of the spectral responses on the red (R), green (G) and blue (B) wavebands. Outlier data comprise the spectral responses of substrate and other species (where such data were available). No outlier data were generated for 1-C. The 'TOTAL' row provides summary statistics for the spectral data collected from each substrate.

Substrate	Transect	Target dataset						Outlier dataset					
		N	R	G	B	R	G	N	R	G	B	R	G
1	A	70	70.8	63.1	42.0	16.7	15.7	70	140.0	25.1	23.6	89.0	24.2
	B	70	69.9	62.8	41.9	11.1	11.2	70	151.9	18.2	19.2	102.9	19.4
	C	27	74.5	69.9	47.2	9.8	10.0	-	-	-	-	-	-
	TOTAL	167	71.1	64.1	42.8	13.6	13.3	140	145.9	22.7	24.7	95.8	23.0
	A	50	65.3	59.7	36.2	15.1	14.3	70	146.7	29.7	29.1	84.7	26.5
2	B	35	72.4	61.1	36.3	12.4	11.4	50	158.6	25.7	25.3	100.2	25.4
	C	70	85.8	68.3	41.7	12.1	11.1	70	161.1	22.9	22.8	102.5	21.1
	D	40	69.8	59.3	34.8	13.2	11.8	70	142.0	28.1	27.2	85.8	24.7
	TOTAL	250	72.7	61.5	36.2	15.8	13.1	330	149.5	27.5	27.0	91.5	25.1
	A	50	85.3	76.8	45.9	13.8	13.8	70	148.7	28.8	24.0	93.9	23.0
3	B	74	75.7	65.1	36.5	14.8	13.2	70	156.8	29.0	29.2	98.6	25.9
	C	45	73.4	63.7	38.0	10.4	10.0	70	131.0	37.0	34.8	80.0	31.8
	TOTAL	169	77.9	68.2	39.7	14.3	13.7	210	145.5	33.5	31.2	90.8	28.2
	A	45	69.2	60.2	35.5	12.2	11.7	70	145.2	23.6	21.2	81.9	19.0
	B	70	67.6	58.6	33.6	18.2	15.2	70	153.4	27.2	25.0	86.3	23.6
4	TOTAL	115	68.2	59.2	34.4	16.1	13.9	140	149.3	25.7	23.3	84.1	21.5
	A	70	71.9	60.8	38.2	10.9	10.1	70	137.6	24.8	24.4	83.5	23.9

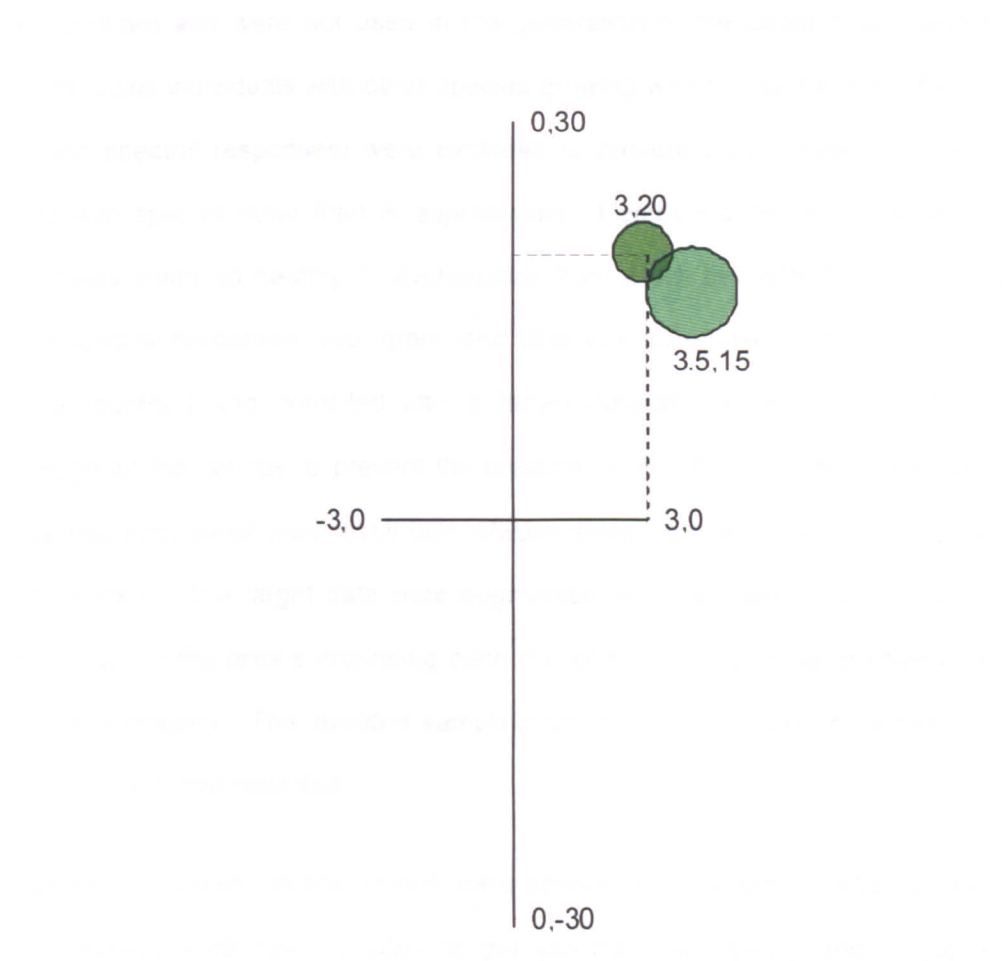


Figure 4-5 Illustration of the collection of target data (*S. supranubius*) from 60 x 6 m transects (axes not to scale). GPS and Cartesian location of individuals were taken if part of their canopy fell within the transect, or if their canopy merged with an individual within the transect.

The locations of the surveyed *S. supranubius* individuals were digitised and overlaid on the digital aerial imagery in ArcMap. Inspection of the data revealed unacceptable non-systematic error in the GPS readings when both the internal and external aerial were used. Consequently the readings were adjusted by converting the Cartesian recordings to a Mercator grid (i.e., polar to rectangular co-ordinate conversion). This procedure could only be used on those individuals where the Cartesian locations were recorded, removing all opportunistic data from the analysis. It is important that no data belonging to other classes (hereafter referred to as ‘outliers’) are inadvertently incorporated into the target training data. For the purposes of this analysis, outliers are pixels belonging to anything other than healthy *S. supranubius* individuals. Thus only the locations of healthy individuals (i.e., those with a dense canopy) were digitised. Individuals recorded as ‘dead’, ‘dying’ or as having a ‘sparse canopy’ were

not digitised and were not used in the generation of the target data. Similarly, *S. supranubius* individuals with other species growing within or surrounding their canopy (mixed spectral responses) were excluded to prevent contamination of the training data with species other than *S. supranubius*. Removal of these individuals left 143 positively identified healthy *S. supranubius* from which to create the target data set. The spectral responses (red, green and blue values) of pixels within each individual were recorded and compiled into a target dataset. Pixels were selected from throughout the canopy to prevent the creation of an artificially small data description (hypersphere) which may result from always selecting pixels from (e.g.) the centre of the canopy. The target data were augmented with data taken from *S. supranubius* individuals in the area surrounding each transect if they could be positively identified from the imagery. This avoided sampling too heavily from any one individual which may bias the data collected.

Statistical analyses (details below) were performed to determine whether there was any spatially systematic variation in the spectral responses of the *S. supranubius* individuals. Spatial variation in classification accuracy is common, although frequently not acknowledged (Foody, 2005). If marked spatial variation in the spectral response of *S. supranubius* existed, training a classifier on cases from across the site is likely to produce a high-volume hypersphere. When applied to unseen imagery, such a classifier would be likely to produce frequent misclassification errors, over-estimating the extent of the target class (i.e., false positive classification [commission] errors). The intra- and inter-substrate variation in *S. supranubius* spectral response was investigated. Because of violation of the normality assumption, all statistical analyses were non-parametric. The analyses were augmented with Cohen's *d* statistics to assess the practical significance of spectral differences. Cohen's *d* is given by

$$d = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(s_1^2 + s_2^2)/2}}$$

where  $\bar{x}$  and  $s$  are the mean and standard deviation of groups 1 and 2. Cohen (1988) provided rough guidelines for interpreting  $d$ : values between 0.2 and 0.5 represent a small effect size, values between 0.5 and 0.8 a medium effect size, and values greater than 0.8 a large effect size ( $d$  can exceed 1).

Analysis of the intra-substrate (i.e., between transects) spectral responses by Mann-Whitney U tests (Appendix B) suggested statistically and practically significant within-substrate spatial variation in the spectral response (red, green and blue) of *S. supranubius* on Substrates 2 and 3. Spatial variation in the spectral of *S. supranubius* in the green and blue colour bands was detected on Substrate 1 (Appendix B). Substrate 4 showed no spatial variation in spectral response, although only two transects were analysed. Analyses could not be performed on Substrate 5 as data were only collected from one transect.

Kruskal–Wallis analyses revealed significant differences in the substrate-specific spectral responses of *S. supranubius* on all three colour bands (R:  $\chi^2 = 28.316$ ,  $p < 0.001$ ; G:  $\chi^2 = 34.180$ ,  $p < 0.001$ ; B:  $\chi^2 = 37.751$ ,  $p < 0.001$  [3 d.p.], critical level reduced to 0.005 using Bonferroni correction). Assessing the pairwise differences with Mann–Whitney U-Tests suggests that the spatial variation in *S. supranubius* spectral response is not constant across the three wavebands (Tables 4-2). However, with large samples, especially when variances are small, statistically significant differences may be reported when little practical or theoretical difference exists.

Tables 4-2 Comparison of substrate-specific spectral responses of *S. supranubius* on the (a) red, (b) green and, (c) blue wavebands respectively. Differences in mean spectral intensity values ranging from 0 to 255 (column minus row). \*\* difference significant at the 0.001 level, \* difference significant at the 0.05 level (pairwise Mann–Whitney U-tests). All values reported to 2 d.p.

Substrate	1	2	3	4	5
1	-				
2	-3.86*	-			
3	-6.89**	-3.03	-		
4	2.79	6.65**	9.68**	-	
5	-0.85	3.01	6.04*	-3.64	-

(a) Red

Substrate	1	2	3	4	5
1	-				
2	1.13	-			
3	-4.10	-5.23	-		
4	-4.89**	3.76	11.99	-	
5	3.23	2.10	7.33	-1.66	-

(b) Green

Substrate	1	2	3	4	5
1	-				
2	4.89	-			
3	3.10	-1.78	-		
4	8.42**	3.54	5.32	-	
5	4.56**	-0.32	1.46	-3.86*	-

(c) Blue

Spatial variations in the spectral response of *S. supranubius* are strongest at the local scale and more spatially consistent at broader (i.e., substrate) scales. Spectral variation at the local scale may result from the shading effects of local topography, or localised variation in canopy colour. Without producing location-specific classifiers for individual areas of the field site, it is not possible to incorporate spectral variation at the local scale into the classification procedure. There are few consistent statistical differences in the spectral signature of *S. supranubius* on the different substrates, and almost no practical differences. In terms of the SVDD classifier this means that the position and volume of the hypersphere around the target data is relatively consistent



between substrates. Therefore, assuming the target data are spectrally separable from the outlier data, a single classifier trained on target data from all five substrates should be suitable for the classification of all substrates. The training dataset for the target class (*S. supranubius*) was created from the red, green and blue values of pixels ( $n = 521$ ) belonging to *S. supranubius* canopies from all five substrates.

### Outlier data

Recent research has shown that the intelligent selection of target training data believed to lie at the boundary of the data distribution can improve classification accuracy (Foody and Mathur, 2004; Mathur and Foody, 2008). It seems reasonable to suggest that when applying an SVM as a binary classifier, the intelligent selection of training cases from the outlier class could also influence classification accuracy. The optimal separating hypersphere will be located such that the distance to training cases of both classes is maximised. Thus if outlier training cases are spectrally dissimilar to the target training cases the SVDD will produce a higher volume hypersphere than if the outlier training cases were spectrally similar to the target cases (Figure 4-6). The hypersphere produced in the latter case will have greater generalisation ability than the former data description, which may misclassify outlier pixels as target.

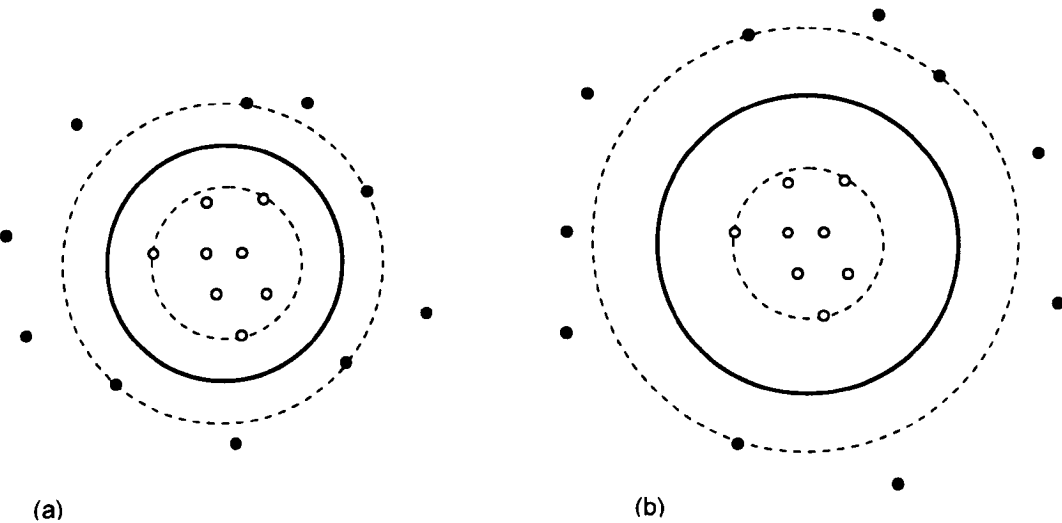


Figure 4-6 Intelligent training of outlier datasets for image classification. When outlier datasets are selected such that they are spectrally similar to the target data a higher generalisation ability is achieved (a) than when spectrally dissimilar outlier training cases are used (b).

*S. supranubius* may be expected to be most spectrally similar to other leguminous shrub species, notably *Adenocarpus viscosus*. Thus four dataset combinations were devised. All four used the same target training data. Dataset A used no outlier training data, whereas datasets B, C, and D used outlier training sets composed of substrate pixels, *A. viscosus* pixels, and all outlying classes (substrate and other species) respectively.

The *A. viscosus* training data (Dataset C) were collected by digitising the locations of positively identified *A. viscosus* individuals and randomly selecting and recording the spectral responses of pixels within their canopies. The quantity of outlier data collected was deliberately restricted so that classification effort was focused on the class of interest (*S. supranubius*). Following Sanchez-Hernandez et al. (2007), the quantity of outlier and target training data were maintained at a ratio of 1:3. Thus with 521 target training cases, 174 outlier training cases were used, with the exception of the *A. viscosus* outlier training dataset which comprised only 65 cases as only a few positive identifications and GPS recordings of these individuals were collected in the field. The spectral characteristics of each outlier training dataset are summarised in Table 4-3.

Table 4-3 The spectral characteristics of the outlier training data (mean[stdev]) in each of the four datasets. Dataset A used no outlier training data, whereas datasets B, C, and D used outlier training sets composed of substrate pixels, *A. viscosus* pixels, and all outlying classes (substrate and other species) respectively.

	Red	Green	Blue
<b>Dataset A</b>	-	-	-
<b>Dataset B</b>	153 (28)	127 (27)	95 (25)
<b>Dataset C</b>	113 (16)	99 (17)	65 (17)
<b>Dataset D</b>	146 (30)	121 (29)	90 (27)

To assess the spectral separability of the target and outlier training classes in each dataset the Bhattacharyya distance (B-distance) was calculated. The B-distance measures the similarity of two discrete probability distributions and is commonly used

to assess the separability of classes during classification (Schmidt and Skidmore, 2003). Larger B-distances indicate greater separability. The B-distance is defined as:

$$B = \frac{1}{8} [\mu_1 - \mu_2]^T \left[ \frac{\Sigma_1 + \Sigma_2}{2} \right]^{-1} [\mu_1 - \mu_2] + \frac{1}{2} \ln \left[ \frac{(\Sigma_1 + \Sigma_2)/2}{\sqrt{|\Sigma_1| |\Sigma_2|}} \right]$$

where  $\mu_1$  and  $\Sigma_1$  are the mean vector and the covariance matrix of class 1. The B-distance is closely related to the probability of accurate classification. There are no predefined thresholds for the degree of class separability represented by B-distances. Therefore, only the relative separability of classes can be assessed. The B-distance of each of the three outlier training datasets from the target training data set was calculated using the Bhattacharyya tool in Matlab (Cao, 2008). B-distances of 2.037, 1.0373 and 1.4294 were obtained for the separability of the target training data from the substrate only, *A. viscosus* only, and mixed outlier training data sets respectively. As anticipated the outlier dataset comprising the spectral responses of *A. viscosus* individuals only was the least separable from the *S. supranubius* class. Although the classification accuracy provided by all three outlier datasets was assessed, it was hypothesised that classifiers trained on the *A. viscosus* outlier data would provide the highest accuracy.

#### 4.3.2 STAGE TWO: CLASSIFIER TRAINING AND ALLOCATION

To ensure maximum classification accuracy was achieved, the performances of several classifier models were assessed. This section describes the classifiers that were employed. All classifications were performed in Matlab R2007a (MathWorks, 2007) using the dd\_toolbox (Tax, 2008). SVDD classification accuracy is largely determined by the selection of suitable kernels and parameter values (Ali and Smith-Miles, 2007). The literature provides little guidance on the selection of appropriate kernel functions and parameters. Most commonly, kernels are selected through a process of trial and error (Ali and Smith-Miles, 2007). This thesis uses the data intensive approach of cross-validation to determine the optimum parameters and

kernel functions. Following Sanchez-Hernandez et al. (2007) three kernels were assessed. These were the Gaussian radial basis function,

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(\frac{-\|\mathbf{x}_i - \mathbf{x}_j\|^2}{p^2}\right)$$

the polynomial kernel,

$$k(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i \cdot \mathbf{x}_j + 1)^p$$

and the exponential kernel

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(\frac{-\|\mathbf{x}_i - \mathbf{x}_j\|}{p^2}\right)$$

For each kernel, the value of the parameter  $p$  must be pre-defined. In the Gaussian radial basis kernel and exponential kernel,  $p$  controls the width of the kernel, whereas in the polynomial kernel function  $p$  determines the order of the kernel. Following Sanchez-Hernandez et al. (2007), the performance of each kernel was assessed using  $p$  values of 1 – 10 inclusive. In addition to  $p$ , the user can define the value of a second parameter,  $C$ , also known as the rejection error. This parameter determines the fraction of the target data that is allowed to lie on the ‘wrong side’ of the data description (i.e., extreme values), and thus allows for outliers in the training samples (Sanchez-Hernandez et al., 2007). In this way it enables the user to control the trade-off between training error and model complexity. However, if too many target training points are allowed to lie beyond the data description (i.e., a large  $C$ ) the optimal hypersphere will produce an over-fitted model with limited generalisation ability (Foody and Mathur, 2006).  $C$  values of 0.1, 0.01, 0.001 and 0.0001 were used (following Sanchez-Hernandez et al., 2007).

As well as testing the SVDD classifier, the performance of the incremental SVDD (INCSVDD) was assessed, as preliminary analyses (not reported) indicated that this technique may produce higher classification accuracies than the standard SVDD. Including the training data combinations described in 4.3.2 this resulted in 960 different classifiers (Figure 4-7).

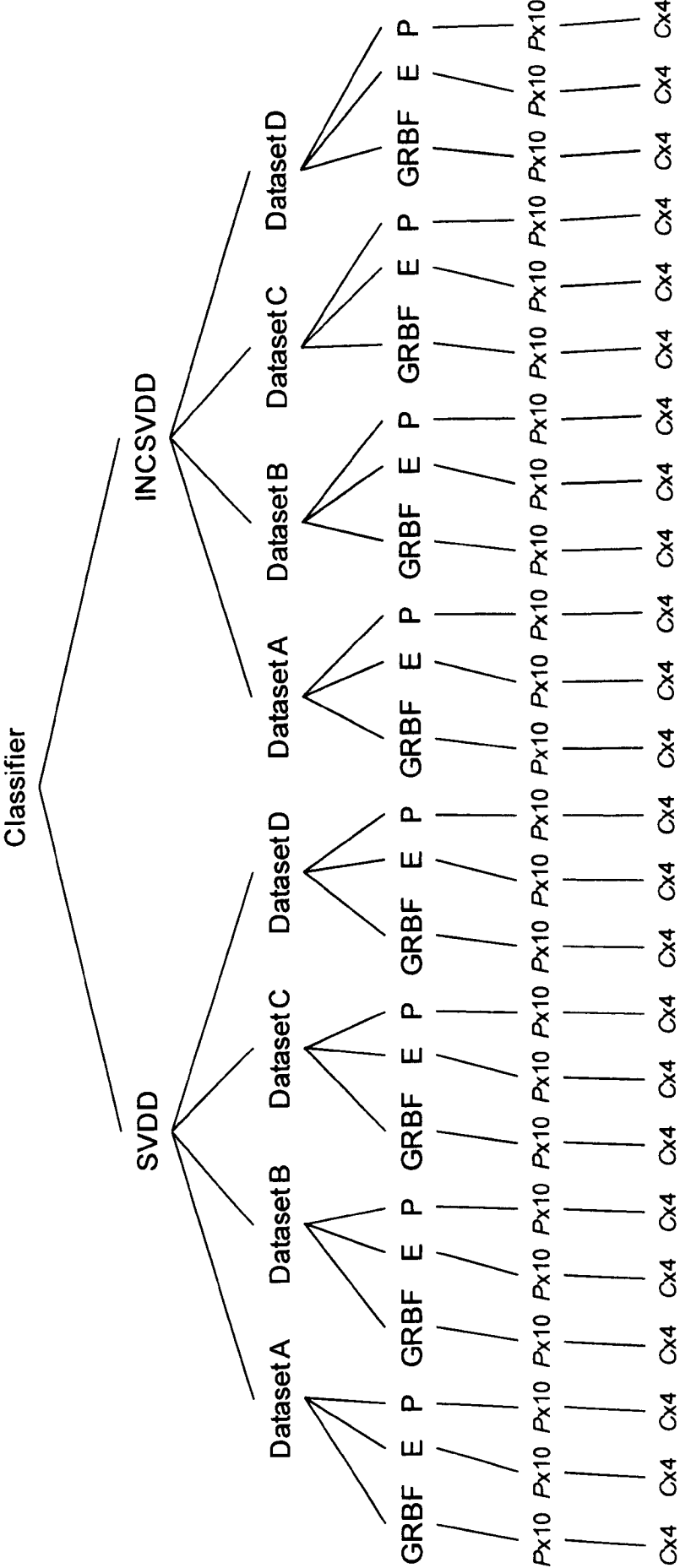


Figure 4-7 The 960 classifier models assessed. The three kernels are abbreviated to GRBF (Gaussian radial basis function), E (exponential) and P (polynomial). See text for description of the parameters  $p$  and  $C$ .

### 4.3.3 STAGE THREE: ACCURACY ASSESSMENT

#### Quantitative assessment of classification accuracy

The accuracy of the thematic map produced by a classification is of great importance, especially if the imagery is to be used in further analysis. Two datasets comprising data known to belong to the target and outlier classes were created (hereafter referred to as the 'testing' datasets). Fifty target testing cases and fifty outlier cases were drawn randomly from each substrate, generating two testing datasets of 250 cases. The pixels used in testing were independent from the training data to avoid any biases in the confidence of classifier accuracy. Because of the small number of *A. viscosus* data, the outlier testing data consisted only of known substrate cases. The accuracy provided by each classifier model is calculated as the proportion of testing pixels (both target and outlier) correctly classified (Foody, 2002).

The testing dataset is small relative to the quantity of data that is being classified. Therefore it is important that the classifier selected has both a high accuracy and good generalisation capacity. However, very high classification accuracies as assessed from the testing data may be a consequence of over-fitting to the training data. Such classifiers are unlikely to classify large quantities of unseen data with high accuracy. Consequently, instead of selecting the highest performing classifier as measured by the testing datasets, statistical analyses were performed to investigate the average response of classifiers to changes in structure (e.g., changes in parameter value). Initially these analyses focused on the major elements of classifier structure; model (i.e., SVDD vs. INCSVDD), kernel and dataset. Selecting criteria that consistently outperformed competing criteria would result in greater confidence in the final classifier. The effect of both parameters ( $p$  and  $C$ ) will be dependent upon selection of other criteria. Therefore the selection of values for  $p$  and  $C$  was made after the selection of classifier model, kernel and dataset.

As the data were measured as proportions, the accuracy values were arcsine transformed prior to analysis. Analyses for paired and related data were used so that

the independent effect of each classifier variable (model, kernel, dataset,  $p$ , and  $C$ ) on classification accuracy could be assessed. Because of normality violation, the non-parametric Friedman's test and Wilcoxon signed rank tests for related samples were employed. Because of large sample sizes, which can overemphasise small effects, the results of the Friedman's test were augmented with calculations of the average classification accuracy differences. Other studies have investigated the performance of SVM classifiers of differing structure (Cortes and Vapnik, 1995; Huang et al., 2002; Sanchez-Hernandez, 2006; Ali and Smith-Miles, 2007; Sanchez-Hernandez et al., 2007). However, as the optimum classifier structure will be data-dependent, direct comparison of the classifier structure selected in this research with the classifier structures used in previous studies will not be made.

Of the 960 classifiers tested, computational errors prevented 69 from producing output. Of the remaining 891 classifiers, ones providing an overall accuracy of  $\geq 85\%$  were deemed to be 'high performance' classifiers. An example of the classification accuracies of the various models is provided in Appendix C. 286 models produced overall classification accuracies of 85% or greater. Seven classifiers achieved overall accuracies of 98% or greater. 80% of the high performing classifiers used the Incremental SVDD model. On average INCSVDD models produced c. 13% higher classification accuracy than their SVDD counterparts (Wilcoxon signed rank test,  $Z = -14.592$ ,  $p < 0.0005$  [3dp]).

Out of the 286 high performing classifiers, almost equal proportions used the four different dataset combinations; datasets A, B, C and D were used by 77, 75, 64 and 70 of the higher performance classifiers respectively. This indicates that the addition of outlier datasets, and the composition of those datasets, was of limited importance. However, statistical analysis indicated that the different datasets may influence the classification accuracies achieved (Friedman's test,  $\chi^2 = 15.528$ ,  $p = 0.001$  [3dp]). When all other classifier variables are kept constant, selecting dataset A (no outlier training data) produced routinely higher classification accuracies than datasets B and D (Wilcoxon signed rank test,  $Z = -2.089$ ,  $p = 0.037$  and  $Z = -3.200$ ,  $p = 0.001$  [3dp])



respectively). Dataset C produced an intermediate level of accuracy which was not significantly different from the accuracy achieved by dataset A ( $Z = -1.845$ ,  $p = 0.065$ ). Yet despite statistical significance, classifiers using dataset A produced classifiers that were on average  $< 1.5\%$  more accurate than like-for-like classifiers using the other datasets. No statistically significant differences in classification accuracy were found when comparing classifiers using the three different outlier training datasets (datasets B-D). Therefore, there appeared to be no notable increase in classification accuracy when outlier training data were included, concurring with Sanchez-Hernandez et al. (2007). Previous studies (Foody and Mathur, 2004; Mathur and Foody, 2008) have suggested that the intelligent selection of target training data can significantly improve classifier performance. However, the use of *A. viscosus* outlier training data (dataset C) did not show improved classifier accuracy over other outlier training datasets. In explanation, it is suggested that the *A. viscosus* outlier training data shared too little feature space with the target training data (Figure 4-8). Such outlier training data would only influence the data description of the target class in the small area of feature space in which the outlier and target distributions overlap. As such, it was decided to perform the classifications without outlier training data as these classifiers produced slightly higher classification accuracies and will require less input to train.

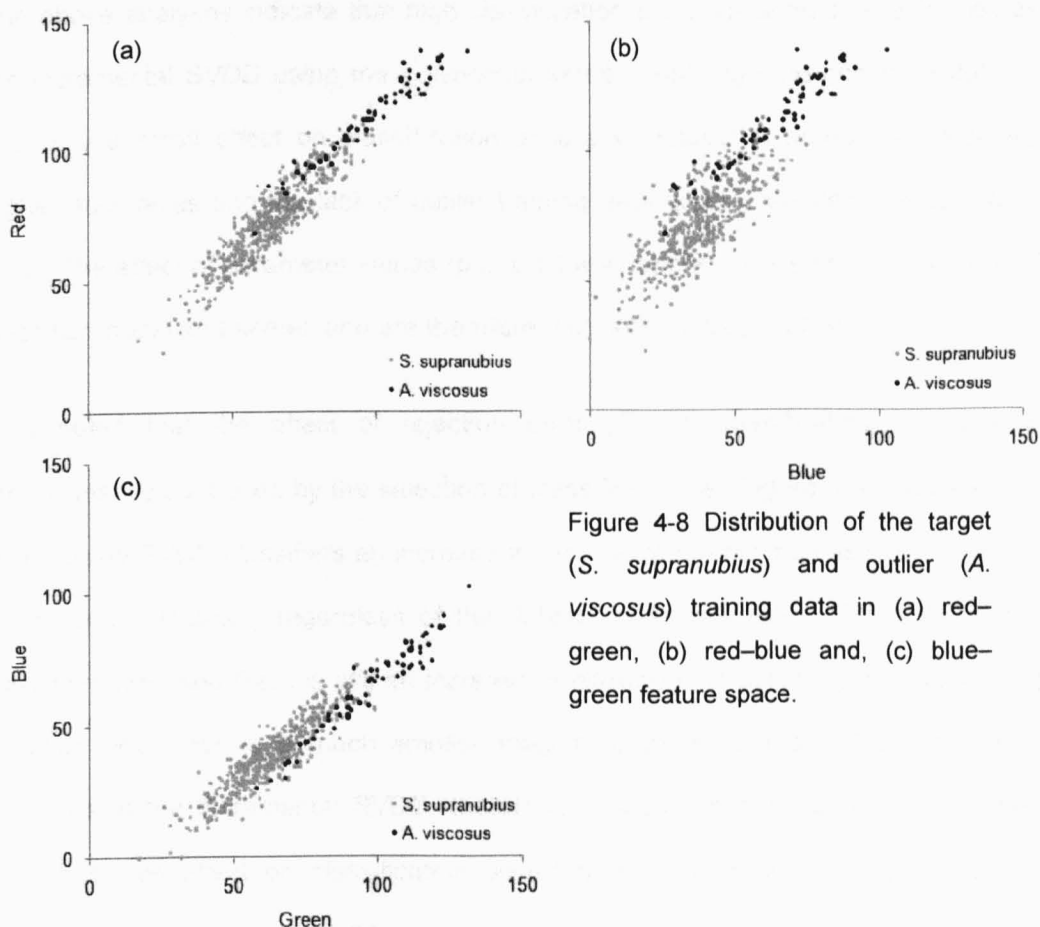


Figure 4-8 Distribution of the target (*S. supranubius*) and outlier (*A. viscosus*) training data in (a) red–green, (b) red–blue and, (c) blue–green feature space.

All three kernel functions were capable of producing high accuracies, although the polynomial was used by 145 of the high performance classifiers, compared to 75 and 66 using the Gaussian RBF and exponential kernel respectively. Pair-wise Wilcoxon tests revealed statistically significant differences between all three kernels (Gaussian-Polynomial:  $Z = -11.098$ ,  $p < 0.0005$ ; Gaussian-Exponential:  $Z = -4.983$ ,  $p < 0.0005$ ; Polynomial-Exponential:  $Z = -11.080$ ;  $p < 0.0005$  [3dp]) with the polynomial kernel providing the highest overall accuracies and the exponential kernel providing the lowest overall accuracies. On average, classifiers using the polynomial kernel produced classification accuracies 13.65% and 13.87% higher than comparable classifiers using the Gaussian and exponential kernels respectively. Selecting a Gaussian kernel over an exponential kernel, however, only provided an increase in classification accuracy of 0.22%. These results concur with Sanchez-Hernandez et al. (2007).

The above analyses indicate that high classification accuracy should be achieved by the incremental SVDD using the polynomial kernel. Although the choice of dataset had only a small effect on classification accuracy, dataset A produced marginally higher accuracies and the lack of outlier training data greatly reduced computational time. The effect of parameter values ( $p$  and  $C$ ) will depend largely upon the choice of classifier model and kernel, and are therefore discussed in this context.

It is noted that the effect of rejection error ( $C$ ) on classification accuracy is predominantly controlled by the selection of classifier model (Figure 4-9, Figure 4-10). When using SVDD classifiers an increase in the rejection error increases the average classification accuracy regardless of the dataset or kernel used. When INCSVDD classifiers are used there is still an increase in classification accuracy with increasing rejection error, but of a much smaller magnitude (maximum of 5.8% difference). When using the incremental SVDD, therefore, it seems that the choice of rejection error has little effect on classification accuracy, concurring with Belousov et al.'s (2002) conclusions that SVM based classification displays a large degree of robustness to variation in parameter values. It is known, however, that large values of  $C$  can result in over-fitting of the classifier to the target data which can reduce the classifier's generalisation capacity (Foody and Mathur, 2006). Thus, when there is no noteworthy difference in classification accuracy (as with the INCSVDD), lower values of  $C$  are preferable.

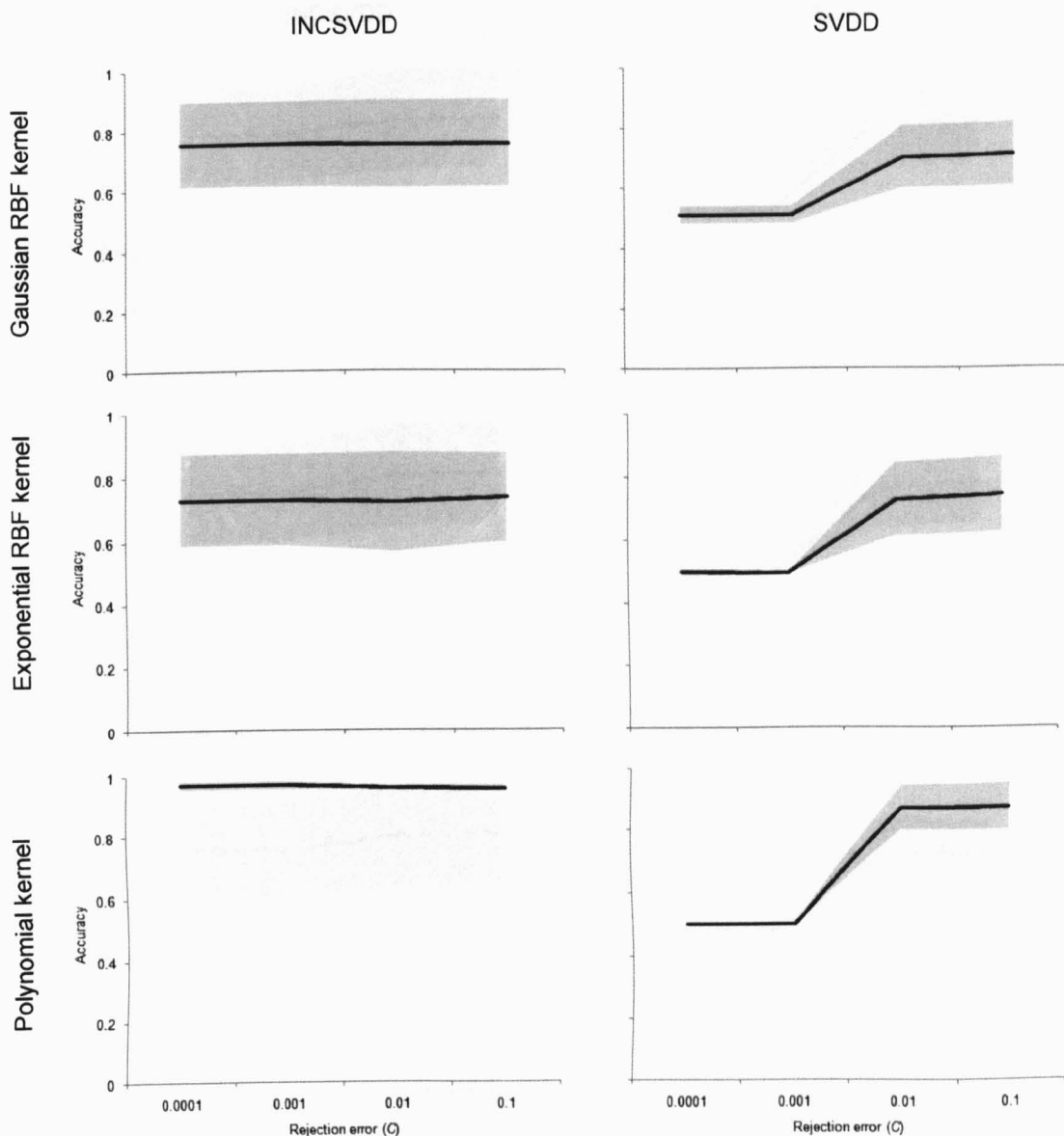


Figure 4-9 The change in overall accuracy (measured as a proportion) as the rejection error ( $C$ ) increases from 0.0001 to 0.1. The results are divided by kernel (Gaussian RBF, Exponential RBF and polynomial) and by classifier model (INCSVDD and SVDD). The black line shows the mean accuracy and the shaded area shows the standard deviation ( $n = 40$ ).

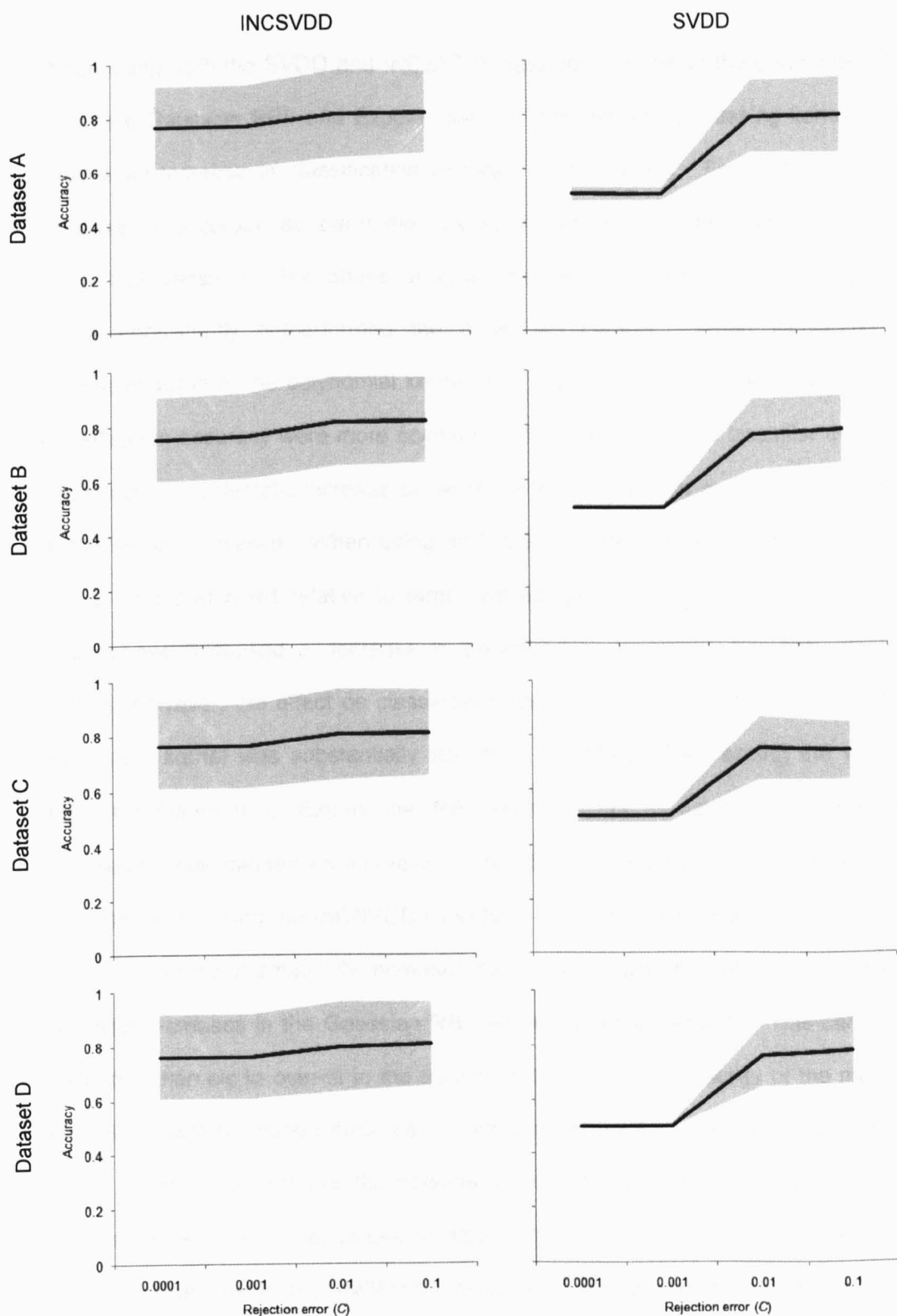


Figure 4-10 The change in overall accuracy (measured as a proportion) as the rejection error ( $C$ ) increases from 0.0001 to 0.1. The results are divided by dataset (A, B, C and D) and by classifier model (INCSVDD and SVDD). The black line shows the mean accuracy and the shaded area shows the standard deviation ( $n = 30$ ).

When using both the SVDD and INCSVDD classifier, increasing the parameter value ( $p$ ) of the Gaussian RBF and Exponential RBF kernels (i.e., increasing kernel width) causes an increase in classification accuracy (Figure 4-11). For both kernels the increase in accuracy as parameter values increase is greatest when using the INCSVDD classifier. The above analyses revealed, however, that the polynomial kernel consistently outperformed the other two kernels. When increasing the parameter value of the polynomial kernel (i.e., the order of the kernel), the effects of classification accuracy were more complex. When using a SVDD classifier there was seemingly no systematic increase or decrease in classification accuracy as the order of the kernel increased. When using an INCSVDD classifier parameter values of 1 and 2 underperformed relative to larger values, whereas increasing the parameter value above 3 caused a decrease in classification accuracy. For both classifier models, however, the effect on classification accuracy of increasing the order of the polynomial kernel was substantially less than the effect of increasing the width of either the Gaussian or Exponential RBF kernels. For all datasets increasing the parameter value caused an increase in classification accuracy. This increase was strongest when using the INCSVDD classifier. Increasing the parameter value of the polynomial kernel dramatically increases the dimensionality of feature space relative to similar increases in the Gaussian RBF and exponential kernels. This can cause polynomial kernels to over-fit to the training data, reducing the ability of the model to accurately classify unseen data (i.e., a reduced generalisation capacity; Cortes and Vapnik, 1995). To minimise the potential of classifier over-fitting, it was decided to limit the range of parameter values considered to between 1 and 6. This increase was strongest when using the INCSVDD classifier. Increasing the parameter values beyond  $p = 6$  only resulted in minimal improvements in classifier accuracy (Figure 4-12).

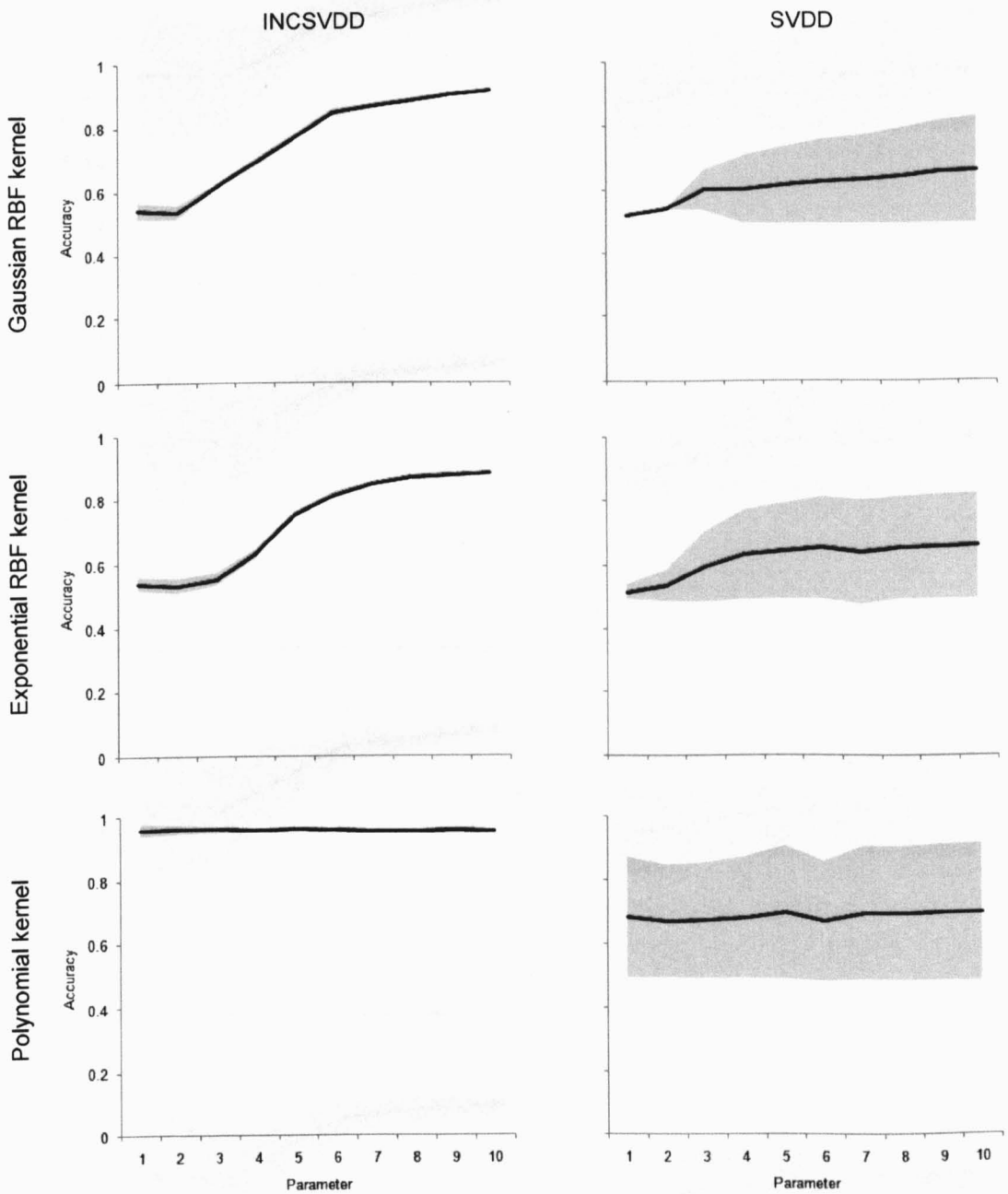


Figure 4-11 The change in overall accuracy (measured as a proportion) as the parameter value ( $p$ ) increases from 1 to 10. The results are divided by kernel (Gaussian RBF, Exponential RBF and polynomial) and by classifier model (INCSVDD and SVDD). The black line shows the mean accuracy and the shaded area shows the standard deviation ( $n = 16$ ).

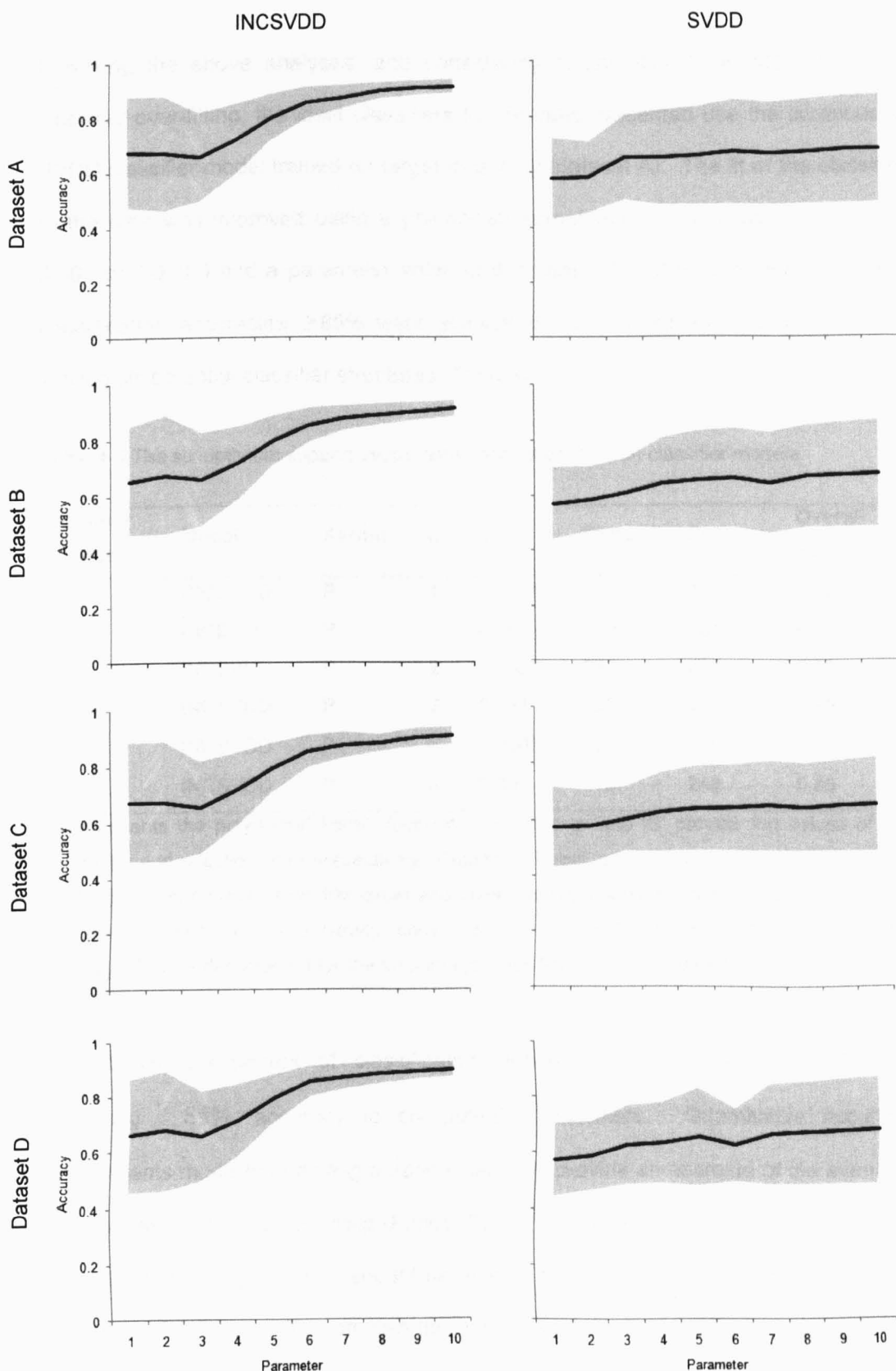


Figure 4-12 The change in overall accuracy (measured as a proportion) as the parameter value ( $p$ ) increases from 1 to 10. The results are divided by dataset (A, B, C and D) and by classifier model (INCSVDD and SVDD). The black line shows the mean accuracy and the shaded area shows the standard deviation ( $n = 12$ ).



Following the above analyses, and considering computation time and the risk of classifier over-fitting, the ideal classifiers for the data presented use the incremental SVDD classifier model trained on target data only (dataset A). The fit of the classifier to the data was improved using a polynomial kernel with a low rejection error (i.e., 0.001 or 0.0001) and a parameter value of 6 or less. The 286 classifiers achieving classification accuracies  $\geq 85\%$  were eliminated until the above criteria were met, leaving six potential classifier structures (Table 4-4).

Table 4-4 The six optimum support vector data description (SVDD) classifier models.

Classifier number	Model	Kernel	p	C	Target	Outlier	Overall accuracy
1	INCSVDD	P	1	0.0001	250	224	0.95
2	INCSVDD	P	1	0.001	250	224	0.95
3	INCSVDD	P	2	0.0001	250	225	0.95
4	<b>INCSVDD</b>	<b>P</b>	<b>2</b>	<b>0.001</b>	<b>250</b>	<b>225</b>	<b>0.95</b>
5	INCSVDD	P	6	0.0001	178	248	0.85
6	INCSVDD	P	6	0.001	178	248	0.85

'P' represents the polynomial kernel function. Columns 'p' and 'C' provide the values of the parameter and rejection error respectively. Columns 'Target' and 'Outlier' contain the number of pixels classified correctly from the target and outlier testing datasets (each containing a total of 250 pixels). The 'Overall accuracy' column shows the total proportion of pixels classified correctly. The model selected for the final image classification is highlighted in bold.

Quantitative assessment of classification accuracy reduced the 286 classifiers producing  $\geq 85\%$  accuracy to six potential classifiers. Quantitative accuracy assessments made from testing datasets can only provide an estimate of the eventual 'correctness' of the thematic map (Foody, 2002). The final classifier should have high generalisation ability in that it should be able to classify previously unseen data to a high level of accuracy. To compare the performance and generalisation ability of the six remaining classifiers, each model was applied to areas of 4 ha (200 x 200 m) randomly selected from each of the five substrates. The classified images were displayed in raster format in ArcMap 9.2 with a cell size of 0.26 x 0.26 m. By comparing the images produced by the six different classifier models it was possible to assess the level of disagreement between the classifiers.

When applied to the 4 ha sample plots the differences in classifications provided by the six optimum classifiers were minimal (Table 4-5). Classification disagreements are primarily attributed to changes in  $p$ , with changes in  $C$  having little impact. Classifiers with a parameter value of 1 (i.e., Classifiers 1 and 2; Table 4-4) had the greatest disagreement with the remaining four classifiers, most notably Classifier 6 (Table 4-5). The largest disagreement between any two classifiers occurred on Substrate 2. However, this disagreement totalled only 6714 cells, corresponding to only 1.17% of the 4 ha image classified (Table 4-4b). Investigation of the classified images suggested that classifiers with  $p = 1$  frequently mis-classified areas of substrate as belonging to the target class. Furthermore, classifiers with  $p = 1$  also incorrectly classified many shaded areas of *S. supranubius* canopy as 'outlier'. These errors are despite the seemingly high accuracy of Classifiers 1 and 2 (Table 4-4) suggesting that, when applied to the current data, incremental SVDD classifiers with polynomial kernels of order 1 have poor generalisation ability. This concurs with other studies concluding that linear class boundaries are rare, with most studies preferring to use non-linear solutions (e.g., Sanchez-Hernandez et al., 2007). On all substrates the disagreement between classifiers with parameter values of 2 and 6 (Classifiers 3–6; Table 4-4) was in the classification of cells as 'target' by the former which were classified as 'outlier' by the latter (Figure 4-13). This is consistent with Table 4-4 which suggests that Classifiers 5 and 6 have a low accuracy on the target class. Thus, the two classifier types identify the same objects as being *S. supranubius* individuals, but models with a parameter value of 6 produce slightly lower estimates of canopy cover (Figure 4-13). Given the increased potential for over-fitting as the order of the polynomial kernel increases, and the minimal differences between classifiers with parameter 2 and 6, a parameter value of 2 was deemed most appropriate given the data. With rejection errors of  $C = 0.0001$  and  $C = 0.001$  producing identical results when  $p = 2$ , the selection of  $C$  was largely arbitrary. A rejection error of 0.001 was chosen to allow for a greater number of outliers in the training data. This gave a final classifier that used the incremental INCSVDD model, trained on target data only using the polynomial kernel of the second order and a rejection error of 0.001.

Table 4-5 Comparison of six optimum support vector data description (SVDD) classifier models. Each classifier was applied to a randomly selected area of 4 ha on each of the five substrates. Values in the top right indicate how many cells were classified as 'outlier' by the column classifier, but classified as 'target' by the row classifier. Values in the lower left indicate how many cells were classified as 'target' by the column classifier, but as 'outlier' by the row classifier. All values are out of a total 570025 cells.

Classifier number	1	2	3	4	5	6
1	-	0	856	856	837	837
2	0	-	856	856	837	837
3	784	784	-	0	0	0
4	784	784	0	-	0	0
5	1355	1355	590	590	-	19
6	1355	1355	571	571	0	-

(a) Classification disagreements on Substrate 1

Classifier number	1	2	3	4	5	6
1	-	0	3583	3583	3510	3578
2	0	-	3583	3583	3510	3578
3	2600	2600	-	0	0	0
4	2600	2600	0	-	0	0
5	3136	3136	609	609	-	68
6	3136	3136	541	541	0	-

(b) Classification disagreements on Substrate 2.

Classifier number	1	2	3	4	5	6
1	-	0	1410	1410	1389	1402
2	0	-	1410	1410	1389	1402
3	761	761	-	0	0	0
4	761	761	0	-	0	0
5	1334	1334	594	594	-	13
6	1334	1334	581	581	0	-

(c) Classification disagreements on Substrate 3

Classifier number	1	2	3	4	5	6
1	-	0	2300	2300	2220	2299
2	0	-	2300	2300	2220	2299
3	938	938	-	0	0	0
4	938	938	0	-	0	0
5	1116	1116	258	258	-	79
6	1116	1116	179	179	0	-

(d) Classification disagreements on Substrate 4

Classifier number	1	2	3	4	5	6
1	-	0	2902	2896	2782	2887
2	0	-	2902	2896	2782	2887
3	1033	1033	-	11	9	9
4	1024	1024	8	-	7	7
5	1534	1534	630	631	-	105
6	1534	1534	525	526	0	-

(e) Classification disagreements on Substrate 5

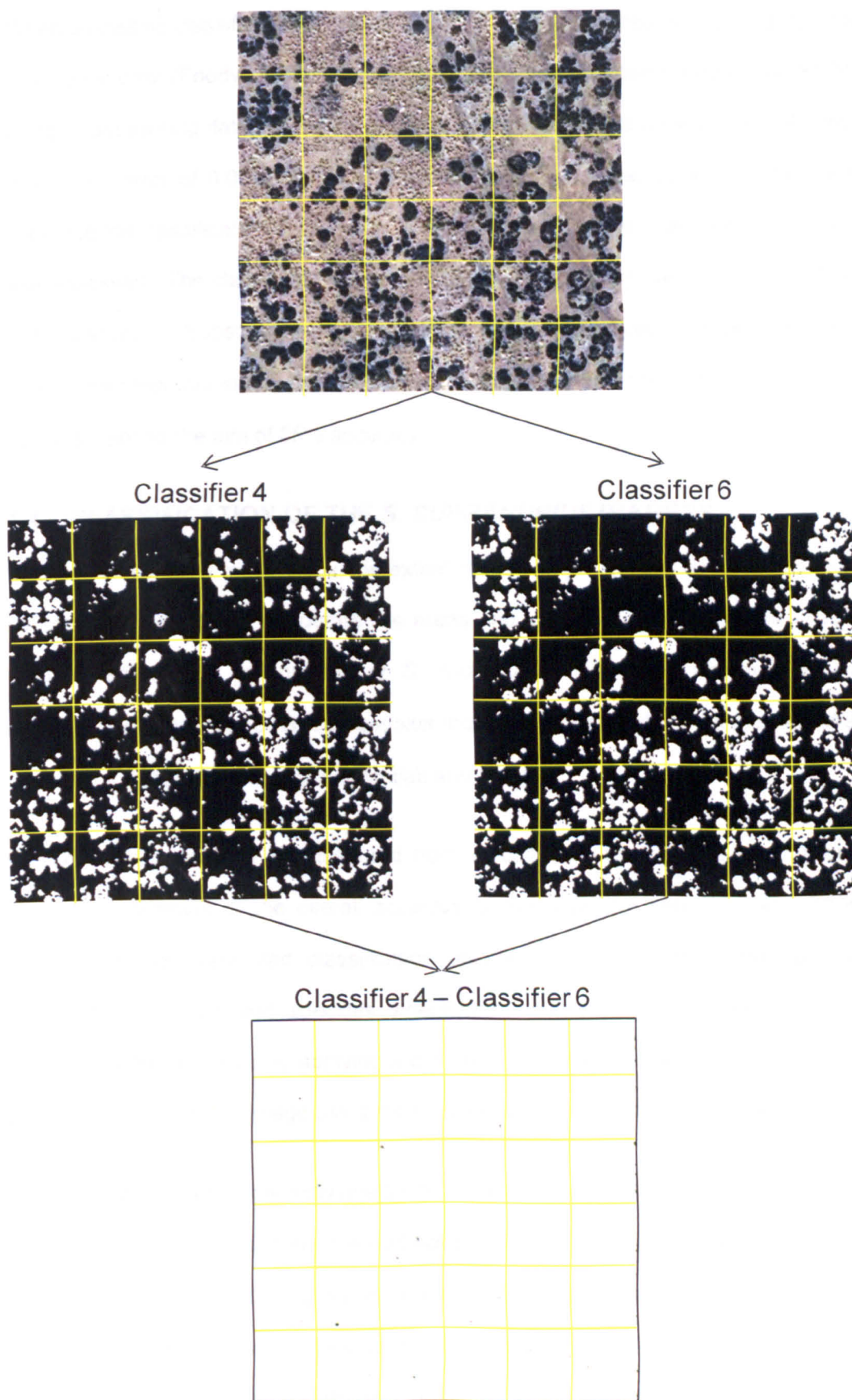


Figure 4-13 Comparison of the classification of a 4 ha sample area on Substrate 1 using optimum Classifiers #4 and #6 (Table 4-4). Subtracting the images produced by the two classifiers reveals minimal disagreements in the resultant maps.

When assessing classifier performance, global measures of accuracy can hide spatial variation in error (Foody, 2002). The performance of the selected classifier (INCSVDD using target training data only and a polynomial kernel with a parameter value of 2 and a rejection error of 0.001) on each of the five substrates was assessed. On each substrate the classification of the 50 target and 50 outlier cases in the testing datasets was assessed. The classification accuracy of the target testing data was 100% on all five substrates. Substrate 1 also had 100% classification accuracy on the outlier class. The other four substrates showed some classification errors in the outlier class, but all exceeded the aim of 85% accuracy.

#### **4.4 CLASSIFICATION OF THE *S. SUPRANUBIUS* IMAGERY**

Because of lava flow morphology, the extent of the sample window classified on each substrate varied. Table 4-6 details the areas and locations of the classified sample windows as well as the number of *S. supranubius* individuals identified by the classifier. The number of *S. supranubius* individuals classified was calculated after the images were post-processed in Matlab and ArcMap.

Classification accuracies as measured from independent testing datasets can only provide an estimate of the overall accuracy of the resultant thematic map. Thus, despite the high estimated classification accuracy of the selected classifier, the classified imagery showed 'speckles' where outlier pixels were misclassified. These 'speckles' were removed by applying a circular morphological filter with a radius of 2 pixels (c. 0.52 m) to the image using the Image Processing Toolbox in Matlab.

During image classification, individuals with adjacent canopies were often classified as a single object. Consequently, merged canopies were manually separated by hand-digitising in ArcMap. During visual analysis of the imagery it was often unclear whether small objects were juvenile *S. supranubius*, or large rocks and boulders. During the December 2007 field trip, very few individuals were observed to have canopy areas of  $< 1 \text{ m}^2$ . Therefore, all classified objects below this size were removed in ArcMap. This action was deemed appropriate because of the large area and

number of individuals mapped. With this quantity of data it was considered more important to reduce the likelihood of commission errors (wrongly classifying a shadow or rock as a *S. supranubius* individual), while accepting some errors of omission (failing to identify a *S. supranubius* individual), i.e., the set of objects identified as *S. supranubius* should have a high probability of being *S. supranubius*. Consequently, any classified objects greater than 1 m<sup>2</sup> that were deemed unlikely to be *S. supranubius* individuals on the basis of their shape and spectral response were manually removed from the dataset.

Table 4-6 Summary information of the classified images. The morphology of Substrate 1 prevented a single large area being classified. Therefore two separate areas were classified.

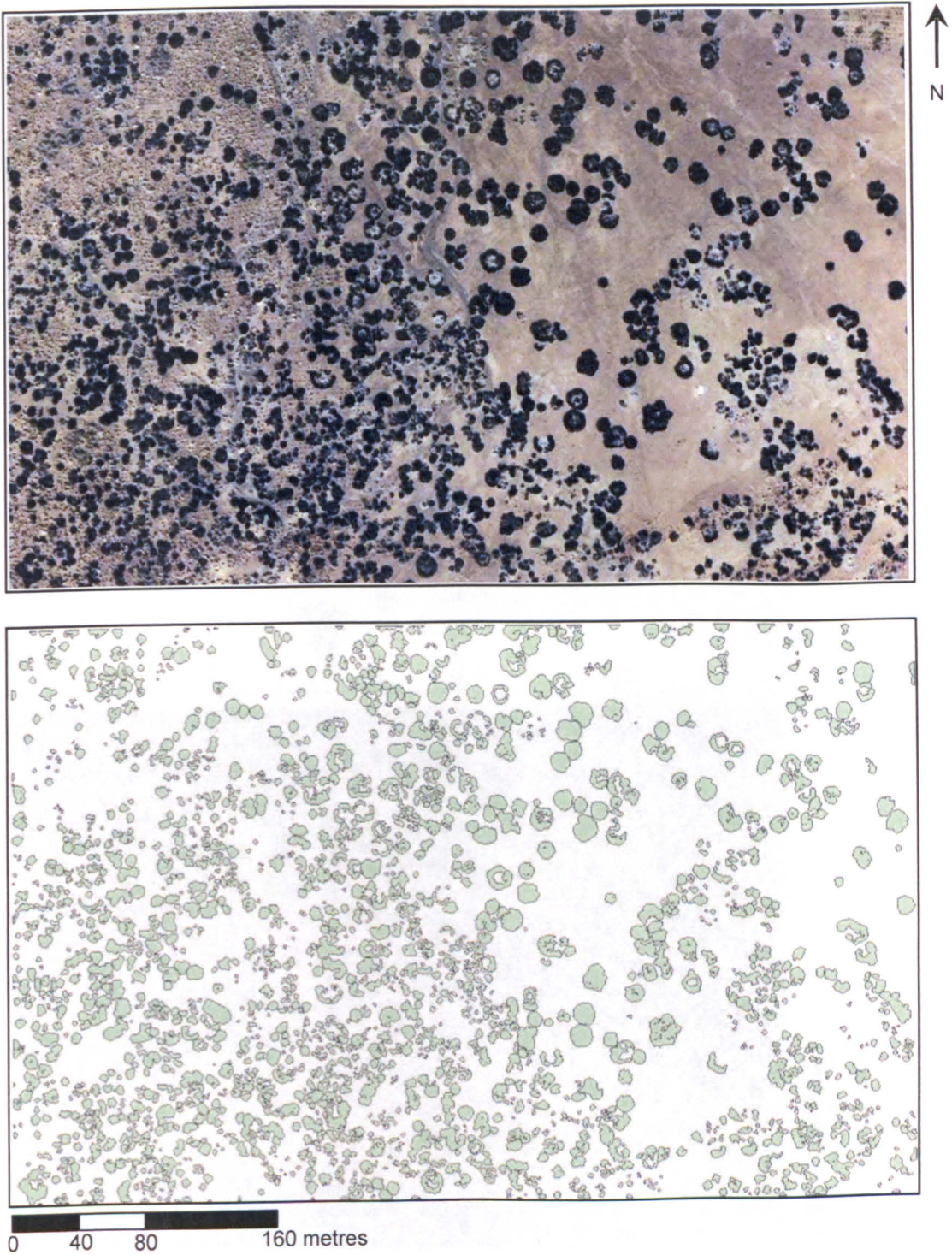
Substrate	ULX / ULY	LRX / LRY	Area (ha)	<i>S. supranubius</i>
1a	340150 / 3124050	340700 / 3123700	19.25	1949
1b	340080 / 3125350	340330 / 3125000	8.75	1029
2	340400 / 3125100	341100 / 3124400	49.00	4880
3	341710 / 3126780	342110 / 3126130	26.00	1967
4	342600 / 3124300	343300 / 3123600	49.00	6478
5	342585 / 3124995	342985 / 3124745	10.00	1174

Figure 4-14 shows the final classified map of *S. supranubius* on Substrates 1 to 5. The classification of *S. supranubius* individuals was verified in the field. Using stratified random sampling, 54 quadrats (each 0.25 ha) covering a total area of 13.5 ha were located within the focal plots (see Table 4-6). Twelve quadrats were located in each of Substrates 1 to 4, and six quadrats were located on Substrate 5. Within these quadrats there were a total of 1572 objects classified as *S. supranubius*. Of these objects, 1407 were verified as being *S. supranubius* individuals. Therefore, 165 objects were incorrectly classified as *S. supranubius*. The majority of the incorrectly classified objects (n = 106) were identified as rocks and accompanying shadows. The remaining objects belonged to other species, either *Pterocephalus lasiospermus* or *Adenocarpus viscosus*. A further 35 *S. supranubius* individuals were identified that were not identified in the image classification. Since the objective of the classification was to reduce the probability of commission errors (i.e., the set of individuals identified as *S. supranubius* should have a high probability of being *S.*

*supranubius* individuals) the error matrices focus on user's accuracy. The overall user's accuracy was 90%, although this varied between 84% and 92% on the various substrates (i.e., commission errors of between 8% and 16%; Table 4-7).



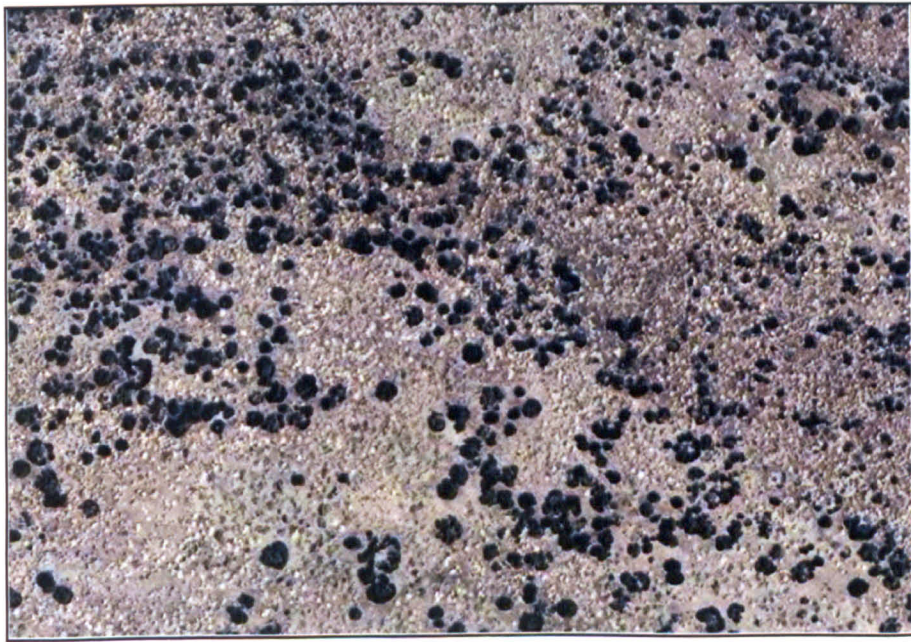
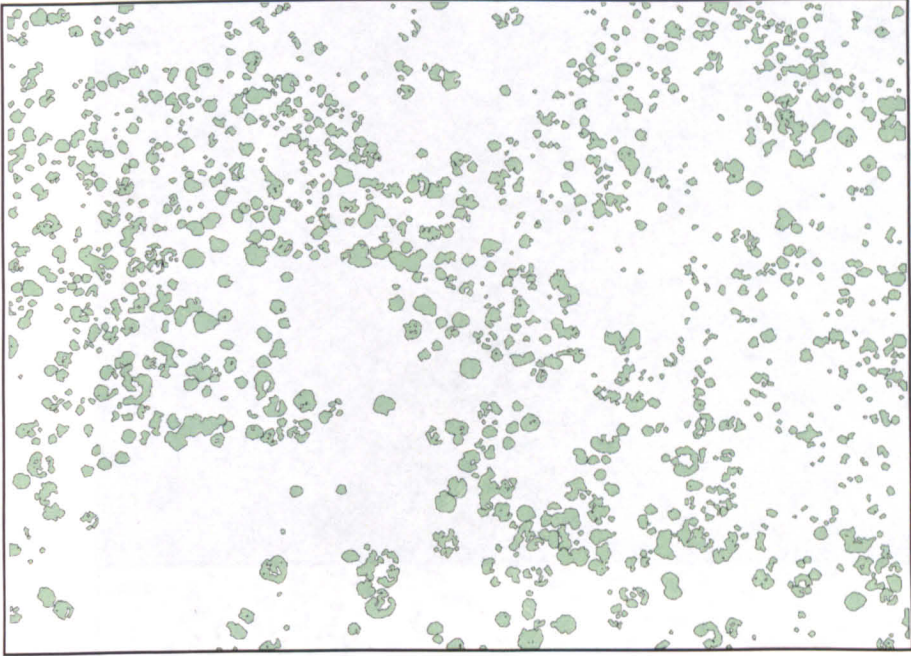
Figure 4-14 Classified imagery of *S. supranubius* on the five focal substrates. Classification performed in Matlab using the Incremental SVDD classifier, training on target data only, using the polynomial kernel with a parameter of 2 and a rejection error of 0.001



(a) Substrate 1a (dimensions: 550 x 350 m)



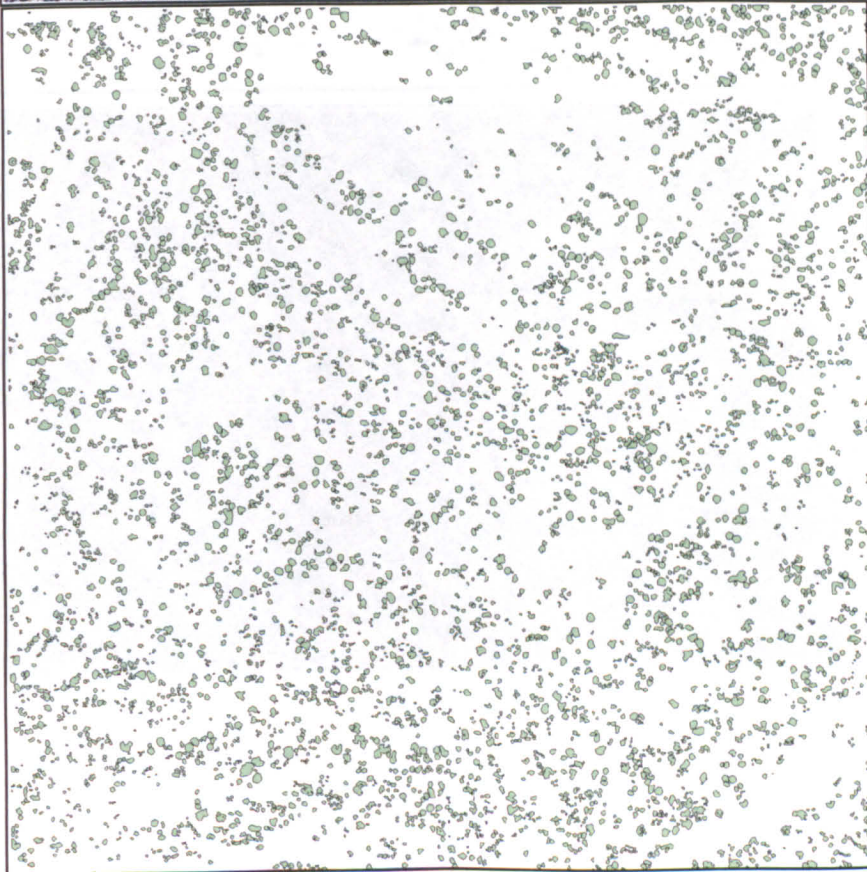
← z



0 20 40 80 metres

(b) Substrate 1b (dimensions 250 m x 350 m)



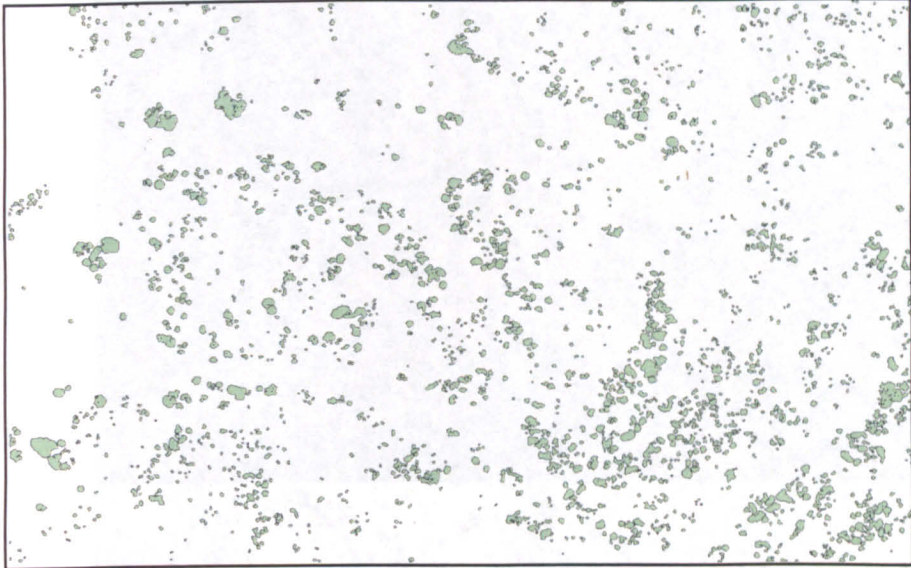


0 50 100 200 metres

(c) Substrate 2 (dimensions 700 m x 700 m)



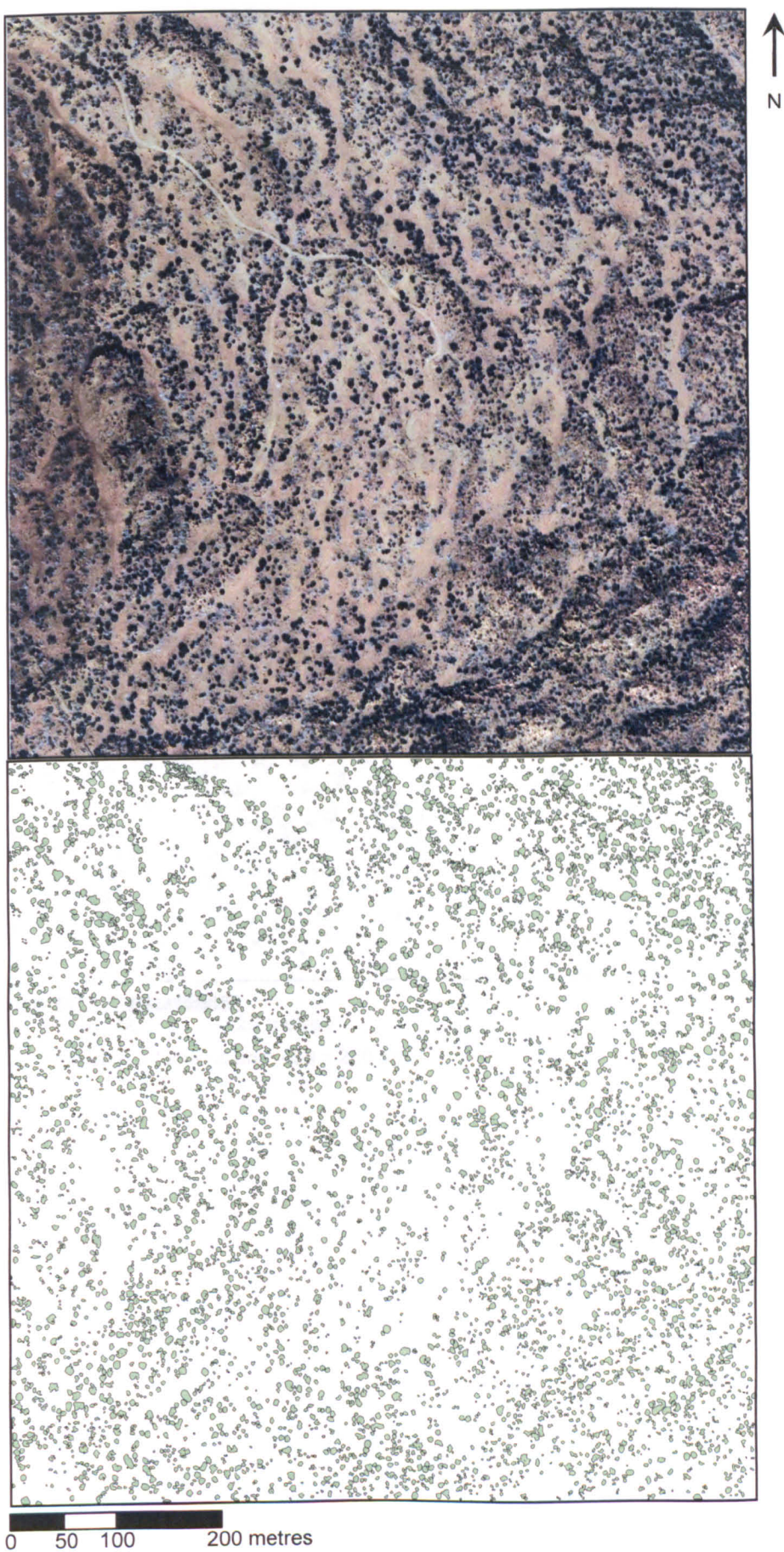
← z



0 30 60 120 metres

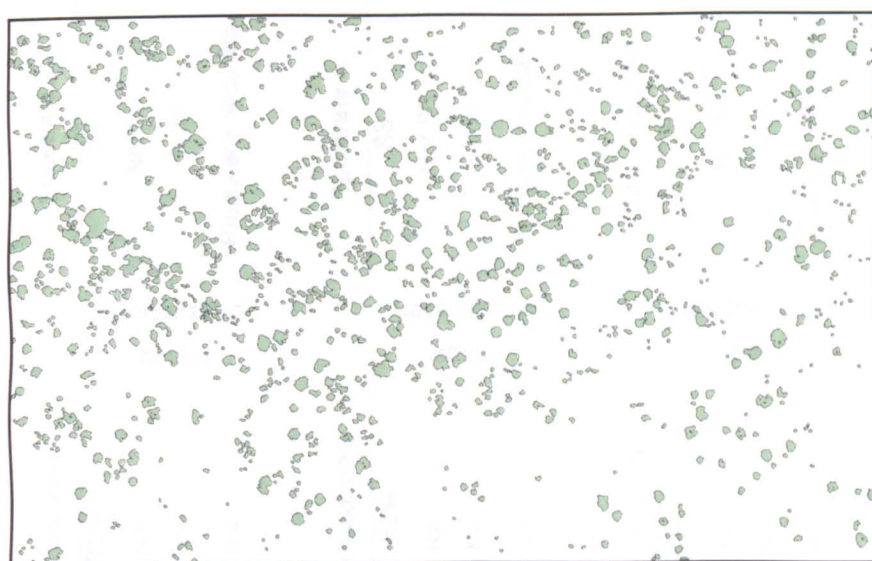
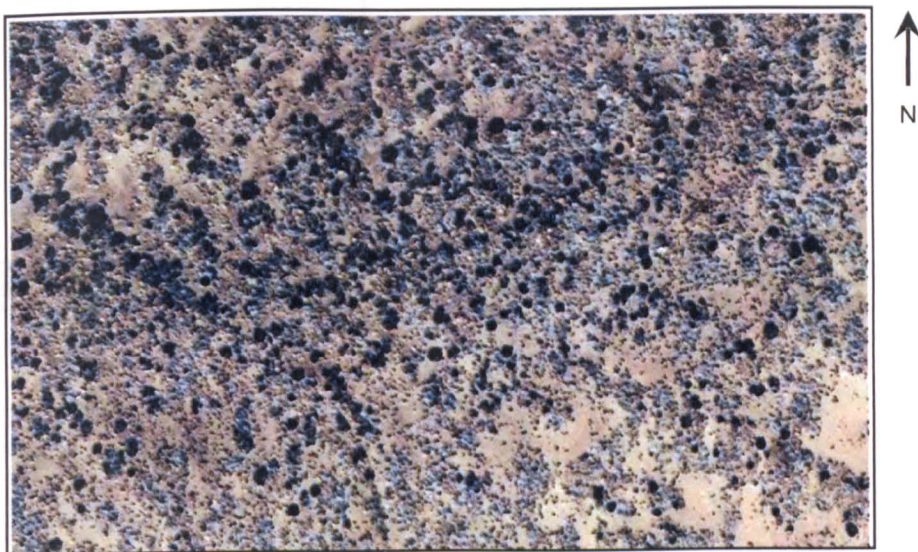
(d) Substrate 3 (dimensions: 400 x 650 m)





(e) Substrate 4 (dimensions: 700 x 700 m)





0 30 60 120 metres

(e) Substrate 5 (dimensions: 400 x 250 m)

Table 4-7 Error matrices for the classification of *S. supranubius* individuals.

Image classification	Substrate 1		Field		Substrate 2		Field		
	SS	Not SS	Total	User's accuracy	SS	Not SS	Total	User's accuracy	
	345	35	380	91%	306	48	354	86%	
	0	N/A	N/A	N/A	17	N/A	N/A	N/A	
	345	N/A	N/A	N/A	323	N/A	N/A	N/A	
Producer's accuracy		100%	N/A	N/A	Producer's accuracy		95%	N/A	N/A
Image classification	Substrate 3		Field		Substrate 4		Field		
	SS	Not SS	Total	User's accuracy	SS	Not SS	Total	User's accuracy	
	221	18	239	92%	388	37	425	91%	
	11	N/A	N/A	N/A	4	N/A	N/A	N/A	
	232	N/A	N/A	N/A	392	N/A	N/A	N/A	
Producer's accuracy		95%	N/A	N/A	Producer's accuracy		99%	N/A	N/A
Image classification	Substrate 5		Field		All substrates		Field		
	SS	Not SS	Total	User's accuracy	SS	Not SS	Total	User's accuracy	
	147	27	174	84%	1407	165	1572	90%	
	3	N/A	N/A	N/A	35	N/A	N/A	N/A	
	150	N/A	N/A	N/A	1442	N/A	N/A	N/A	
Producer's accuracy		98%	N/A	N/A	Producer's accuracy		98%	N/A	N/A

On all substrates except Substrate 5 the target accuracy of 85% was exceeded. The accuracy on Substrate 5 was only marginally below the target. For analysis in subsequent chapters the thematic maps were converted into point patterns using ArcMap 9.2, using the centroid of the canopy to represent the point location of the shrub. This produced a mapped point pattern, the properties of which describe the horizontal spatial structure of the *S. supranubius* population. Point approximation is intended to represent the rooting location of the shrub. Clearly, however, the rooting location will not always be at the centre of the canopy. Thus the process of reducing the thematic map to points will have incorporated some error into the final mapped point pattern.

## **4.5 CONCLUSIONS**

Following extensive cross-validation of different classifier structures, an area of 162 ha was classified using an incremental SVDD model, trained on 521 target cases, polynomial kernel with parameter 2 and a rejection error of 0.001. By testing so many classifier structures it was possible to assess the general effect of different factors on classification accuracy. Noteworthy results include that the polynomial kernel produced consistently greater classification accuracies than either the Gaussian RBF or exponential kernel. Secondly, when training the classifier, incorporating training data on the outlier classes caused an unexpected reduction in classification accuracy.

## CHAPTER 5: THE EFFECT OF EXTENT ON PATTERN ANALYSES USING $g(r)$ AND $L(r)$

---

*We can no longer...cling to the belief that the scale on which we view systems does not affect what we see... (Wiens, 1999 p. 371)*

### 5.1 INTRODUCTION

The concept of scale is fundamental in ecological studies (Levin, 1992; Wiens, 1999; Schneider, 2001). Three attributes of scale are particularly important: grain, focus and extent. Grain describes the area represented by each observational unit, focus describes the area of the analytical unit, and extent refers to the total geographic area being investigated. Variations in all three can affect the results of an analysis, its comparability with other studies and our perception of the processes controlling ecological phenomena. This chapter considers the effect of the last measure, extent, on the accuracy of spatial point pattern analyses. The following terminology applies in the chapter. The area of each mapped plot defines the *extent* of the analysis, and *scale* refers to the intervals and distances of spatial autocorrelation reported by the  $g(r)$ -function and the  $L(r)$ -function (a commonly used transformation of Ripley's  $K(r)$ ). These functions are hereafter referred to as  $g(r)$  and  $L(r)$ .

It has long been recognised that the choice of extent can dramatically alter results and interpretations in ecological systems (Gehlke and Biehl, 1934). Studies of broad-scale species richness have shown that the scale of observation can greatly affect estimates of alpha (Tylianakis et al., 2006) and beta (Kallimanis et al., 2008) species richness, and the predicted relationships between species richness and external influences such as disturbance (Hill and Hamer et al., 2004) and environmental factors (Rahbek and Graves, 2001; Foody, 2004; Hurlbert and White, 2005). Thus, through their selection of extent, researchers may instigate interpretative bias. Nonetheless, there is no pre-defined 'correct' extent at which to study a system or its components.



Point pattern data are increasingly being collected in plant communities and analysed using spatial point pattern statistics (Law et al., 2009; Section 1.3.4). Both  $g(r)$  and  $L(r)$  are commonly used to draw inferences about the biotic and abiotic processes driving plant population spatial structure. However, for a detailed analysis and meaningful interpretations, the robustness of the functions must be understood (Freeman and Ford, 2002). Despite the increasing frequency with which these techniques are being applied, only a few published articles have considered their robustness. Freeman and Ford (2002) found that both missing data and measurement errors could affect the magnitude and scale of the spatial patterns identified by  $L(r)$ . Zenner and Peck (2009) provide the only known attempt, to my knowledge, to assess the effect of changing extent on the performance of spatial point pattern analyses. They concluded that an extent of 0.5 ha was sufficient to characterise the spatial structure of managed forests using Ripley's  $K(r)$ . However, Zenner and Peck did not directly compare the performance of Ripley's  $K(r)$  in windows of different extent. Instead the authors compared the patterns indicated by Ripley's  $K(r)$  at 0.5 ha with other metrics (such as dbh distribution and tree density) measured in windows of 0.05 – 1 ha in extent. Their conclusions were based on the assumption that spatial structures estimated at 1 ha represented the 'true' population structure. Despite the frequency with which such small extents are used (Figure 1-3), they represent only a minute fraction of the range of most species. No studies have considered the performance of  $g(r)$  and  $L(r)$  at larger extents and how they compare to the spatial patterns detected at more commonly published extents. This comparison is timely as, with the growing availability of aerial photographs and image classification software, the potential to digitally map extensive areas of plant populations is increasing (e.g., Moustakas et al., 2008; Chapter 4).

Details of the analytical procedures used in 109 studies employing univariate spatial point pattern analyses are summarised in Appendix A. About 77% of the articles ( $n = 84$ ) use plot extents of less than 5 ha, and just over half of all articles ( $n = 56$ ) use a plot extent of 1 ha or less (see Figure 1-3). Several factors are likely to drive the

selection of small plot areas. Firstly, most studies aim to conduct analyses in homogeneous areas to eliminate the confounding effect of environmental heterogeneity (Wagner and Fortin, 2005; Getzin et al., 2008). These conditions will be more easily satisfied in smaller areas. Secondly, most studies collect data manually (Section 1.3.4). Mapping small extents reduces both the time and cost requirements of a field-based study. Furthermore, reducing the area mapped will usually reduce the number of units (e.g., shrubs) being mapped, potentially reducing the propagation of location ( $xy$ ) measurement errors that have been shown to influence the detection of spatial patterns (Freeman and Ford, 2002).

### **5.1.1 AIMS AND OBJECTIVES**

This chapter investigates the effect of varying plot extent on the identification of spatial patterns using  $g(r)$  and  $L(r)$ . Three main hypotheses are outlined:

*Hypothesis 1:* Quantitative pattern detection: Changes in plot extent will affect the spatial pattern described by  $g(r)$  and  $L(r)$ .

*Hypothesis 2:* Spatial consistency of quantitative pattern detection: The geographical location of the plot will affect the spatial pattern described by  $g(r)$  and  $L(r)$ .

*Hypothesis 3:* Qualitative pattern detection: Changes in plot extent will affect the interpretation of 'significant' patterning based upon Monte Carlo simulations of the CSR null model.

## **5.2 METHODS**

### **5.2.1 STUDY AREA AND DATA COLLECTION**

Two sites were selected: the Majua and the Montaña lava flows (see Table 2-1). These substrates were selected as they had the largest mapped area of *S. supranubius* individuals (49 ha each). To maintain consistency with other chapters these are referred to as Substrates 2 and 4 respectively. Six experimental extents were selected; 0.0625 ha (25 x 25 m), 0.25 ha (50 x 50 m), 1 ha (100 x 100 m),

2.25 ha (150 x 150 m), 4 ha (200 x 200 m) and 6.25 ha (250 x 250 m). Ten replicate plots of each experimental extent were randomly located within both of the 49 ha plots. Table 5-1 provides summary statistics for the ten replicate plots at each extent on both substrates (more detail is provided in Appendix D). According to published sample size recommendations (see Section 1.3.4), extents of 1 ha and greater (in this system) have enough individuals to provide robust descriptions of spatial pattern. Pseudo-replication within extents was avoided by rejecting plots that overlapped substantially with another plot. Plots were randomly re-located if they overlapped with another plot by 25% or more. There were no pair-wise overlaps of greater than 25% at plot extents of 2.25 ha and lower. Some overlaps of greater than 25% had to be accepted at the two largest extents. On Substrate 2 there were four quadrat pairs with an overlap of greater than 25% at the 4 ha extent, and 10 quadrat pairs with an overlap of greater than 25% at the 6.25 extent. The corresponding numbers of pair-wise overlaps on Substrate 4 were three and five respectively.

Table 5-1 Average number of shrubs per replicate plot at each extent. Standard error is given in parentheses

Substrate	Extent					
	0.0625 ha	0.25 ha	1 ha	2.25 ha	4 ha	6.25 ha
2	6.9	25.6	97.9	218.9	377.2	619.4
	(1.3)	(4.2)	(11.2)	(22.4)	(27.4)	(31.8)
4	5.8	31.2	128.5	300.7	500.8	802.4
	(0.8)	(2.0)	(8.0)	(16.7)	(22.9)	(24.4)

5.2.2 SPATIAL POINT PATTERN ANALYSIS

The pattern of all *S. supranubius* individuals within each plot was assessed using  $g(r)$  and  $L(r)$ . At each experimental extent the empirical  $g(r)$  and  $L(r)$  calculated in the ten replicate plots were combined to produce a single  $g(r)$  and  $L(r)$ . A ring width of 3 m was used when calculating  $g(r)$ . As the replicate plots within each extent class are congruent (i.e., the same size and shape), and the processes generating the patterns are hypothesised to be spatially consistent within each substrate, the combined

function was calculated as the unweighted average of the individual function values at each scale (Diggle, 2003: Eq. 4.20 on Page 52; Illian et al., 2008; Law et al., 2009). Alternative methods for the analysis of replicated spatial point patterns can be found in Mateu (2001) and Bell and Grunwald (2004). The pattern for each full 49 ha plot was also calculated. The results for the 49 ha plots were assumed to provide the most accurate, i.e., most 'correct', assessment of *S. supranubius* spatial pattern. This is hereafter referred to as the 'reference' pattern.

All analyses were performed in Programita (Wiegand and Moloney, 2004). The commonly used estimator of  $K(r)$  (from which  $L(r)$  and  $g(r)$  are derived) proposed by Ripley (1976, 1981) incorporates a weighting factor to account for edge effects (Getis and Franklin, 1987; Haase, 1995). Although the functions applied include weighting factors to correct for edge effects, additional measures were taken to ensure the results reported are accurate. To minimise edge effects indices should not be calculated up to the maximum scale of the sample window, unless the sample window is very large. Previous studies have recommended limiting the interpretation of  $g(r)$  and  $L(r)$  to distances equal to a quarter of the length of the plot (Baddeley and Turner, 2005). Others have been less conservative, interpreting functions up to scales equalling two thirds the side length of the plot (Fortin, 1999). In the following analyses  $g(r)$  and  $L(r)$  were interpreted to a maximum scale that equalled half the side length of the plot (i.e., spatial patterns in plots of 50 x 50 m were only analysed up to scales of  $r = 25$  m).

The main aim of this chapter is to investigate whether the extent of analyses influences the detection of spatial pattern. Spatial patterns are frequently used to infer the operation of ecological processes. Thus, the analyses in this chapter focus on the performance of  $g(r)$  and  $L(r)$  at scales that are ecologically meaningful to *S. supranubius* individuals. The maximum scale of interaction between *S. supranubius* individuals is estimated at 22 m (see Section 7.3.2 for more detail). To allow some leeway,  $g(r)$  and  $L(r)$  are calculated to a maximum scale of 30 m (in plots  $\geq 1$  ha). Additionally,  $g(r)$  and  $L(r)$  are not calculated for scales  $< 2$  m. At these scales

the influence of canopy extent will affect the reliable calculation of the statistics and may result in the incorrect detection of regularity (i.e. soft-core effects, Wiegand et al., 2006, see Section 6.1).

### **Preliminary analysis: environmental heterogeneity**

Preliminary analyses were conducted to assess the presence of large-scale environmental heterogeneity on both substrates as this may influence the pattern of individuals. The distribution of the largest adults ( $\geq 30 \text{ m}^2$ ) across the entire plot (49 ha) was compared to the null model of complete spatial randomness (CSR) using  $g(r)$  and  $L(r)$ . Based on the assumption that large adults persist only in environmentally benign areas, deviation of their distribution from CSR at large scales ( $r > 20 \text{ m}$ ) indicates the presence of environmental heterogeneity (Getzin et al., 2008). Further justification of this technique is provided in Chapter 7 (see Section 7.3.2). The spatial structure of large ( $\geq 30 \text{ m}^2$ ) individuals was investigated to a maximum scale of 50 m. A ring with of 4 m was used in the calculation of  $g(r)$  as this gave a relatively smooth function. Spatial pattern was evaluated by comparing the empirical  $g(r)$  and  $L(r)$  to the 5<sup>th</sup>-lowest and 5<sup>th</sup>-highest value of 999 Monte Carlo simulations of the CSR null model, generating approximately 99% simulation envelopes.

### **Analysis 1: the effect of extent on the estimation of pattern trend**

The unweighted average function calculated at each extent was compared to the reference function. Both the ability of analyses at different extents to identify the major pattern trend (i.e., aggregation or dispersion) and the scale-dependent accuracy of function estimates were considered. Greater differences between the average function and the reference function indicate less accurate pattern detection. The height of  $g(r)$  is taken as a measure of pattern strength, with larger values indicating the detection of stronger (aggregative) patterns (following Barbeito et al., 2009; Getzin et al., 2008).

## **Analysis 2: the spatial consistency of pattern detection**

This analysis considers the effect of extent on the accuracy and spatial consistency of pattern detection. Two aspects of pattern detection are considered: (un)reliability and pattern strength.

### ***Analysis 2a: the effect of extent on the (un)reliability of pattern detection***

For each extent, the standard deviation in the  $g(r)$  and  $L(r)$  values was calculated at each scale across the ten replicates. The use of the standard deviation allows the analysis to investigate the range within which the empirical  $g(r)$  and  $L(r)$  functions may be expected to lie if individual plots were performed at each of the experimental extents. A large standard deviation indicates unreliable pattern detection for that scale.

### ***Analysis 2b: the effect of extent on the magnitude and scale of the dominant pattern***

This analysis considered whether extent affects the magnitude of the dominant pattern detected (i.e., the maximum value  $g(r)$  and  $L(r)$  attains) and the scale at which the dominant pattern occurs. For the ten plots at each extent the maximum values of  $g(r)$  and  $L(r)$ , and scale at which the maximum value occurred were recorded. The average magnitude and scale of the dominant pattern was calculated for each extent (Figure 5-1) and compared to the magnitude and scale of the dominant pattern detected at 49 ha. For each calculation of the average magnitude and scale the coefficient of variation is reported. This is calculated as the ratio of the standard deviation to the mean, and is a useful measure for comparing the variability of datasets when their means differ greatly (see Table 5-2 and Table 5-3).

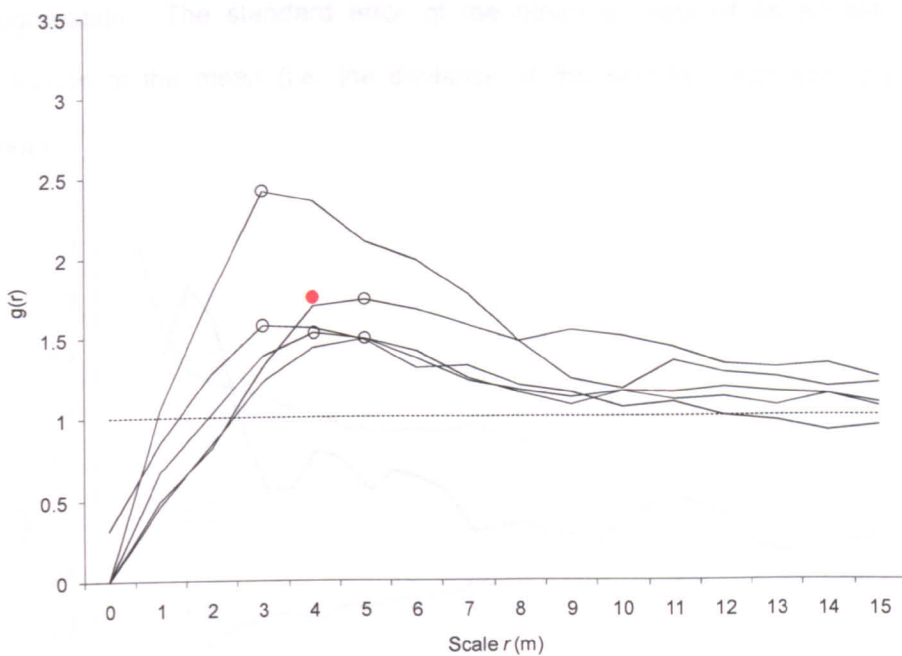


Figure 5-1 Calculation of the average magnitude and scale of the dominant pattern from several replicate plots. The maximum values of each replicate (open circles) are averaged to find the average magnitude and scale of the pattern (solid red circle).

### Analysis 3: the effect of extent on pattern detection by Monte Carlo simulation envelopes

The majority of studies determine the presence of spatial pattern by comparing the empirical  $L(r)$  or  $g(r)$  to envelopes created from Monte Carlo simulations of a specified null model (Section 1.3.4). 'Significant' pattern is reported to occur at the scales where the empirical function falls outside the simulation envelope. This analysis investigated whether plot extent influences spatial pattern interpretation based upon Monte Carlo simulations. The results focused on the detection of significant aggregation, as analyses reveal the pattern of *S. supranubius* to be predominantly aggregative (see Section 5.3.2). For each replicate plot approximately 99% simulation envelopes were created from the 5<sup>th</sup>-highest and 5<sup>th</sup>-lowest value of 999 Monte Carlo simulations of the CSR null model. Each Monte Carlo simulation used the same number of individuals as in the observed pattern. For each plot the number of 1 m intervals at which the empirical  $g(r)$  and  $L(r)$  exceeded the simulation envelope was recorded (Figure 5-2). For each extent the results for the ten plots were combined to give the average number of scales at which Monte Carlo techniques detected

aggregation. The standard error of the mean is reported as an estimate of the precision of the mean (i.e. the deviation of the sample mean from the population mean).

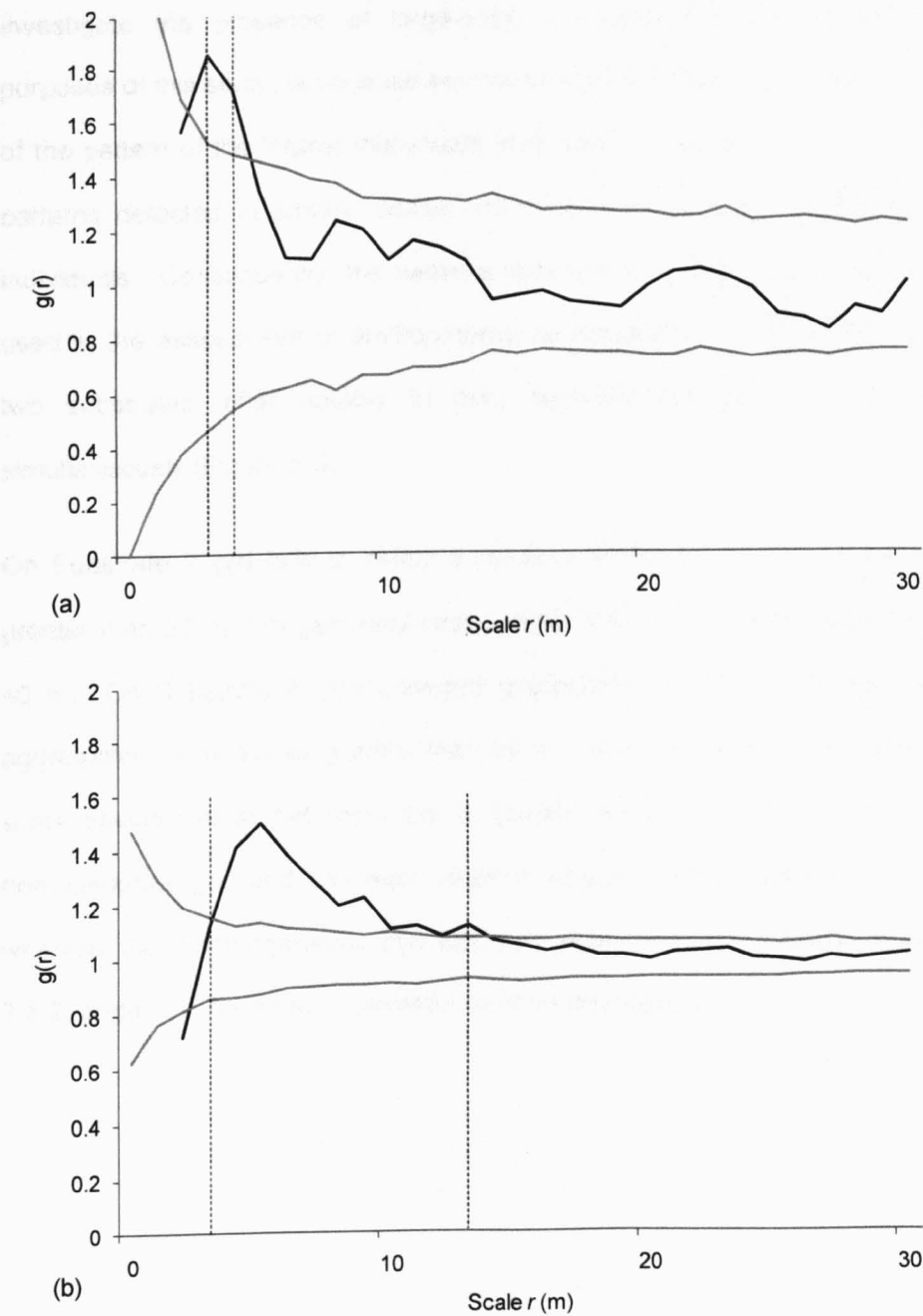


Figure 5-2 Calculation of the number of scales at which  $g(r)$  exceeds the 99% CSR simulation envelopes. On graph (a)  $g(r)$  only exceeds the simulation envelope at 4 and 5 m, thus significant aggregation is detected at two scales. On graph (b)  $g(r)$  exceeds the simulation envelope between 4 and 14 m (inclusive), thus significant aggregation is detected at 11 scales.



## 5.3 RESULTS

### 5.3.1 ENVIRONMENTAL HETEROGENEITY

The pattern of mature individuals ( $\geq 30 \text{ m}^2$ ) was compared to the CSR null model to investigate the presence of large-scale environmental heterogeneity. For the purposes of this study, large-scale environmental heterogeneity is defined by deviation of the pattern of the largest individuals from CSR at scales  $r > 20 \text{ m}$ . It is likely that patterns detected at smaller scales will result from biological interactions between individuals. Consequently, the patterns detected at scales of less than 20 m are not used in the assessment of environmental heterogeneity. To determine whether the two substrates differ notably in their heterogeneity  $g(r)$  and  $L(r)$  are plotted simultaneously (Figure 5-3).

On Substrate 2  $g(r)$  fails to detect a random distribution of individuals at all scales greater than 20 m.  $L(r)$  generally confirms this trend, but detects aggregation beyond 40 m. On Substrate 4,  $g(r)$  detected aggregation at 24 m, whereas  $L(r)$  detected aggregation at all scales greater than 20 m. Both  $g(r)$  and  $L(r)$  suggest that large-scale environmental heterogeneity is greater on Substrate 4. Consequently, the homogeneous  $g(r)$  and  $L(r)$  were used to analyse spatial patterns on Substrate 2, whereas the inhomogeneous  $g(r)$  and  $L(r)$  were used on Substrate 4 (see Section 7.3.2, page 187, for more information on the inhomogeneous  $g(r)$ ).

### 5.3 THE EFFECT OF EXTENT ON THE ESTIMATION OF PATTERNS (195-40)

The effect of extent on pattern detection by  $g(r)$

Substrate 2 (Figure 5-4) - Analyses at 49 ha identified strong aggregation at scales 2 to 7 m, a peak in nearest-neighbour distance at 4 m. At scales between 10 and 20 m,  $g(r)$  approaches a value of 1 indicating a tendency towards a random pattern.

Substrate 4 (Figure 5-4) - Analyses at 49 ha identified a peak at 2 m, a peak in nearest-neighbour distance at 4 m. At scales between 10 and 20 m,  $g(r)$  approaches a value of 1 indicating a tendency towards a random pattern.

Analyses at 1 ha and 25 ha identified a peak at 2 m, a peak in nearest-neighbour distance at 4 m. At scales between 10 and 20 m,  $g(r)$  approaches a value of 1 indicating a tendency towards a random pattern.

Analyses at 1 ha and 25 ha identified a peak at 2 m, a peak in nearest-neighbour distance at 4 m. At scales between 10 and 20 m,  $g(r)$  approaches a value of 1 indicating a tendency towards a random pattern.

Analyses at 1 ha and 25 ha identified a peak at 2 m, a peak in nearest-neighbour distance at 4 m. At scales between 10 and 20 m,  $g(r)$  approaches a value of 1 indicating a tendency towards a random pattern.

Analyses at 1 ha and 25 ha identified a peak at 2 m, a peak in nearest-neighbour distance at 4 m. At scales between 10 and 20 m,  $g(r)$  approaches a value of 1 indicating a tendency towards a random pattern.

Analyses at 1 ha and 25 ha identified a peak at 2 m, a peak in nearest-neighbour distance at 4 m. At scales between 10 and 20 m,  $g(r)$  approaches a value of 1 indicating a tendency towards a random pattern.

Analyses at 1 ha and 25 ha identified a peak at 2 m, a peak in nearest-neighbour distance at 4 m. At scales between 10 and 20 m,  $g(r)$  approaches a value of 1 indicating a tendency towards a random pattern.

Analyses at 1 ha and 25 ha identified a peak at 2 m, a peak in nearest-neighbour distance at 4 m. At scales between 10 and 20 m,  $g(r)$  approaches a value of 1 indicating a tendency towards a random pattern.

Analyses at 1 ha and 25 ha identified a peak at 2 m, a peak in nearest-neighbour distance at 4 m. At scales between 10 and 20 m,  $g(r)$  approaches a value of 1 indicating a tendency towards a random pattern.

Analyses at 1 ha and 25 ha identified a peak at 2 m, a peak in nearest-neighbour distance at 4 m. At scales between 10 and 20 m,  $g(r)$  approaches a value of 1 indicating a tendency towards a random pattern.

Analyses at 1 ha and 25 ha identified a peak at 2 m, a peak in nearest-neighbour distance at 4 m. At scales between 10 and 20 m,  $g(r)$  approaches a value of 1 indicating a tendency towards a random pattern.

Analyses at 1 ha and 25 ha identified a peak at 2 m, a peak in nearest-neighbour distance at 4 m. At scales between 10 and 20 m,  $g(r)$  approaches a value of 1 indicating a tendency towards a random pattern.

Analyses at 1 ha and 25 ha identified a peak at 2 m, a peak in nearest-neighbour distance at 4 m. At scales between 10 and 20 m,  $g(r)$  approaches a value of 1 indicating a tendency towards a random pattern.

Analyses at 1 ha and 25 ha identified a peak at 2 m, a peak in nearest-neighbour distance at 4 m. At scales between 10 and 20 m,  $g(r)$  approaches a value of 1 indicating a tendency towards a random pattern.

Analyses at 1 ha and 25 ha identified a peak at 2 m, a peak in nearest-neighbour distance at 4 m. At scales between 10 and 20 m,  $g(r)$  approaches a value of 1 indicating a tendency towards a random pattern.

Analyses at 1 ha and 25 ha identified a peak at 2 m, a peak in nearest-neighbour distance at 4 m. At scales between 10 and 20 m,  $g(r)$  approaches a value of 1 indicating a tendency towards a random pattern.

Analyses at 1 ha and 25 ha identified a peak at 2 m, a peak in nearest-neighbour distance at 4 m. At scales between 10 and 20 m,  $g(r)$  approaches a value of 1 indicating a tendency towards a random pattern.

Analyses at 1 ha and 25 ha identified a peak at 2 m, a peak in nearest-neighbour distance at 4 m. At scales between 10 and 20 m,  $g(r)$  approaches a value of 1 indicating a tendency towards a random pattern.

Analyses at 1 ha and 25 ha identified a peak at 2 m, a peak in nearest-neighbour distance at 4 m. At scales between 10 and 20 m,  $g(r)$  approaches a value of 1 indicating a tendency towards a random pattern.

Analyses at 1 ha and 25 ha identified a peak at 2 m, a peak in nearest-neighbour distance at 4 m. At scales between 10 and 20 m,  $g(r)$  approaches a value of 1 indicating a tendency towards a random pattern.

Analyses at 1 ha and 25 ha identified a peak at 2 m, a peak in nearest-neighbour distance at 4 m. At scales between 10 and 20 m,  $g(r)$  approaches a value of 1 indicating a tendency towards a random pattern.

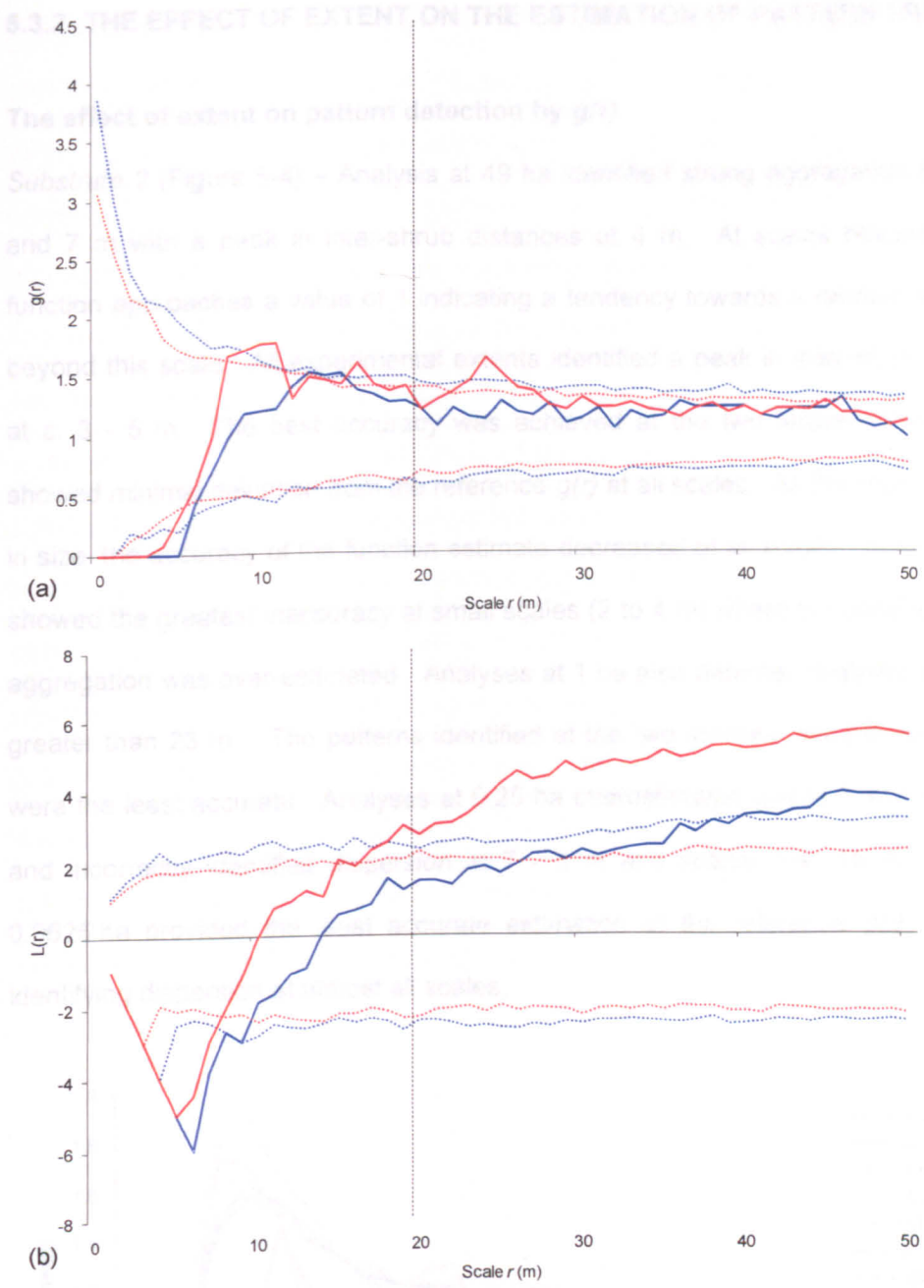


Figure 5-3 Comparing the heterogeneity on Substrate 2 (blue lines) and Substrate 4 (red lines) as assessed by (a) the homogeneous  $g(r)$  and, (b) the homogeneous  $L(r)$ . On both graphs the solid lines plot the function for all mature individuals ( $\geq 30 \text{ m}^2$ ) against their respective 99% simulation envelopes (dotted lines). Deviation from CSR at scales greater than 20 m (represented by the vertical lines) is taken as evidence for large-scale heterogeneity.

### 5.3.2 THE EFFECT OF EXTENT ON THE ESTIMATION OF PATTERN TREND

#### The effect of extent on pattern detection by $g(r)$

Substrate 2 (Figure 5-4) – Analysis at 49 ha identified strong aggregation between 3 and 7 m with a peak in inter-shrub distances at 4 m. At scales beyond 7 m the function approaches a value of 1 indicating a tendency towards a random distribution beyond this scale. All experimental extents identified a peak in inter-shrub distances at c. 3 - 5 m. The best accuracy was achieved at the two largest extents, which showed minimal deviation from the reference  $g(r)$  at all scales. As the plots decreased in size, the accuracy of the function estimate decreased at all scales. Extents of 1 ha showed the greatest inaccuracy at small scales (2 to 4 m) where the strength of shrub aggregation was over-estimated. Analyses at 1 ha also detected dispersion at scales greater than 23 m. The patterns identified at the two smallest experimental extents were the least accurate. Analyses at 0.25 ha overestimated  $g(r)$  at 3 and 10 – 11 m, and incorrectly identified dispersion at 7 - 8 m and scales over 15 m. Plots of 0.0625 ha provided the least accurate estimation of the reference  $g(r)$ , incorrectly identifying dispersion at almost all scales.

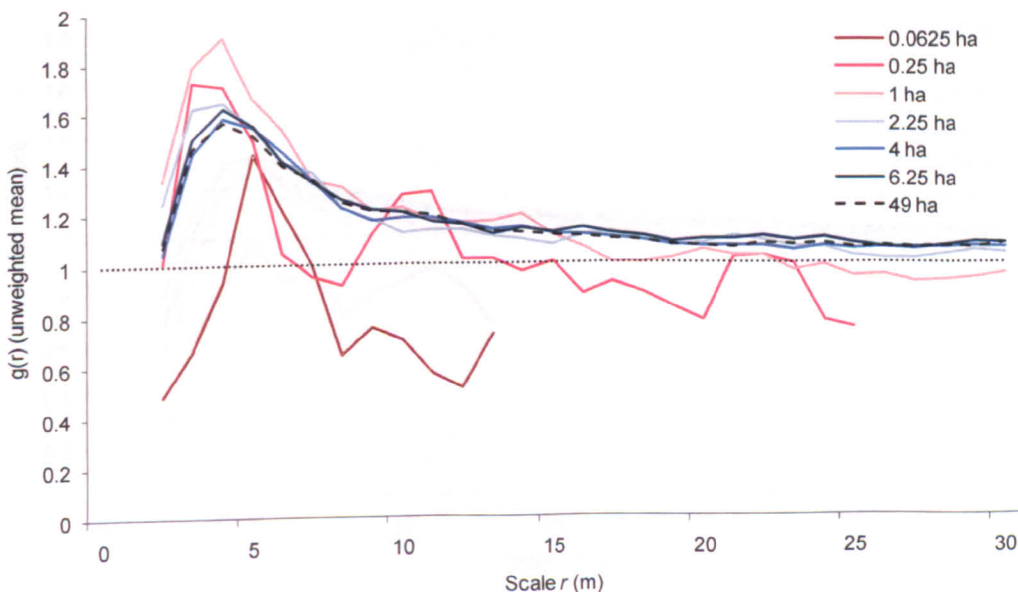


Figure 5-4 Unweighted mean homogeneous  $g(r)$  calculated from ten replicate plots at each of the six experimental extents plotted against the  $g(r)$  calculated at 49 ha on Substrate 2. The dotted line at  $g(r) = 1$  indicates the expected value under CSR.

*Substrate 4* (Figure 5-5) – Analysis at 49 ha identified notable aggregation between 4 and 6 m, peaking at 5 m. At larger scales  $g(r)$  decreases and approaches a value of 1 indicating an increasingly random distribution of individuals at scales of c. 12 m and greater. All experimental extents identified the same peak in shrub densities at 5 m, although at 0.0625 ha this peak was not strong enough to be considered aggregative. The accuracy of the estimates decreased with extent. The accuracies provided by the two largest extents were similar at all scales. Both showed the greatest inaccuracy at small scales (2 – 3 m) with accuracy improving as the scale increased. Plots of 2.25 ha also provided reasonably accurate estimates of  $g(r)$ . The functions produced at extents of 1 ha and larger identified the same patterning trend as at 49 ha, decreasing towards a value of  $g(r) = 1$  at a similar rate as the reference pattern. Analyses at 0.25 ha overestimated  $g(r)$  at small scales (2 – 3 m) and underestimated  $g(r)$  at most other scales. At scales larger than 20 m, dispersion was incorrectly identified. Analyses at 0.0625 ha under-estimated  $g(r)$  at all scales and incorrectly detected dispersion at all scales. In general, the functions produced at all extents were more accurate than at comparable extents on Substrate 2.

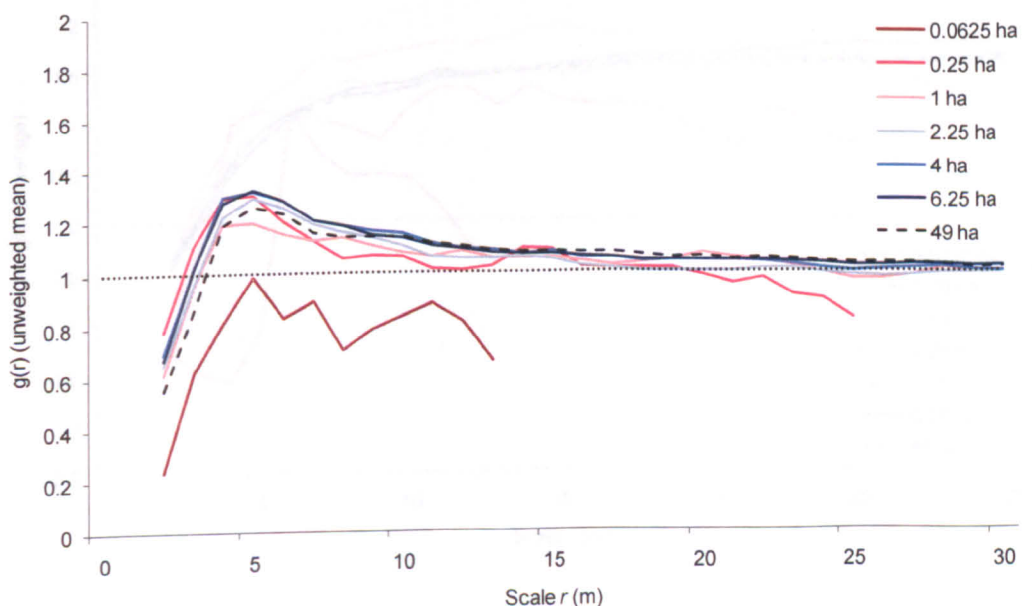


Figure 5-5 Unweighted mean inhomogeneous  $g(r)$  calculated from ten replicate plots at each of the six experimental extents plotted against the  $g(r)$  calculated at 49 ha on Substrate 4. The dotted line at  $g(r) = 1$  indicates the expected value under CSR.



### The effect of extent on the estimation of pattern trend by $L(r)$

*Substrate 2* (Figure 5-6) – Analyses at 49 ha show that the shrubs are dispersed at the smallest scales, but increase in density to become aggregated by 4 m. The function stabilises between 14 and 27 m. Analyses at all experimental extents described the initial dispersion of shrubs at small scales. The accuracy of the average function, however, decreased with decreases in plot extent. The three largest extents provided the most accurate estimations, detecting a similar spatial pattern trend to that observed at 49 ha. The average function calculated at 1 ha was also reasonably accurate although overestimated the value of  $L(r)$  at all scales  $r < 25$  m, detecting a peak in aggregation at 15 m. Analyses at the two smallest extents were the least accurate, failing to identify the correct spatial pattern trend. Analyses at 0.25 ha generally replicated the reference pattern up to 7 m, although over-estimated  $L(r)$  at 4 – 5 m, but underestimated  $L(r)$  at subsequent scales. At 0.0625 ha the estimated function bore little similarity to the correct spatial pattern, strongly under-estimating  $L(r)$  at almost all scales.

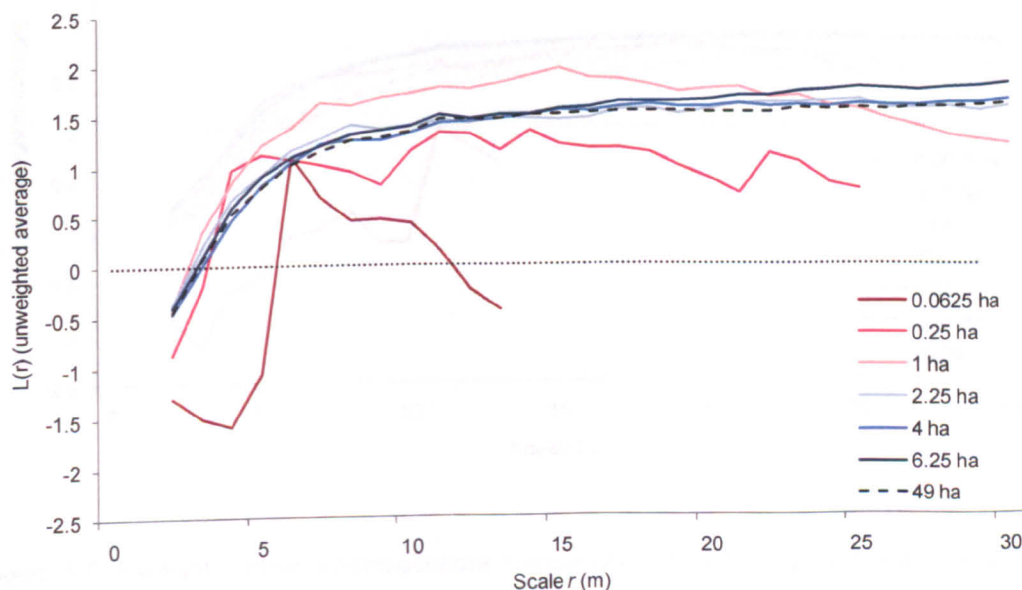


Figure 5-6 Unweighted mean homogeneous  $L(r)$  calculated from ten replicate plots at each of the six experimental extents plotted against the  $L(r)$  calculated at 49 ha on Substrate 2. The dotted line at  $L(r) = 0$  indicates the expected value under CSR.

Substrate 4 (Figure 5-7) – When analysed at 49 ha,  $L(r)$  detected dispersion at scales less than c. 6 m. Beyond this scale the pattern of shrubs becomes aggregated. The function continues to increase until it reaches scales of c. 23 m where it stabilises. Analyses conducted at 0.25 ha and larger correctly identified shrub dispersion up to scales of 5 m. Analyses at 0.0625 ha, however, incorrectly detected dispersion at all scales. The two largest extents (4 and 6.25 ha) correctly replicated the aggregation of individuals at scales of c. 20 m and above. These extents provided the most accurate estimation of the reference function, although accuracy decreased slightly at smaller scales ( $r < 10$  m). Analyses at 1 and 2.25 ha underestimated the strength of aggregation at large scales ( $r > 20$  m). Plots of 0.25 ha in extent underestimate  $L(r)$  at all scales  $r > 5$  m. The average function calculated at the smallest extent strongly under-estimated the reference function value at all scales. All extents provided a better estimate of pattern trend on Substrate 4 compared to Substrate 2.

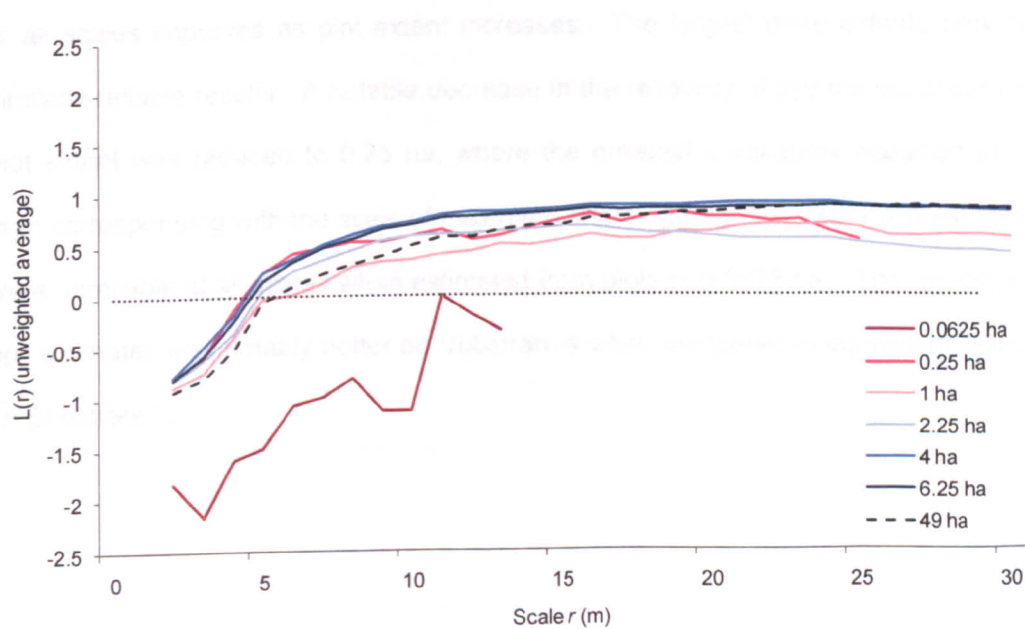


Figure 5-7 Unweighted mean inhomogeneous  $L(r)$  calculated from ten replicate plots at each of the six experimental extents plotted against the  $L(r)$  calculated at 49 ha on Substrate 4. The dotted line at  $L(r) = 0$  indicates the expected value under CSR.

### 5.3.3 THE EFFECT OF EXTENT ON THE SPATIAL CONSISTENCY OF PATTERN DETECTION

#### **The effect of extent on the (un)reliability of pattern detection by $g(r)$**

*Substrate 2* (Figure 5-8) – The reliability of  $g(r)$  estimates improves with increasing plot extent. The reliability at the two largest extents is very similar. Within each extent, except 0.0625 ha, the least reliable estimation of  $g(r)$  occurs at the scale of shrub aggregation. Unreliability increases notably when plot extent decreases to 1 ha. At this plot extent estimates of  $g(r)$  are quite unreliable up to scales of 15 m. Decreasing the plot extent to 0.25 ha reveals a peak in unreliability at 3 – 4 m and a notable increase in unreliability at 10 m which is not observed at the other extents. Analyses at 0.0625 ha are unreliable at all scales.

*Substrate 4* (Figure 5-9) – As observed on Substrate 2, the reliability of  $g(r)$  estimates at all scales improves as plot extent increases. The largest three extents provided similarly reliable results. A notable decrease in the reliability of  $g(r)$  did not occur until plot extent was reduced to 0.25 ha, where the greatest unreliability occurred at 3 – 4 m corresponding with the scale of maximum shrub aggregation. Function estimates were unreliable at all scales when estimated from plots of 0.0625 ha. The reliability of  $g(r)$  estimates was notably better on Substrate 4 when compared to equivalent extents on Substrate 2.

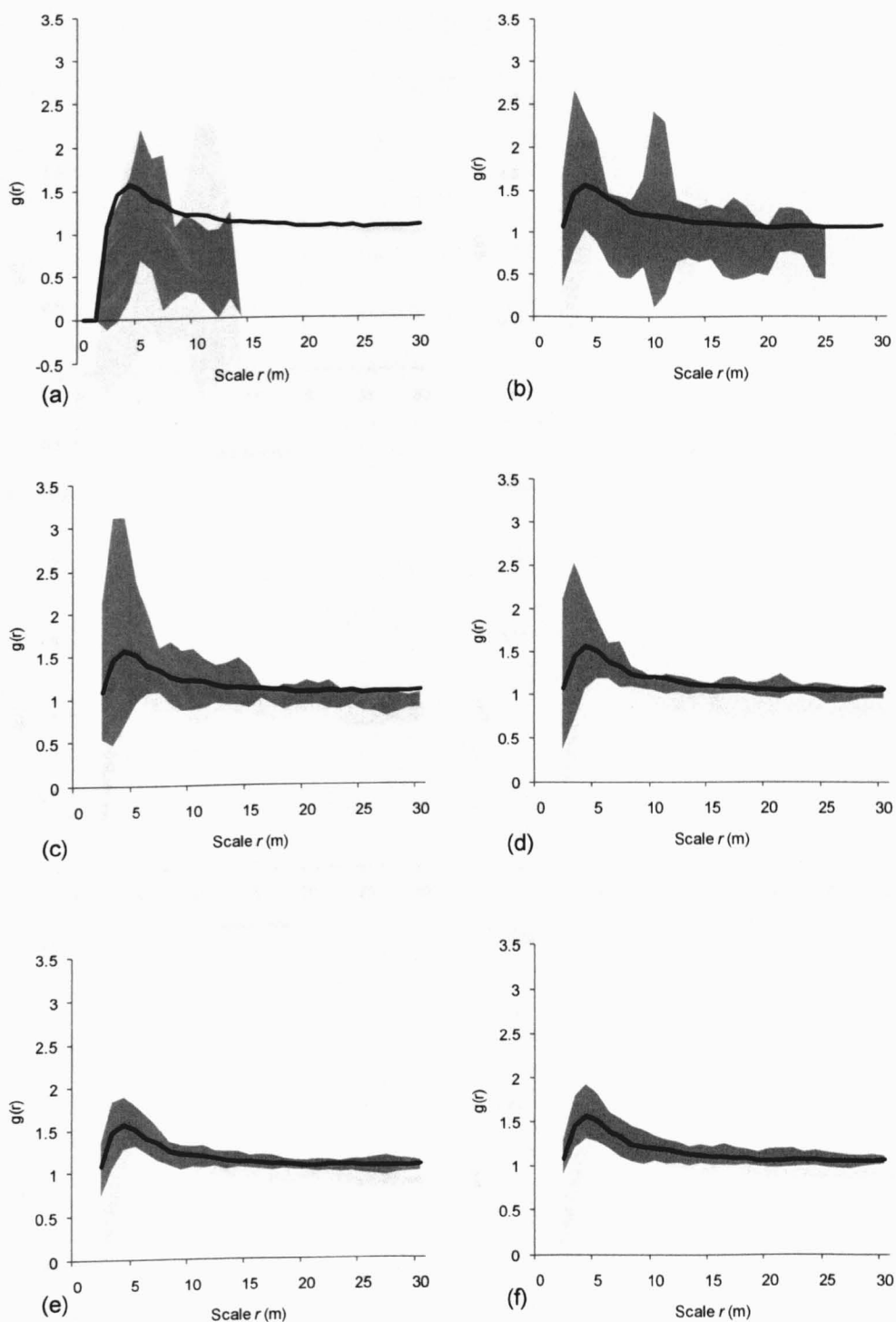
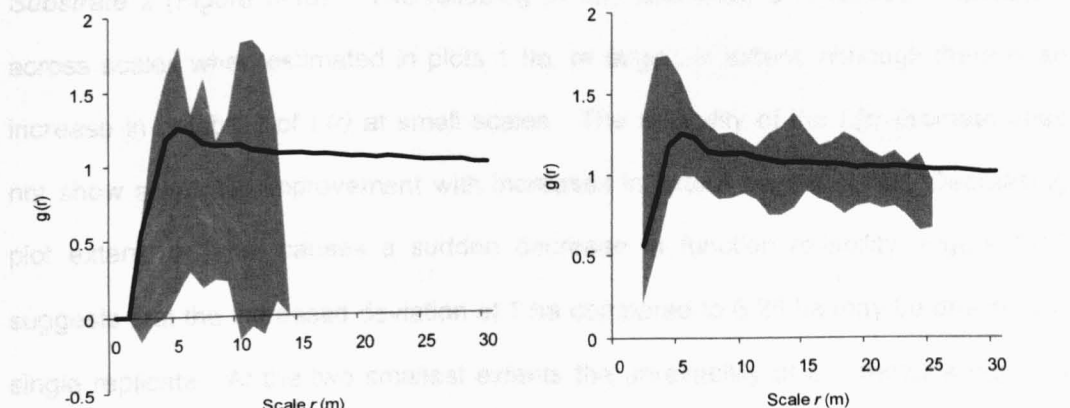


Figure 5-8 The unreliability of  $g(r)$  estimates calculated on Substrate 2 from ten replicate plots at six extents, (a) 0.0625 ha, (b) 0.25 ha, (c) 1 ha, (d) 2.25 ha, (e) 4 ha and (f) 6.25 ha. The grey area represents the standard deviation of  $g(r)$  estimates across ten replicates at each scale. The black line shows the reference  $g(r)$  calculated at 49 ha.



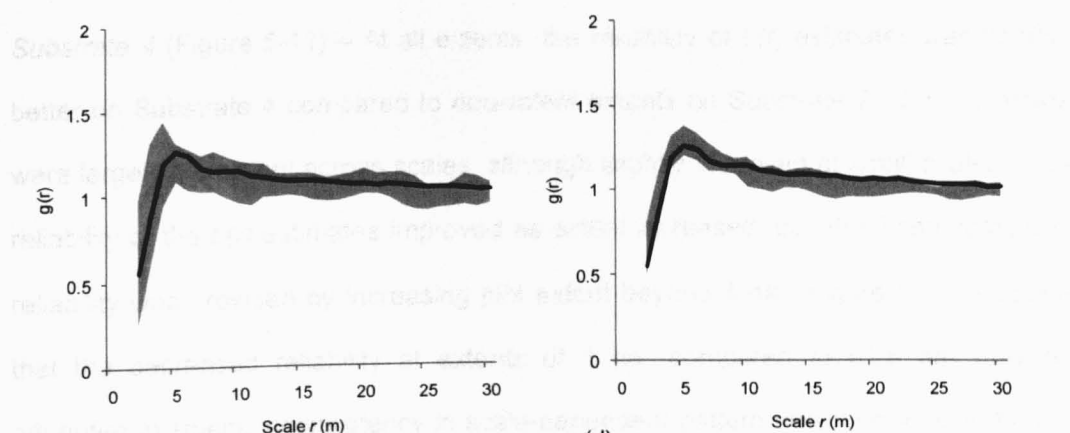
The effect of extent on the (un)reliability of  $g(r)$  estimates by 1.7

Substrate 2 (Figure 5-10) – The reliability of  $g(r)$  estimates calculated on Substrate 2



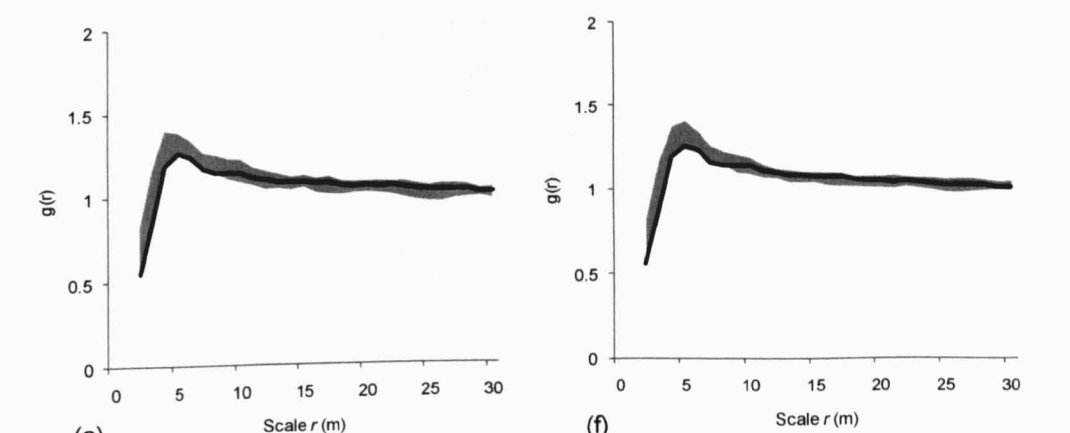
(a)

(b)



(c)

(d)



(e)

(f)

Figure 5-9 The unreliability of  $g(r)$  estimates calculated on Substrate 4 from ten replicate plots at six extents, (a) 0.0625 ha, (b) 0.25 ha, (c) 1 ha, (d) 2.25 ha, (e) 4 ha and (f) 6.25 ha. The grey area represents the standard deviation of  $g(r)$  estimates across ten replicates at each scale. The black line shows the reference  $g(r)$  calculated at 49 ha.

### **The effect of extent on the (un)reliability of pattern detection by $L(r)$**

*Substrate 2* (Figure 5-10) – The reliability of  $L(r)$  estimates is remarkably consistent across scales when estimated in plots 1 ha, or larger, in extent, although there is an increase in reliability of  $L(r)$  at small scales. The reliability of the  $L(r)$  estimate does not show a notable improvement with increases in extent beyond 4 ha. Decreasing plot extent to 1 ha, causes a sudden decrease in function reliability. Figure 5-12 suggests that the increased deviation at 1 ha compared to 6.25 ha may be driven by a single replicate. At the two smallest extents the unreliability of  $L(r)$  remains high but becomes more variable with scale.

*Substrate 4* (Figure 5-11) – At all extents, the reliability of  $L(r)$  estimates was notably better on Substrate 4 compared to equivalent extents on Substrate 2.  $L(r)$  estimates were largely consistent across scales, although slightly improved at small scales. The reliability of the  $L(r)$  estimates improved as extent increased, but little improvement in reliability was provided by increasing plot extent beyond 4 ha. Figure 5-12 suggests that the decreased reliability at extents of 1 ha (compared to 6.25 ha) may be attributed to spatial inconsistency in scale-dependent pattern detection across the ten replicates.

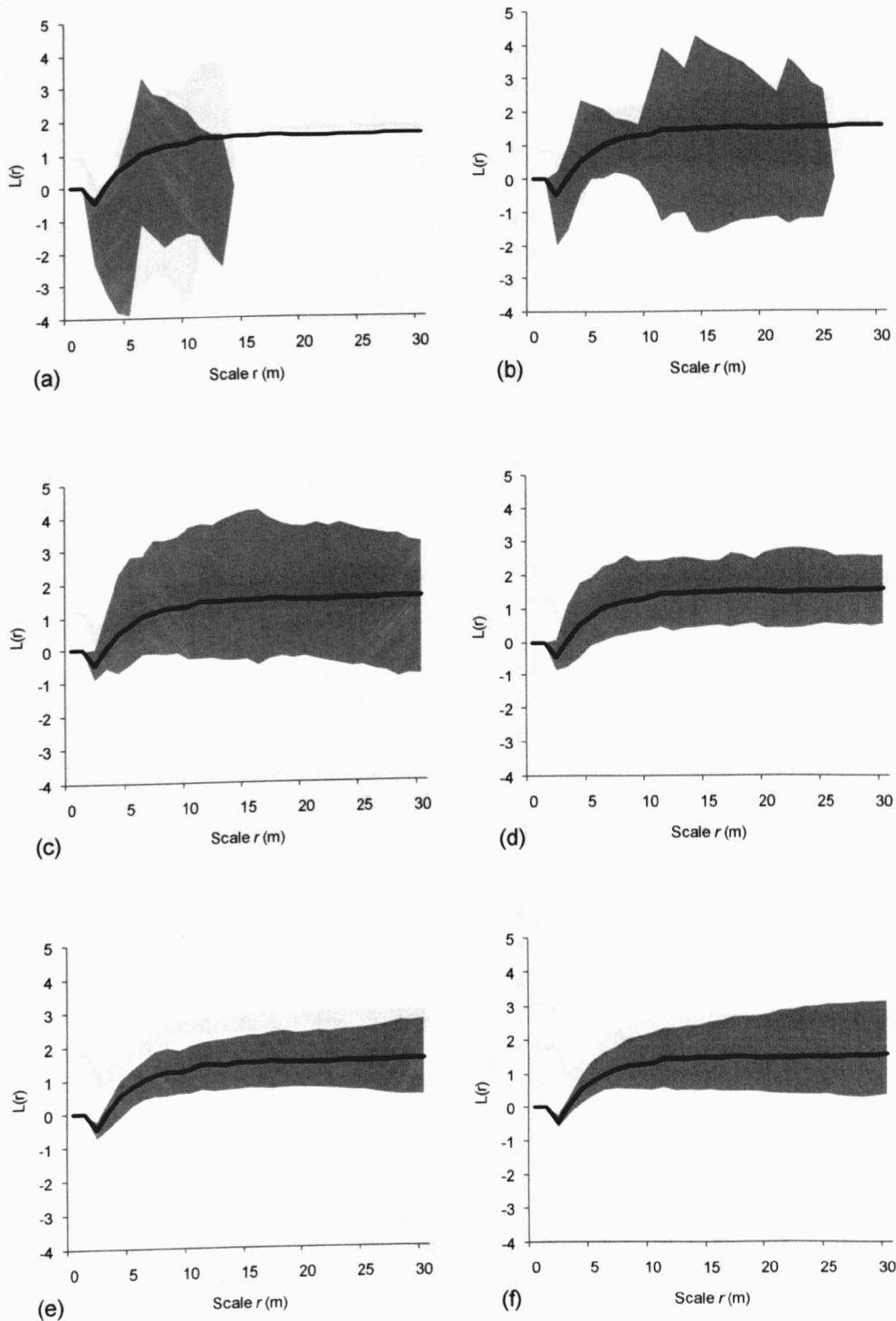


Figure 5-10 The unreliability of  $L(r)$  estimates calculated on Substrate 2 from ten replicate plots at six extents, (a) 0.0625 ha, (b) 0.25 ha, (c) 1 ha, (d) 2.25 ha, (e) 4 ha and (f) 6.25 ha. The grey area represents the standard deviation of  $L(r)$  estimates across ten replicates at each scale. The black line shows the reference  $L(r)$  calculated at 49 ha.

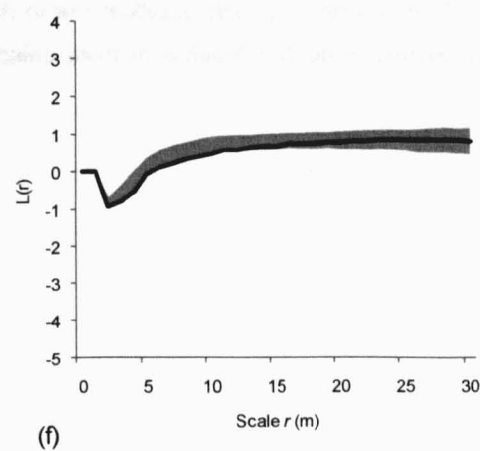
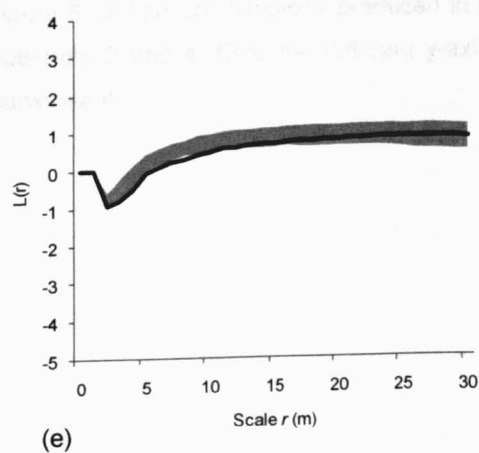
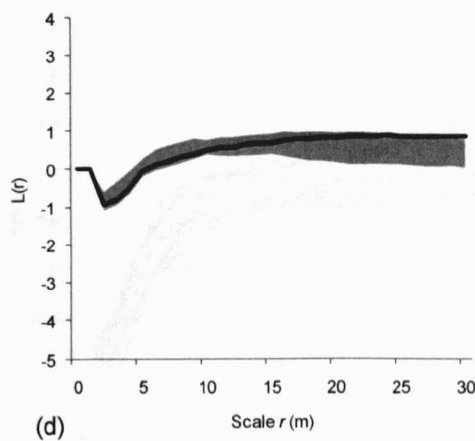
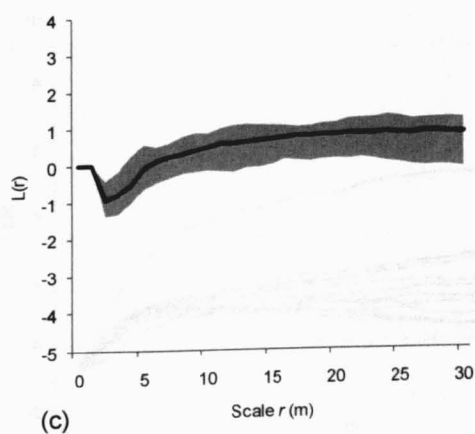
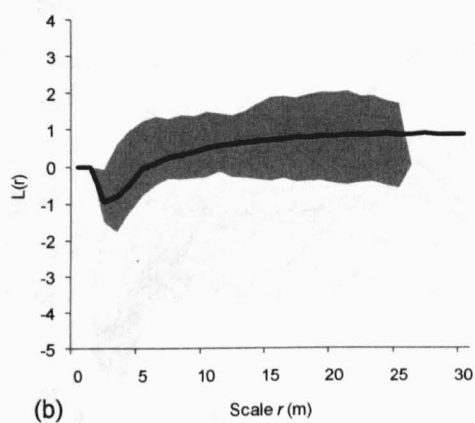
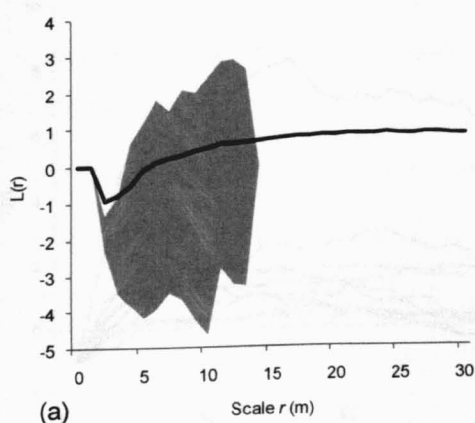


Figure 5-11 The unreliability of  $L(r)$  estimates calculated on Substrate 4 from ten replicate plots at six extents, (a) 0.0625 ha, (b) 0.25 ha, (c) 1 ha, (d) 2.25 ha, (e) 4 ha and (f) 6.25 ha. The grey area represents the standard deviation of  $L(r)$  estimates across ten replicates at each scale. The black line shows the reference  $L(r)$  calculated at 49 ha.

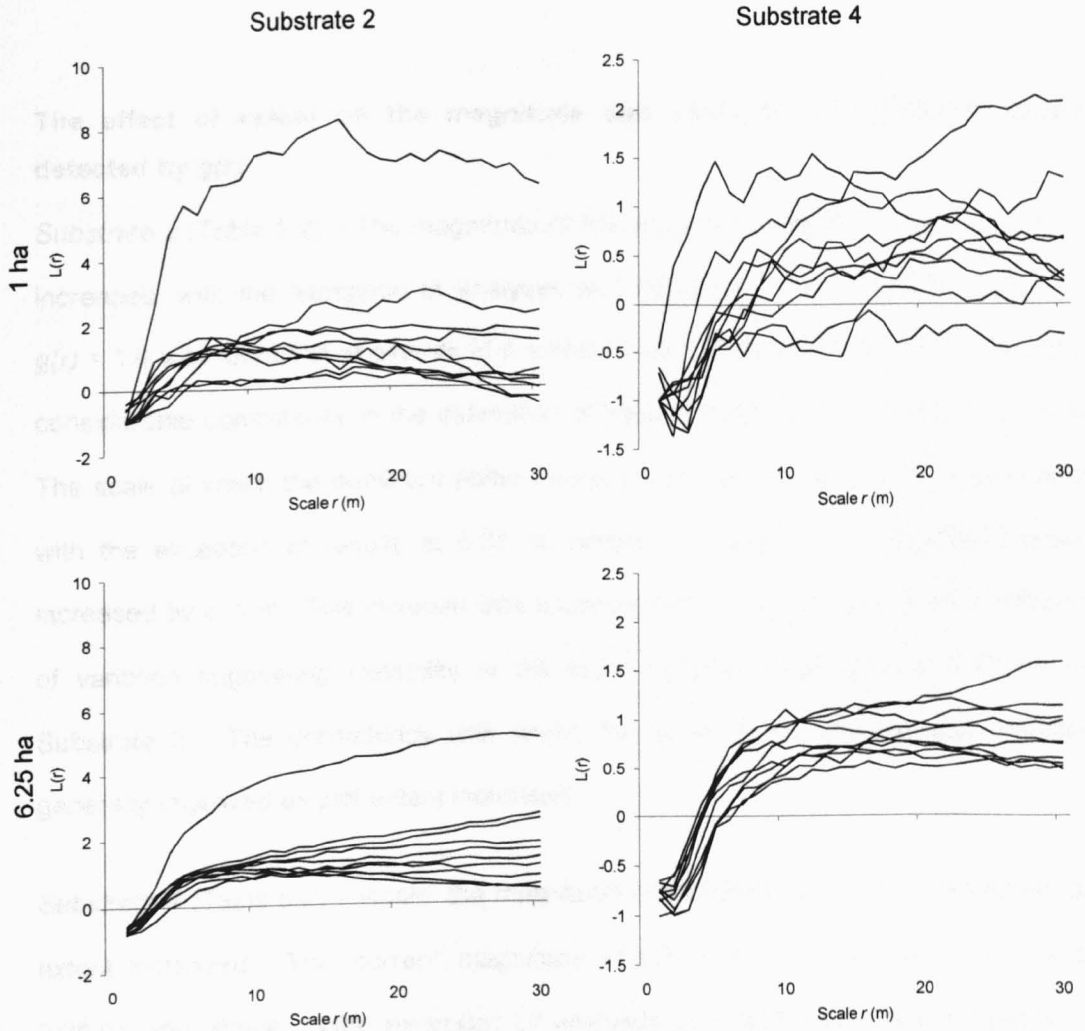


Figure 5-12 The  $L(r)$  functions produced in each of ten replicate plots of 1 ha and 6.25 ha on Substrate 2 and 4. Note the different y-axis scales used to present  $L(r)$  on Substrate 2 and Substrate 4.

**The effect of extent on the magnitude and scale of the dominant pattern detected by  $g(r)$**

*Substrate 2* (Table 5-2) – The magnitude of the dominant pattern decreased as extent increased, with the exception of analyses at 0.0625 ha. The correct magnitude of  $g(r) = 1.6$  was identified at extents of 4 and 6.25 ha. At these extents there was also considerable consistency in the estimation of pattern magnitude across the ten plots. The scale at which the dominant pattern occurs was approximated well at all extents with the exception of results at 0.25 ha where the scale of the dominant pattern increased by c. 1 m. This increase was accompanied by an increase in the coefficient of variation suggesting instability in the scale of peak magnitude at 0.25 ha on Substrate 2. The consistency with which the scale of the pattern was detected generally improved as plot extent increased.

*Substrate 4* (Table 5-2) – Again, the magnitude of the dominant pattern decreased as extent increased. The 'correct' magnitude of  $g(r) = 1.3$  is detected at extents of 2.25 ha and above. With exception of analyses at 0.0625 ha, the consistency of magnitude detection was better on Substrate 4 than on Substrate 2. The scale of the dominant pattern changed considerably with extent. At the three smallest extents the scale of the strongest pattern was over-estimated. This was accompanied by an increase in the coefficient of variation suggesting instability in the identification of pattern scale at these extents.

Table 5-2 The average magnitude and scale of the dominant pattern detected across ten replicates at each of six experimental extents using  $g(r)$ . The mean and coefficient of variation (in parentheses) are reported to 1 dp. The magnitude and scale of the dominant pattern detected at 49 ha is also shown.

Substrate	Extent	Magnitude	Scale
2	0.0625 ha	1.8 (0.4)	4.3 (0.4)
	0.25 ha	2.0 (0.5)	5.3 (0.6)
	1 ha	2.0 (0.6)	4.2 (0.3)
	2.25 ha	1.8 (0.5)	4.3 (0.2)
	4 ha	1.6 (0.2)	4.2 (0.2)
	6.25 ha	1.6 (0.2)	4.1 (0.1)
	49 ha	1.6	4
4	0.0625 ha	1.8 (0.5)	7.2 (0.4)
	0.25 ha	1.6 (0.2)	6.2 (0.7)
	1 ha	1.4 (0.2)	8.1 (0.9)
	2.25 ha	1.3 (0.1)	5.4 (0.3)
	4 ha	1.3 (0.0)	4.8 (0.2)
	6.25 ha	1.3 (0.1)	4.9 (0.2)
	49 ha	1.3	5

#### The effect of extent on magnitude and scale of dominant pattern detected by $L(r)$

*Substrate 2* (Table 5-3) – Although changing extent had little effect on pattern magnitude, the magnitude detected became increasingly consistent as extent increased. At all extents above and including 0.25 ha the magnitude of dominant pattern detected was over-estimated. The magnitude detected was much lower at 0.0625 ha. Increases in extent did, however, improve the approximation of the scale of the dominant pattern.

*Substrate 4* (Table 5-3) – Analyses approximated the reference pattern magnitude at extents of 1 ha and above. These estimates became increasingly consistent as extent increased. The consistency of magnitude detection was generally better on Substrate 4 compared to Substrate 2. Pattern magnitude was over-estimated at 0.0625 and 0.25 ha. Increases in extent appear to cause an increase in the scale of the dominant pattern, although the relationship is not as clear as observed on Substrate 2.

Table 5-3 The average magnitude and scale of the dominant pattern detected across ten replicates at each of six experimental extents using  $L(r)$ . The mean and coefficient of variation (in parentheses) are reported to 1 dp. The magnitude and scale of the dominant pattern detected at 49 ha is also shown.

Substrate	Extent	Magnitude	Scale
2	0.0625 ha	0.7 (1.2)	7.4 (0.3)
	0.25 ha	2.2 (1.2)	12.0 (0.6)
	1 ha	2.2 (1.0)	16.6 (0.5)
	2.25 ha	1.9 (0.6)	22.7 (0.3)
	4 ha	1.8 (0.6)	21.7 (0.4)
	6.25 ha	2.0 (0.6)	22.6 (0.4)
	49 ha	1.6	30.0
4	0.0625 ha	1.7 (1.3)	8.3 (0.3)
	0.25 ha	1.3 (0.8)	17.8 (0.3)
	1 ha	1.0 (0.6)	23.2 (0.3)
	2.25 ha	0.8 (0.4)	15.7 (0.4)
	4 ha	1.0 (0.3)	23.2 (0.3)
	6.25 ha	1.0 (0.3)	19.7 (0.4)
	49 ha	0.9	24.0

### 5.3.4 THE EFFECT OF EXTENT ON PATTERN DETECTION BY MONTE CARLO SIMULATION ENVELOPES

This analysis focuses on extents 1 to 6.25 ha on Substrate 4 as previous analyses suggest these plots have greater consistency in the average quantitative pattern description by both  $g(r)$  and  $L(r)$  compared to equivalent extents on Substrate 2. Maximum stability in quantitative pattern detection across extents was desired so that any differences in qualitative pattern detection via Monte Carlo envelopes could be more certainly attributed to changing extent and/or sample size rather than real differences in spatial pattern.

Both  $g(r)$  and  $L(r)$  detected fewer scales of significant spatial pattern as the plot extent decreased (Figure 5-13). The detection of significant aggregation at fewer scales in small plots could be due to one of two effects: a real decrease in the magnitude of aggregation (i.e., an increase in the empirical function value), or an increase in the width of the simulation envelope (or both) (see Figure 3-3).



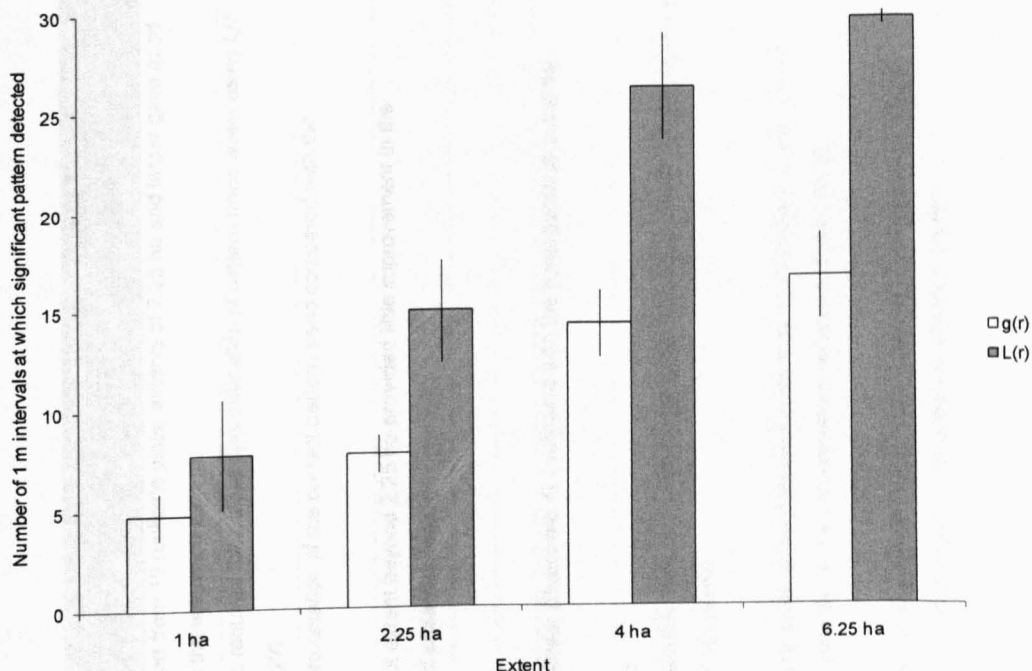


Figure 5-13 Number of scales at which a significant pattern is detected by  $g(r)$  and  $L(r)$  on Substrate 4 at extents of 1 ha and greater, as measured by 99% Monte Carlo simulation envelopes. Averages are based on ten replicates at each extent. Error bars represent the standard error of the mean.

The average simulation envelope width for both  $g(r)$  and  $L(r)$  increased as the extent of the plot decreased. At most scales the increase in the width of the simulation envelope as extent decreased was greater than the change in the function value. Thus, for analyses by both  $g(r)$  and  $L(r)$  the reduced detection of 'significant' spatial pattern in small extents can be primarily attributed to an increase in the width of the simulation envelope rather than an actual change in the magnitude of aggregation. The exception appears to be the increased detection of aggregation between analyses at 2.25 and 4 ha where increases in  $g(r)$  and  $L(r)$  exceeded any decrease in the simulation envelope at most scales suggesting a real increase in aggregation magnitude may be occurring.

Table 5-4 summarises the main findings of this chapter.

Table 5-4 Summary of the main findings in Chapter 5

	SMALL 0.0625 ha			LARGE 6.25 ha		
	0.25 ha			4 ha		
	1 ha			2.25 ha		
The effect of extent on the estimation of pattern trend (Analysis 1).				When estimated from 10 replicate plots, extents of 2.25 ha and larger gave good estimations of the reference pattern.		
				Larger extents required for a close approximation of pattern trend when using $L(r)$ compared to $g(r)$ .		
	At equivalent extents, analyses on the heterogeneous substrate (Substrate 4) provided a closer approximation of the correct pattern trend compared with the homogeneous substrate (Substrate 2).					
The effect of extent on the spatial consistency of pattern detection (Analysis 2a).	The reliability of both functions decreased as plot extent decreased.			Increasing plot extent beyond 2.25 ha provided little improvement in the consistency of pattern detection		
	$L(r)$ and $g(r)$ inconsistent at all scales.	$L(r)$ inconsistency greatest at large scales.				
		$g(r)$ inconsistency was greatest at small scales.				
	At equivalent extents, spatial consistency in pattern detection was better on the heterogeneous substrate (Substrate 4) compared with the homogeneous substrate (Substrate 2).					
The effect of extent on the magnitude and scale of the dominant pattern detected (Analysis 2b).	The magnitude of the dominant $g(r)$ decreased as plot extent increased.					
	Detection of the scale of the dominant pattern by $g(r)$ was worse on the heterogeneous substrate (Substrate 4), especially at smaller plot extents.			On both substrates, detection of the scale of the dominant pattern by $g(r)$ improved as plot extent increased.		
	The magnitude of the dominant $L(r)$ showed little consistent trend with changing plot extent.					
				Detection of the scale of the dominant pattern by $L(r)$ improved as plot extent increased, especially on the homogeneous substrate (Substrate 2).		
				Decreased detection of spatial pattern at small extents attributed to the widening of simulation envelopes because of smaller sample sizes rather than real changes in the detection of spatial pattern.		

## 5.4 DISCUSSION

The majority of studies using second-order spatial statistics to investigate the spatial patterns of woody species use small plot extents (Chapter 1, Appendix A). However, the results presented in this chapter suggest that analyses conducted on data collected from plots of small extent may provide inaccurate descriptions of spatial pattern.

### 5.4.1 THE EFFECT OF EXTENT ON QUANTITATIVE PATTERN DETECTION

On both substrates the greatest agreement with the reference spatial pattern occurred when the function was estimated at large experimental extents. The detection of pattern trend was, however, consistently better on Substrate 4 compared with Substrate 2. The most accurate replication of pattern trend occurred at 6.25 ha on Substrate 2 and at 4 ha on Substrate 4 (this analysis was indistinguishable from analyses at 6.25 ha). Decreasing plot extent had a considerable effect on the spatial pattern identified by  $g(r)$ . Extents of 2.25 ha and greater identified similar spatial structure to analyses at 49 ha, when estimated from ten replicate plots. Thus, if spatial pattern analyses are applied merely as an exploratory technique and interpretation is not intended to go beyond the visual assessment of graphical output, it seems that increasing sample effect beyond extents of 2.25 ha offers little improvement in pattern detection in the current ecosystem, as long as replicate plots are used. On Substrate 4 analyses at 1 ha also provided a good estimation of spatial pattern. Smaller extents, however, did not identify the correct pattern trend. The performance of the different extents differed when investigating  $L(r)$ . For visual interpretation, analysis by  $L(r)$  generally requires greater extents ( $\geq 4$  ha) than analysis by  $g(r)$ . This is attributed to the cumulative nature of  $L(r)$ , as discrepancies in spatial pattern at small scales will be propagated at larger scales.

Unlike Zenner and Peck's (2009) study, this research investigated the patterns detected by  $L(r)$  and  $g(r)$  when applied to windows of different extent. The results suggest that, for accurate pattern description, a much larger extent is required than the 0.5 ha recommended by Zenner and Peck. Furthermore, pattern detection by both

$g(r)$  and  $L(r)$  required much greater plot extents than used in the majority of published studies. An extent of 2.25 ha exceeds the plot size used in 75 (71%) of the studies in Appendix A (for which information on plot extent was available,  $n = 106$ ). The selection of plot extent will largely depend on the study species, with greater areas needed as the size of individual plants increases, and/or plant density decreases. However, even studies of arid shrubs and trees, which are typically low density, have predominantly used small plot extents. Eighteen of the studies reviewed in Appendix A investigate the patterns of woody species in arid ecosystems. Eleven of the 17 studies for which information was available used an average plot extent of less than 2.25 ha. All of these studies used an average plot extent of  $\leq 1$  ha. Of these only three used more than one replicate plot (Haase et al., 1997; Schenk et al., 2003; Meyer et al., 2008).

In the present study system, an extent of 2.25 ha corresponds to an average sample size per plot of 219 and 301 on Substrates 2 and 4 respectively. Summed over ten replicate plots the estimated functions at 2.25 ha were created from a total of 2189 and 3007 individuals on Substrates 2 and 4 respectively. This sample size far exceeds any previous assertions of minimum sample size requirements (Eccles et al., 1999; Plotkin et al., 2000; Malkinson et al., 2003; Wiegand et al., 2007a; Meyer et al., 2008; Jacquemyn et al., 2009). Of the 75 studies using extents less than 2.25 ha (Appendix A), information on sample size was available from 37 studies. In fifteen of these articles the geometric mean sample size per plot fell below the minimum sample size of 70 recommended by Wiegand et al. (2007a). However, 29 studies performed at least one analysis on plots containing fewer than 70 individuals and 19 performed at least one analysis on fewer than 30 individuals. The validity of pattern interpretations generated at such low sample sizes and small extents is questioned. Formal tests comparing the pattern detected when varying numbers of points are removed from a known pattern would help elucidate the role of sample size in pattern detection.

#### **5.4.2 THE SPATIAL CONSISTENCY OF QUANTITATIVE PATTERN DESCRIPTION AT ECOLOGICALLY MEANINGFUL SCALES**

Because of differences in the reliability of analyses by  $g(r)$  and  $L(r)$  they are discussed separately.

##### **The spatial consistency of $g(r)$**

Compared with smaller extents, analyses at 2.25 ha and larger were remarkably spatially consistent (Figure 5-8 and Figure 5-9). At these extents, there was little spatial inconsistency in the detection of dominant pattern magnitude (Table 5-2). Analyses at 4 and 2.25 ha, and larger, on Substrates 2 and 4 respectively, identified dominant patterns of the same strength as analyses at 49 ha (Table 5-2). At smaller extents the magnitude of the dominant pattern in the average plot was over-estimated. Correct approximations of the scale of the dominant pattern were achieved at extents of 4 ha. However, even at this extent the plots varied quite widely in the scale of the dominant pattern. These results suggest that reasonably reliable estimates of spatial pattern could be obtained from individual plots with a minimum extent of 2.25 ha. There was, however, some spatial inconsistency in the scale of the dominant pattern. The scale of the dominant (maximum)  $g(r)$  indicates the modal inter-shrub distance. Inaccurately quantifying this distance could affect interpretations of dominant ecological processes. It is therefore recommended that when using single or few replicates, especially when plot sizes are small, interpretation is limited to the major trends and range of distances indicated by the analysis rather than specific distance classes.

At all plot extents, the greatest inconsistency in  $g(r)$  occurred at small scales (Figure 5-8 and Figure 5-9). Patterns at these scales are presumed to represent interactions between individual shrubs. The inconsistency at small scales increased as plot extent decreased. At the smallest extents (0.0625 and 0.25 ha) the sample sizes are often very small (Appendix D) and thus inconsistency may in part be caused by a lack of statistical power. At 1 ha, however, where the results are assumed to be statistically robust, there is considerable inconsistency in  $g(r)$  at small scales, especially on

Substrate 2. This result is particularly concerning as 1 ha is a commonly employed extent (Figure 1-3). At larger extents ( $\geq 2.25$  ha) the sample size is large enough to assume that the results are statistically robust (Appendix D). This implies, therefore, that the detection of local interactions may vary with the extent of the plot.

Therefore, as plots decrease in extent, the inference of biological processes from single plots may depend increasingly upon the locations sampled. The small-scale inconsistency of  $g(r)$  on Substrate 2 is of particular interest as analyses suggest that this substrate has little large-scale environmental heterogeneity. The results on Substrate 4, despite its apparent heterogeneity, were more spatially consistent. Much current statistical research focuses on methods to remove the confounding effects of large-scale environmental heterogeneity on the spatial signal generated by biological interactions (Law et al., 2009). One of the commonly recommended techniques is to define homogeneous sub-regions within the larger heterogeneous pattern and to calculate the spatial structure within these separately (Pélissier and Goreaud, 2001). These separate functions can be combined into a single 'master' function (Illian et al., 2008). Law et al. (2009) recommend this technique if the aim of the study is to understand spatial autocorrelation at small scales. However, the current study suggests that even within homogeneous subplots of congruent shape and size, there is inconsistency in the detection of small-scale spatial structure. If it is assumed that plots of c. 1 ha contain, on average, a large enough sample size ( $n \approx 100$ ) to produce statistically sound assessments of spatial structure, as published studies suggest, then it appears that there may be real differences in spatial pattern when assessed at this extent. Two possible explanations are suggested: the presence of small-scale heterogeneity, and the effect of remotely sensed data collection. Both these explanations have important consequences for the conduct of spatial pattern studies and are therefore discussed in turn.

### **Remotely sensed data vs. manual field collection**

Although manually mapping plant distributions over small areas may take considerable time, it usually results in high accuracy maps of stem distributions.

Remotely sensed data enable large extents to be mapped quickly. This may, however, come at the expense of fine-scale accuracy.

A major source of error when using remotely sensed data occurs when the centre of mass of the canopy is used to represent stem locations. Canopies are not necessarily symmetrically distributed around the central stem, and effects such as wind-throw may dramatically alter their relative position. If the approximation of rooting point shows spatially systematic error, this could cause spatial inconsistency in the quantification of small scale spatial structures. This must be considered when analysing spatial patterns obtained from remotely sensed data. Surveys in December 2007 and November 2009 did not provide any evidence of wind-throw affecting *S. supranubius* individuals. Furthermore, canopies were observed to be largely symmetrical about the central stem. Therefore, error in the estimation of *S. supranubius* rooting points is considered to be both small and spatially random and should not be driving the observed spatial inconsistency in small-scale patterns.

When using remotely sensed data large extents are recommended to average any spatially systematic error in the data. If remotely sensed data were collected at the extents routinely used in field collections (e.g., Malkinson et al., 2003; Malkinson and Kadmon, 2007) small scale errors in the estimation of rooting locations may over-ride the true signature of plant–plant interaction.

### **Small-scale heterogeneity and the importance of replicates**

Although relatively homogeneous at the large scale, analyses indicate that environmental heterogeneity may be present at small scales. Large-scale heterogeneity has been shown to interact with plant dynamics to generate quantitatively different spatial patterns (Getzin et al., 2008; Chapter 7). It is therefore feasible that heterogeneity at smaller scales could have a similar effect. Individuals may inhabit different resource micro-habitats (Beckage and Clark, 2003). Small plots may sample different micro-habitats, whereas large plots will incorporate multiple micro-habitats and thus average their impact on species distributions.

Manually collecting data on the distribution of arid shrubs may have a dual effect on spatial pattern reliability as the demanding environment may limit the extents covered, which, because of the typically low density populations, may result in inadequate sample sizes. Fourteen of the studies in Appendix A estimated the spatial patterns of woody species in arid areas using manually collected data. Seven of these studies estimated spatial patterns from single plots (per environmental context) of small extent ( $< 0.5$  ha; Haase et al., 1996; Eccles et al., 1999; Zavala-Hurtado et al., 2000; Malkinson et al., 2003; Schurr et al., 2004; Malkinson and Kadmon, 2007; Biganzoli et al., 2009). The remaining studies either used multiple small plots, or single large plots ( $> 2.25$  ha). Spatial pattern reliability may also be a problem in higher density systems, such as temperate forests. In such systems, achieving an adequate sample size will require much smaller areas to be mapped (depending on plant size) than if the same sample size were desired in an arid system. As plot extents decrease the variation in small-scale heterogeneity within each plot will tend to decrease and estimates of spatial pattern will become less reliable representations of the population as a whole.

In the present study, decreasing plot extent resulted in an increased inconsistency in pattern detection by  $g(r)$  and  $L(r)$ . This is attributed to heterogeneity in the small-scale pattern. For example, a plot may have spatially consistent intensity (i.e., homogeneous), but individuals may be arranged in clusters of varying size. Whereas larger plots will detect aggregation at the average cluster size, small plots will pick up the variation in cluster size and so produce seemingly inaccurate results. Clearly, identifying and mapping small-scale heterogeneity in any system is unfeasible in most cases, requiring too much time and inconceivable amounts of knowledge of the species' resource requirements and fine-scale distribution of these resources (but see Chapter 8). Therefore, caution is urged when using small plot extents, even in seemingly homogeneous areas. Greatest concern is raised where plots of small extent are used to infer the operation of biological processes and extrapolate these to the dynamics of much more extensive populations. Unless a considerable portion of



the focal population (population here may refer to a forest or a species on a particular substrate) of interest can be mapped, an accurate estimate of spatial pattern will require replicate plots across the area of interest. However, despite the tendency of published studies to be based on sample plots with small extents, very few studies use replicate plots. More than two-thirds ( $n = 75$ ) of the 109 articles summarised in Appendix A estimated spatial pattern from single plots per environmental context. Of the remaining 34 articles, 15 used only two replicate plots per environmental context.

The number of replicate plots should increase as the extent of those plots decreases. In the current system, the reliability analyses reveal a strong need to increase the number of replicate plots used once the extent of the plots decreases from 2.25 to 1 ha. However, when mapping individuals by hand, increasing plot replication over large and/or multiple areas will considerably increase the time and financial costs of the study. In situations where spatial patterns are necessarily quantified from single small plots it is strongly recommended that interpretations of patterns are cautious and are not over-extrapolated. It is also recommended that strong inferences about ecological processes are not made from spatial structure alone. Where possible, spatial pattern investigation and interpretation should be supported with *a priori* knowledge, with the ultimate aim of experimentally verifying the operation of inferred processes (Perry et al., 2006).

### **The spatial consistency of $L(r)$**

Spatial inconsistency in spatial pattern detection by  $g(r)$  was greatest at small scales. When assessed by  $L(r)$  spatial inconsistency in pattern detection was greatest at large scales reflecting the accumulation of inconsistencies occurring at smaller scales. As inconsistencies in pattern detection by  $L(r)$  are primarily at large scales, it is recommended that  $L(r)$  is only interpreted up to the scales of interest derived from *a priori* hypotheses.

### 5.4.3 THE EFFECT OF EXTENT ON PATTERN DETECTION BY MONTE CARLO SIMULATION ENVELOPES

As the extent of the plot decreased, 'significant' aggregation was detected at fewer scales by both  $g(r)$  and  $L(r)$ . This is attributed to an increase in the width of the simulation envelope rather than an actual decrease in the magnitude of aggregation, with the possible exception of increased detection of aggregation as extent increased from 2.25 to 4 ha, which may be driven by real increases in the empirical  $g(r)$  and  $L(r)$ . Increasing sample size narrows the simulation envelopes as a random simulation of a few individuals shows a greater degree of apparent spatial variation than a random simulation of many individuals. This research suggests that the lower sample sizes associated with smaller extents will increase the width of the simulation envelope and may prevent notable patterns from being detected (i.e., Type II error). This effect was much more pronounced when analysing patterns with  $L(r)$  because of the cumulative nature of the statistic.

Kenkel (1988) pioneered the use of Monte Carlo simulation envelopes to detect spatial pattern using  $K(r)$  (Loosmore and Ford, 2006). This method of interpreting spatial point pattern statistics dominates the literature (Section 1.3.4). Furthermore, the majority of articles limit interpretation to the scales at which the empirical function exceeds the simulation envelope limits, with no assessment of the height of the curve in relation to other curves or the values expected under the null model. Only six of the articles in Appendix A compare the magnitude of empirical function curves (see Section 1.3.4). Of these studies, the majority visually compared curves, with only one study (Peterson and Squiers, 1995) making numerical comparisons. Studies basing analyses and interpretations solely upon Monte Carlo techniques may be at risk of misinterpreting the range of spatial autocorrelation. Furthermore, the strong dependence of Monte Carlo simulation envelopes on sample size will make it difficult to make meaningful comparisons of spatial patterns formed from different numbers of individuals.

The reliance of spatial pattern analyses on Monte Carlo simulation envelopes has been previously questioned (Diggle, 2003). Loosmore and Ford (2006) present a critique of the use of Monte Carlo simulation envelopes claiming that their use to determine whether, and at what scale, an observed pattern deviates from a specified null model is invalid because of incorrect type 1 error rates. Simulation envelopes are constructed from, at every distance, the maximum and minimum (or, e.g., the 5<sup>th</sup>-highest and 5<sup>th</sup>-lowest) value of the pattern statistic ( $g(r)$  or  $L(r)$ ) calculated from a number of simulated patterns. Consequently, simulation envelopes are constructed from the results of many simulated patterns, each contributing to the envelope over different distances. Therefore, many tests are being performed concurrently at each distance class (Loosmore and Ford, 2006). This simultaneous inference yields underestimated Type I errors rates and can lead to Type I errors being made, especially when the empirical function is close to the simulation envelope.

Studies have recommended alternative techniques, such as the accumulated deviation of an observed function from a theoretical statistic, to assess the departure of an observed pattern from a specified null model (e.g., Plotkin et al., 2000; Law et al., 2009). Despite these recommendations, the majority of studies continue to assess spatial pattern solely by comparing observed functions to Monte Carlo simulation envelopes. This chapter recommends that Monte Carlo envelopes are treated as an analogue to statistical assessment via p-values; they provide an indication of the importance of the pattern. Primary interest, however, should be in biological importance, which may be assessed using the magnitude of an effect rather than its statistical significance (Nakagawa and Cuthill, 2007). Monte Carlo simulation envelopes do not allow for the calculation of statistical power or effect sizes. Therefore, analysis of pattern by Monte Carlo envelopes should be supplemented by the direct comparison of empirical functions, and examination of the deviation the observed pattern has from its distribution under the null model, an analogue of effect sizes.

## 5.5 CONCLUSIONS

Part of the allure of techniques such as  $L(r)$  and  $g(r)$  is the apparent ease with which they are able to provide information on the scale of departure of an observed pattern from a hypothetical spatial model (Loosmore and Ford, 2006). However, the increasing use of spatial point pattern statistics to gain ecological insights into the processes driving plant population structure has not been accompanied by critical analysis of the accuracy of the techniques, resulting in much variation in the methods and their application. This chapter investigated how variations in plot extent (and associated variations in sample size) affect both the quantitative and qualitative results of spatial point pattern analysis by  $L(r)$  and  $g(r)$ .

According to previously published studies, sample sizes in excess of 70 individuals should provide reliable estimates of spatial pattern when estimated by  $g(r)$  and  $L(r)$ . In the present system this corresponds to a plot of c. 0.7 ha in extent. However, analyses suggest that larger plot extents (a minimum of 2.25 ha) are needed to obtain accurate descriptions of *S. supranubius* spatial pattern because of spatial inconsistency in the quantitative detection of small-scale pattern. These inconsistencies were most pronounced at the smallest extents, but still evident when plots contained relatively large sample sizes ( $n > 100$ ). Assuming the statistical detection of pattern is robust at these sample sizes, it is hypothesised that on both substrates these inconsistencies arise from the effects of undetected heterogeneity at small scales. Mapping the distribution of small-scale heterogeneity is not usually feasible, especially if knowledge of the focal species life history is limited. Furthermore, small-scale heterogeneity will not necessarily be present in all systems, and the scale of any heterogeneity will be location dependent. Therefore, it is impossible to define a plot extent at which the effect of small-scale heterogeneity will be averaged out. Consequently, a cautious approach is recommended to prevent the incorrect extrapolation of localised patterns to whole communities. Randomly distributed replicate plots of congruent shape and size should be used, even in areas that appear to be homogeneous at broad scales. When sample plots are small and

there is no potential to either increase their size or number, researchers are urged to ensure their interpretations only consider the major pattern trends and scales, and should acknowledge the potential effect of location on their results.

When assessing the 'significance' of a pattern, it is recommended that Monte Carlo envelopes are treated as an analogue to statistical assessment via p-values; i.e., they provide an indication of the importance of the pattern. This assessment should be supplemented by direct comparison of empirical functions, and examination of the deviation the observed pattern has from its distribution under the null model, an analogue of effect sizes. Assessments of this kind are vital if spatial patterns are to be meaningfully compared.

## CHAPTER 6: THE CONSEQUENCES OF POINT VERSUS REAL-SHAPE APPROXIMATION IN SPATIAL PATTERN ANALYSIS

---

### 6.1 INTRODUCTION

Almost all studies of plant population spatial structure approximate the locations of individuals as dimensionless points (see Appendix A). Such techniques are probably favoured because of the relative ease with which data can be collected, and the reliance of most software on this data format. Plants, however, are not dimensionless; they are discrete entities occupying a non-zero, finite space within the landscape. The space they occupy, both above- and below-ground, largely determines the intensity and scale of biological processes operating between individuals. It is only in recent years that researchers have begun to consider the statistical and interpretative consequences of representing individuals as points (Wiegand et al., 2006; Barbeito et al., 2008; Muller-Landau et al., 2008; Rossi et al., 2009). Point approximation is now considered by some to be one of the major limitations in contemporary fine-scale ecological pattern analysis (Wiegand et al., 2006).

Representing individuals as points is valid when the size of the plants is small relative to the spatial scales being studied. However, most ecologists are typically concerned with the interactions occurring at small scales where plant–plant interactions are assumed to occur. At these scales point approximation may obscure real spatial structures. There are three primary consequences of representing individuals as points instead of objects: the hard-core distance and soft-core effect (Matérn, 1986) and the aggregation effect (Wiegand et al., 2006; see also Fehmi and Bartolome, 2001). Wiegand et al.'s (2006) software Programita provides two techniques to control for these effects. These effects and techniques are summarised below.

#### **The hard-core distance**

The hard-core distance (HCD) applies in populations of non-overlapping individuals. As the distance under consideration by  $g(r)$  decreases, the HCD is the radius at which

the pattern detected begins to be influenced by the fact that the shrubs cannot overlap. In the simplest situation of a uniform population of shrubs with circular canopies, the HCD corresponds to the canopy diameter (Figure 6-1a). If approximated as points, randomisation during Monte Carlo simulations may locate individuals closer together than their canopy extents would in reality allow (Figure 6-1b). Consequently, distances less than the canopy diameter will occur less frequently in the observed pattern than expected under a random distribution leading to the incorrect detection of dispersion at small scales. In the more realistic case of a mixed size population, the HCD corresponds to the maximum canopy diameter. The populations used in this chapter (and thesis) are of mixed size. Therefore, the hard-core distance refers to the canopy diameter of the largest individuals. Consequently, it is possible for real interactions to occur at scales less than the hard-core distance, although the probability of such interactions occurring will be small and will decrease rapidly as the distance considered by  $g(r)$  decreases. At scales below the hard-core distance a soft-core effect occurs (see below). This definition is in contrast to the description of hard-core distances and hard-core effects in Wiegand et al. (2006) which apply to the less realistic situation of uniform shrub populations.

### **The soft-core effect**

Soft-core effects occur at all distances below the HCD in populations of non-overlapping individuals of variable size (Figure 6-1c). As the distance considered by  $g(r)$  decreases the effect of the inability of shrubs to overlap on the pattern detected increases. When individuals are approximated as points, individuals that are separated by short distances will occur less frequently than expected under a random distribution. Consequently, the soft-core effect produces an ever more pronounced dispersion of shrubs as distance decreases (Figure 6-1d).

### **Non-overlapping Monte Carlo simulations**

Both the hard- and soft-core effect can be mitigated by preserving the size and shape of individuals (i.e., real-shape analysis), and by preventing the overlap of individuals during Monte Carlo simulations (see Section 6.2.5).

### **Aggregation effect**

An aggregation effect may occur when non-overlapping individuals vary greatly in size. Large individuals may occupy a considerable proportion of the available space, forcing smaller individuals into the intervening gaps (Figure 6-1e). If the pattern of small individuals were analysed, the restriction of small individuals to gaps between large individuals may be incorrectly diagnosed as aggregation when in fact they may be regularly distributed within the available space (Figure 6-1e). Thus, calculation of  $g(r)$  may be distorted if the spatial structure of older individuals is non-random at the scales at which plant–plant interactions may be occurring (Wiegand et al., 2006).

### **Space restriction (free-space analysis)**

Wiegand et al.'s (2006) grid-based software Programita provides a technique of space restriction to control for the aggregation effect. The space available for individuals to establish, and therefore their spatial structure, may be limited by the presence of older individuals. By masking the location of older individuals and preventing individuals being placed in occupied locations during Monte Carlo simulations, the “genuine” pattern of the younger individuals may be quantified (Figure 6-1f).

Considering the size and shape of individuals is important if accurate interpretation of the small-scale spatial structure of the community is required (Purves and Law, 2002). The interpretative effects described above are most extreme when populations consist of non-overlapping individuals, such as most arid shrubs, especially when individuals reach large sizes, as observed in *S. supranubius*.

Throughout this chapter the following abbreviations are used:



$g_{pp}(r)$  – the pair-correlation function calculated from a pattern of shrubs represented as points (point analysis).

$g_{rs}(r)$  – the pair-correlation function calculated from a pattern of shrubs where the size and shape of individuals has been preserved (real-shape analysis).

$g_{fs}(r)$  – the pair-correlation function calculated from a pattern of shrubs where the size and shape of individuals has been preserved and the space made inaccessible by larger individuals has been masked (free-space analysis).

$HCD$  – the hard-core distance.

$HCD_g$  – the hard-core distance estimated from the convergence of  $g_{pp}(r)$  and  $g_{rs}(r)$  (see Section 6.2.2).

### 6.1.1 PREVIOUS STUDIES

The earliest attempt to account for the size of objects when investigating their spatial

organisation was by Clark & Westoby (1959). They used a point pattern analysis to investigate the spatial distribution of plants in a field. They found that the distribution was not random, but that there was a significant aggregation of plants. This was due to the fact that the plants were of different sizes, and that the larger plants were more likely to be aggregated. This is an example of the 'aggregation effect'.

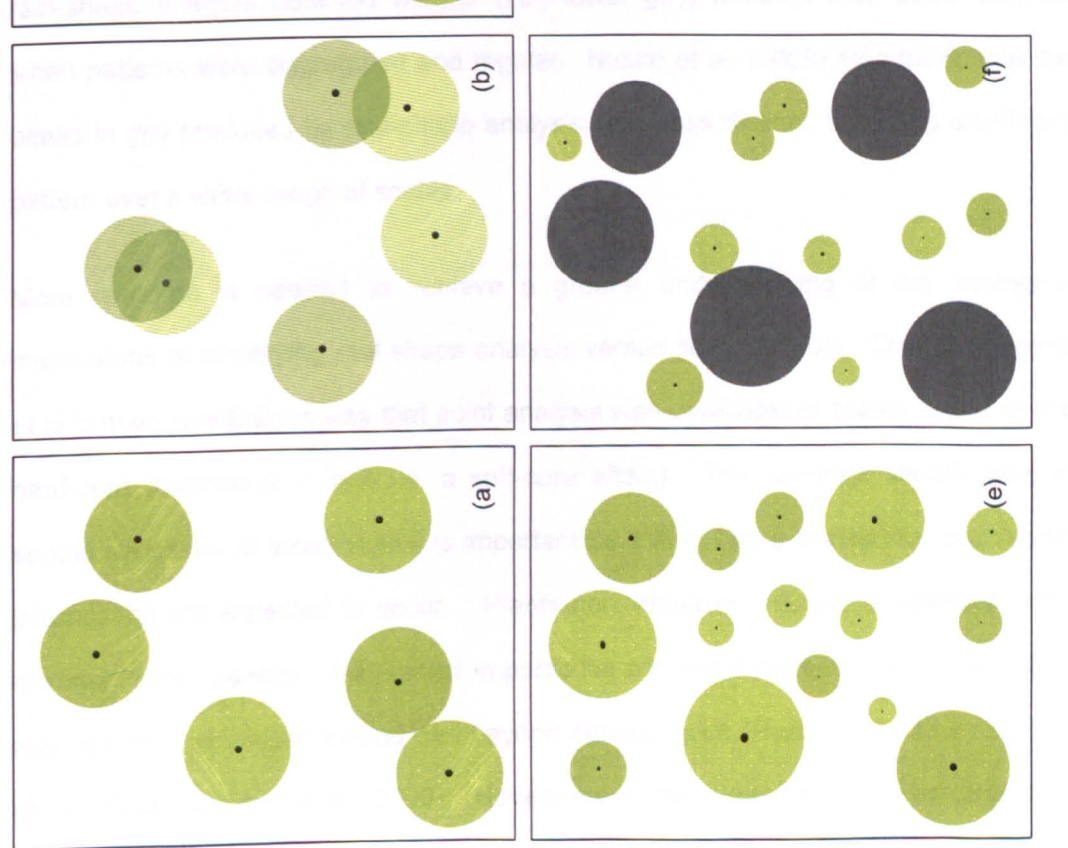


Figure 6-1 The effects of using 'real shape' analysis compared with point approximation. The hard-core distance (a) may cause point analyses to incorrectly identify dispersion at scales less than the hard-core distance (b). Similarly, the soft-core effect (c), may cause point analysis to detect dispersion at small scales (d). Large individuals may restrict the establishment of younger individuals to intervening available space. When analysed on their own, this may cause young individuals to appear aggregated when they may in fact be regularly distributed within the available space (e). This so-called 'aggregation effect' may be controlled for by masking (black circles) the location of inaccessible space (f).

### 6.1.1 PREVIOUS STUDIES

The earliest attempt to account for the size of objects when investigating their spatial organisation was provided by Simberloff (1979), who developed a corrected form of Clark and Evans' (1954) nearest neighbour statistic. A few years later, Prentice and Werger (1985) adapted the CSR null model to account for the average size of objects during calculation of Ripley's  $K(r)$ . However, despite the rapid growth in the application of second-order pattern analyses (such as Ripley's  $K(r)$ ) and acknowledgement of the problems associated with point approximation, only three studies (to my knowledge) have since attempted to extend traditional pattern analysis to consider objects' sizes and shapes (Wiegand et al., 2006; Barbeito et al., 2008; Nuske et al., 2009). Both Wiegand et al. (2006) and Barbeito et al. (2008) concluded that point pattern analyses detected weaker spatial structure than real-shape analysis, and that the differences between the two types of analysis increased as the size of the individuals increased. Using simulated data, Nuske et al. (2009), however, found that real-shape analysis detected weaker (i.e., lower  $g(r)$ ) patterns than point analysis when patterns were aggregated and regular. Nuske et al. (2009) also found that the peaks in  $g(r)$  produced by real-shape analysis were less distinct, detecting significant pattern over a wider range of scales.

More research is needed to achieve a greater understanding of the ecological implications of employing real-shape analysis versus point analysis. One of Wiegand et al.'s main conclusions was that point analysis was unreliable at scales less than the hard-core distance ( $r < hcd$ , i.e. a soft-core effect). The accurate identification of spatial structures at small scales is important as it is at these scales that plant–plant interactions are expected to occur. Plants can, however, interact at scales beyond the limit of their canopy. Arid shrubs in particular are noted for their laterally extensive root systems, which can extend well beyond canopy limits (Hartle et al., 2006; Barbier et al., 2008; Caldwell et al., 2008). Nevertheless, there has been little emphasis on whether and how point approximation affects the detection of spatial structures at scales exceeding the hard-core distance. This may result from uncertainty in

detecting the hard-core distance. In this chapter a technique for detecting the hard-core distance is proposed (see Section 6.2.2), allowing analyses to consider the effects of point approximation versus real-shape analysis on the detection of spatial structure both above and below the hard-core distance.

This chapter also investigates the effect of masking inaccessible space on the patterns detected. Wiegand et al. (2006) provide the only other investigation, of which I am aware, of the potential impact of masking inaccessible areas. In their study the locations of other dominant species were masked whilst the patterns of the focal species were investigated. They found that dispersive patterns were weaker when the areas covered by other species were masked. Unlike Wiegand et al.'s (2006) study, this chapter masks the locations believed to be made inaccessible by older *S. supranubius* individuals, under the assumption that larger individuals are also older.

### **6.1.2 AIMS AND OBJECTIVES**

This chapter investigates the quantitative differences in the pattern detected by three data representation techniques: point analysis, real-shape analysis and free-space analysis. Three hypotheses are outlined:

*Hypothesis 1:* The method of data representation (i.e. point or object) will affect the type of pattern detected (i.e., dispersed, random or aggregated).

*Hypothesis 2:* The method of data representation will affect the magnitude and scale of the strongest pattern.

*Hypothesis 3:* The method of data representation will affect the interpretation of ecological processes.

## **6.2 METHODS**

### **6.2.1 DATA COLLECTION**

The following analyses were performed on the *S. supranubius* individuals on Substrate 2 (49 ha plot; see Section 2.3.2). The point locations of individuals were

estimated by taking the co-ordinates of the centre of the canopy. The size and shape of individuals for real-shape and free-space analyses were taken from the classified imagery with a resolution of 1 m<sup>2</sup>.

### 6.2.2 IDENTIFYING THE HARD-CORE DISTANCE

When the distribution of a population of sparse, non-overlapping shrubs of mixed size is represented as a point pattern, soft-core effects will cause  $g(r)$  to be underestimated at scales below the hard-core distance. Conversely, when the size and shape of individuals is preserved the  $g(r)$  will be overestimated at scales below the hard-core distance because there is an increasing chance of the presence of points within the same bush as the focal bush. Consequently, the  $g(r)$  produced by real-shape analysis ( $g_{rs}(r)$ ) and by point pattern analysis ( $g_{pp}(r)$ ) will differ at scales up to the hard-core distance. Specifically:  $g_{pp}(r) < g_{rs}(r)$  for  $r < HCD$

Assuming that only a very small proportion of canopies within the population will be in contact with neighbouring canopies,  $g_{rs}(r)$  and  $g_{pp}(r)$  are expected to become equivalent at, or close to, the hard-core distance:  $g_{pp}(r) = g_{rs}(r)$  for  $r \approx HCD$

Thus the scale at which  $g_{rs}(r)$  and  $g_{pp}(r)$  converge provides an estimate of the hard-core distance. This method for identifying the hard-core distance only applies in univariate analyses and will work best for sparse communities. *S. supranubius* data were used to test this method (Figure 6-2). Individuals were divided into different datasets of different canopy area. The hard-core distance was estimated ( $\hat{HCD}$ ) as the canopy diameter of the largest individual within the dataset (see Section 6.1). Both a real-shape and a point pattern analysis were run on each dataset. The hard-core distance was then estimated graphically ( $\hat{HCD}_g$ ) as the scale of convergence between  $g_{rs}(r)$  and  $g_{pp}(r)$ . In all datasets the graphical estimate of the HCD corresponded (within 0.5 m) with the maximum canopy diameter.

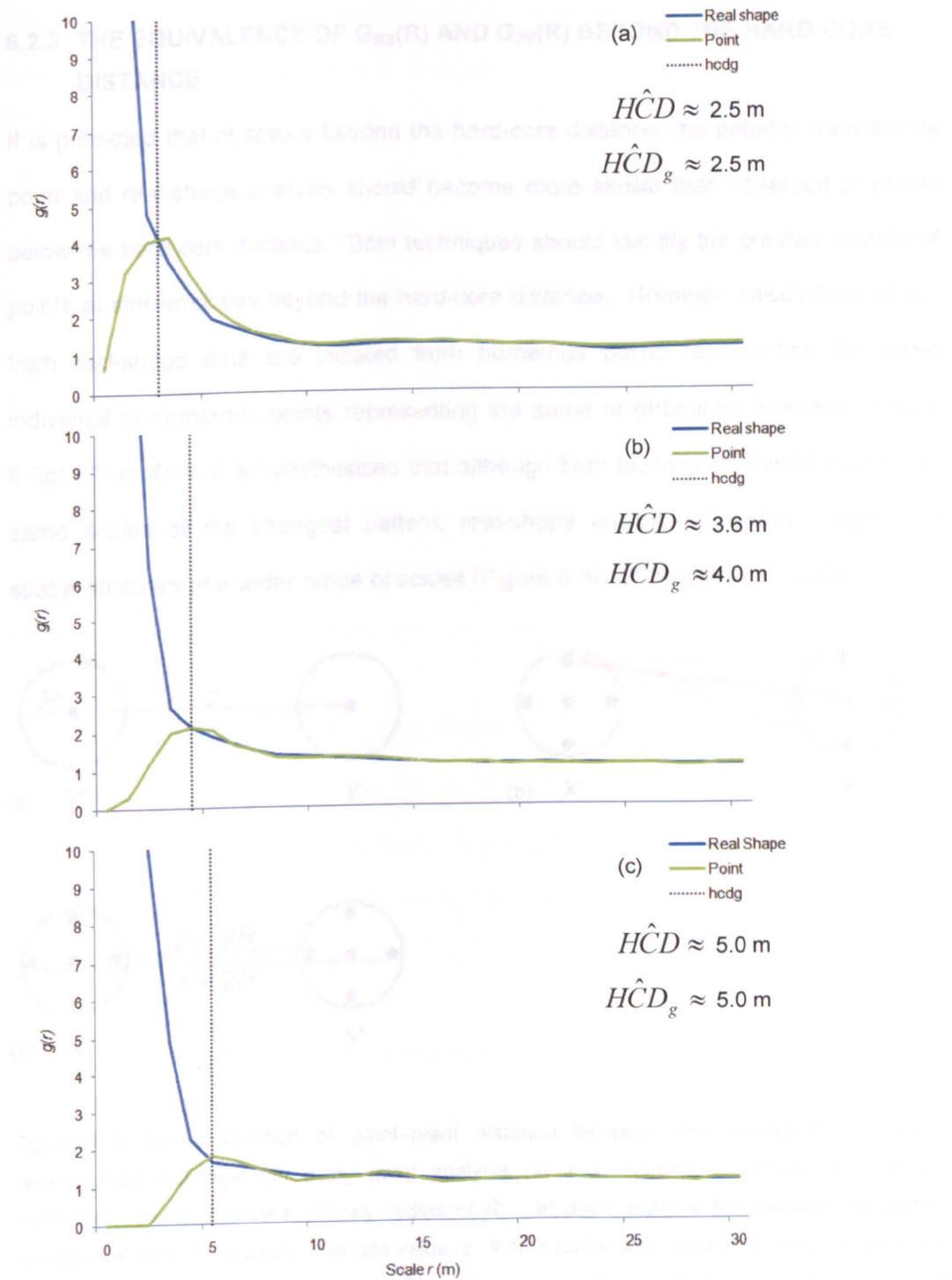


Figure 6-2 Calculation of the hard-core distance from 1) the maximum observed canopy diameter ( $\hat{HCD}$ ) and, 2) from the scale of convergence of the real-shape and point  $g(r)$  ( $\hat{HCD}_g$ ). The vertical lines illustrate the scale of convergence of the real-shape and point  $g(r)$  for individuals with canopy areas between (a)  $1 \leq x \leq 5$ , (b)  $5 \leq x \leq 10$ , and (c)  $10 \leq x \leq 20$ , where  $x$  is canopy area in  $m^2$ . The estimates of  $\hat{HCD}$  and  $\hat{HCD}_g$  are shown on each graph.



### 6.2.3 THE EQUIVALENCE OF $G_{RS}(R)$ AND $G_{PP}(R)$ BEYOND THE HARD-CORE DISTANCE

It is predicted that at scales beyond the hard-core distance, the patterns detected by point and real-shape analysis should become more similar than observed at scales below the hard-core distance. Both techniques should identify the greatest density of points at similar scales beyond the hard-core distance. However, calculations of  $g(r)$  from real-shape data are initiated from numerous points representing the same individual to numerous points representing the same neighbouring individual (Figure 6-3b). Therefore, it is hypothesised that although both techniques should identify the same scales of the strongest pattern, real-shape analysis may detect significant spatial structure at a wider range of scales (Figure 6-3c; cf. Nuske et al., 2009).

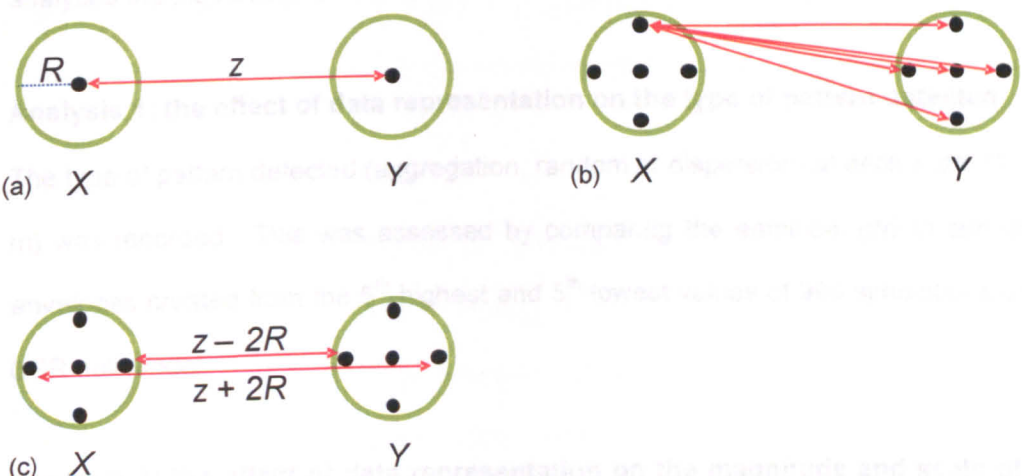


Figure 6-3 The calculation of plant–plant distance between one individual (X) and a neighbouring individual (Y) using point analysis (a) and real-shape analysis (b) and (c). Individuals X and Y have a canopy radius of  $R$ . In point analysis the distance separating individual X from Y is a single, discrete value,  $z$ . When real-shape analysis is used, X and Y are composed of multiple adjacent points. The distance separating X from Y is no longer a single value, but multiple distances separating each point in X from each point in Y (b). Although the average distance separating X and Y in (b) is still  $z$ , a high density of points will be detected at scales ranging from  $z - 2R$  to  $z + 2R$  (c).

### 6.2.4 ANALYSES

The third hypothesis addressed in this chapter is that the method of data representation will affect the interpretation of ecological processes. As demographic

processes are often life-stage specific, the signature of certain processes will only be evident in the spatial pattern of individuals of a certain age. Understanding how patterns change with age can also provide important information on processes such as intra-specific competition. No information on the age of individuals is available. The size of individuals is, however, considered to be a good approximation of relative age. Consequently, all *S. supranubius* individuals were assigned to one of three size classes before analysis: small, medium-sized or large. These size classes correspond to canopy areas ( $x$ ) of  $1 \leq x < 2.8 \text{ m}^2$ ,  $2.8 \leq x < 20.6 \text{ m}^2$  and  $20.6 \leq x \text{ m}^2$  respectively. The size classes had sample sizes of 432, 1711 and 329 respectively. Further information on, and justification of these size class divisions can be found in Chapter 7 (see Section 7.3.2). Analyses were performed on each size class respectively. In all analyses the distribution of individuals was compared with the null model of CSR.

#### **Analysis 1: the effect of data representation on the type of pattern detected**

The type of pattern detected (aggregation, random or dispersion) at each scale (1 – 30 m) was recorded. This was assessed by comparing the empirical  $g(r)$  to simulation envelopes created from the 5<sup>th</sup>-highest and 5<sup>th</sup>-lowest values of 999 simulations of the CSR null model.

#### **Analysis 2: the effect of data representation on the magnitude and scale of the strongest aggregation**

The height of the empirical  $g(r)$  above the CSR expectation ( $g(r) = 1$ ) can be used to quantify the strength of aggregation at any one scale (e.g., Barbeito et al., 2009; Getzin et al., 2008). Larger magnitudes of  $g(r)$  indicate stronger patterns. An alternative technique defines the strongest aggregation as occurring at the scale at which the distance between the empirical  $g(r)$  and the upper simulation envelope is greatest (Wiegand et al., 2006). Another possible technique would calculate the ratio of the distance between the empirical  $g(r)$  and the upper simulation envelope and the width of the simulation envelope, with larger ratios indicating stronger patterns. All three measurements are employed and compared in this chapter.



### **Analysis 3: the effect of data representation on the interpretation of ecological processes**

In addition to the patterns detected within each size class, understanding how patterns change with the age/size of individuals can provide important information on the processes structuring the population. For each technique (point, real-shape and free-space) the empirical  $g(r)$  for each size class was plotted simultaneously.

#### **6.2.5 ANALYTICAL PROCEDURES**

##### **Point pattern analyses**

Because of differences in data density, different ring widths were used to construct  $g(r)$  in the different size classes. Analysis of the small, medium-sized and large size classes used ring widths of 3, 2 and 4 m respectively, as these produced relatively smooth functions in all cases, while maintaining detail at small spatial scales.

##### **Real-shape analyses**

The grid-based software Programita (Wiegand and Moloney, 2004) has been extended to enable the analysis of objects (Wiegand et al., 2006). Individual shrubs are approximated by a group of adjacent cells on a categorical raster map. Each shrub may occupy several adjacent cells, depending upon its size and shape. As cell size decreases, the accuracy with which the size and shape of canopies can be mapped increases. However, very small cell sizes increase computational time. Therefore, a minimal resolution should be selected, depending upon the biological questions being addressed. In the present analysis, the cell size was equivalent to the smallest *S. supranubius* canopy area (i.e., 1 m<sup>2</sup>). A formal point pattern is generated from the categorical raster map. A point is created at the centre of each cell that is part of a shrub. Therefore, whereas the smallest individuals (1 m<sup>2</sup>) are represented by a single point, larger individuals may be represented by several adjacent points. The number of points is therefore much higher than the number of objects, and consequently much higher than the number of points in the conventional point pattern analysis. For calculation of the real-shape  $g(r)$  a ring width of 2 m was used when

analysing the pattern of small individuals, and a ring width of 1 m was used in the analysis of medium-sized and large individuals. These ring widths produced relatively smooth functions, while maintaining detail at small spatial scales.

Simulations of the CSR null model are constructed by rotating (by 0, 90, 180 or 270 degrees, with equal probability), mirroring (or not) and shifting the location of individual shrubs. Field observations of the focal system confirm that *S. supranubius* canopies rarely overlap. Therefore, objects were not allowed to overlap during Monte Carlo simulations of CSR, providing a more realistic simulation of the conditions observed in the field. Unlike points, randomised objects may fall partially outside the study region during Monte Carlo simulations. This would reduce the proportion of occupied cells in the null model (reducing first-order intensity), producing a (positive) bias towards aggregation (Wiegand et al., 2006). Programita provides several options for mitigating this effect (see Wiegand et al., 2006). In the present analyses a torus correction was applied, which wraps the window so that individuals that overlap the window edge also appear on the opposite side of the sample window. This correction, however, breaks individuals into two smaller individuals which can create a bias towards aggregation (Wiegand et al., 2006). Therefore, guard areas were also applied which prevented calculation of  $g(r)$  in the edges of the window. The width of the guard area should be selected so that it is greater than the diameter of most plants, but not so big that it dramatically reduces the sample window extent. Therefore, guard areas were selected to exceed the diameter of 90% of the individuals within each size class. This corresponded to guard area widths of 2 cells, 5 cells and 8 cells when analysing the small, medium-sized and large individuals respectively.

### **Free-space analyses**

As with the real-shape analysis, individuals were represented on a raster map with a resolution of 1 m<sup>2</sup>, and shrubs were not allowed to overlap during Monte Carlo simulations of the null model. Further to the real-shape analysis, the calculation of  $g(r)$  and the simulations of the null model excluded areas that were considered to be inaccessible to the individuals being studied. Thus, when analysing the distribution of

the small individuals, the locations of the medium-sized and large individuals were masked. Similarly, when analysing the pattern of the medium-sized individuals the locations of the large individuals were masked. Free-space analysis was not performed for large individuals as the small and medium-sized individuals are assumed to be younger and therefore could not have driven the establishment of the large individuals (i.e., the analysis would be equivalent to the real-shape analysis of large individuals).

## **6.3 RESULTS**

### **6.3.1 THE EFFECT OF DATA REPRESENTATION ON THE TYPE OF PATTERN DETECTED**

#### **Small individuals**

All three analyses identified aggregation of small individuals at almost all scales, both above and below the hard-core distance of 2 m. The functions produced by the real-shape and free-space analyses are noted for being almost identical in shape (Figure 6-4a).

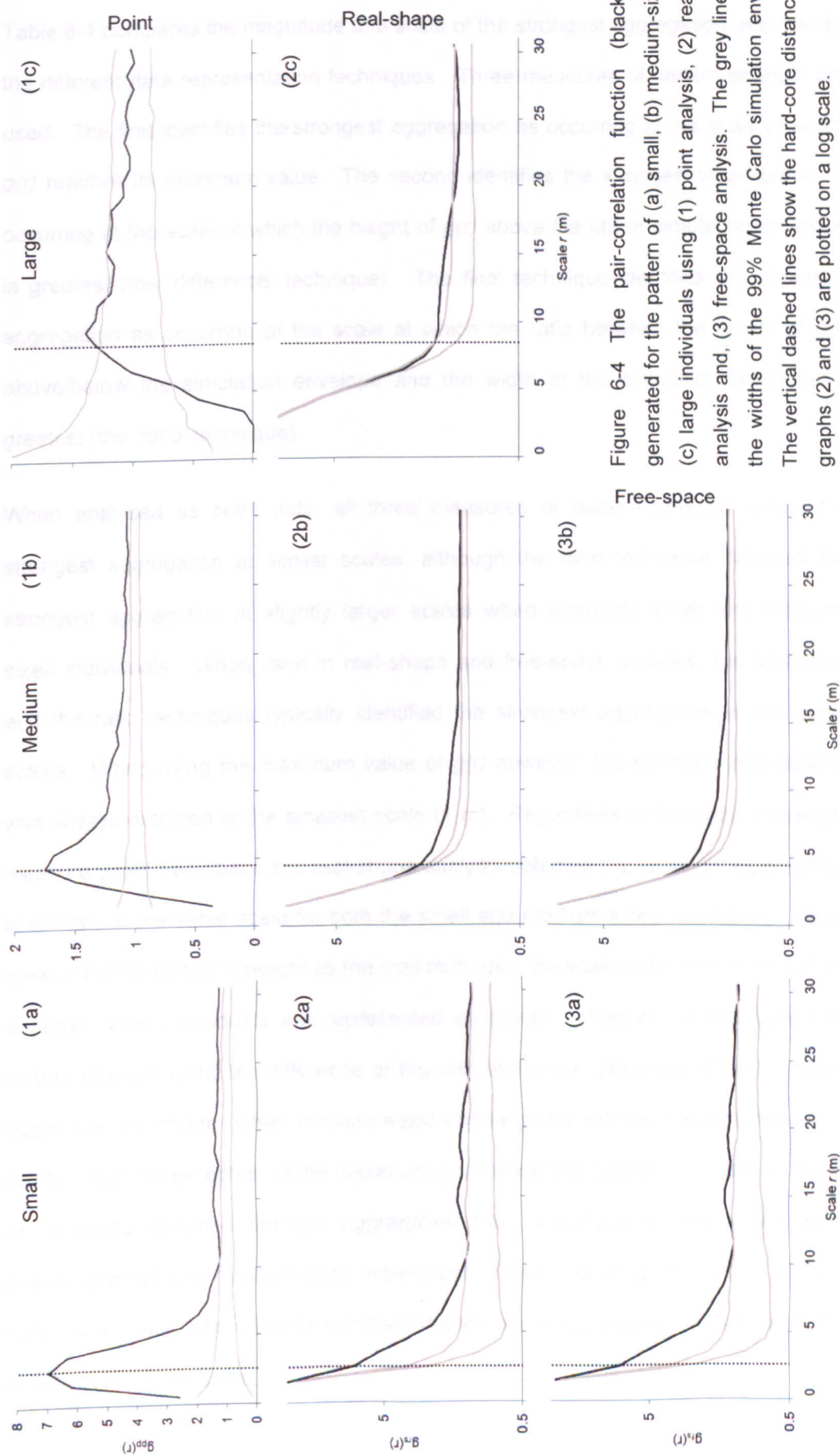
#### **Medium-sized individuals**

All three analyses identified aggregation at all scales greater than the hard-core distance of 4 m (Figure 6-4b). Slight differences in pattern were identified at scales below the hard-core distance (i.e. a soft-core effect). Point analysis detected dispersion at 1 m and a random distribution at 2 m. Real-shape analysis detected a random distribution at 1 m whereas free-space analysis detected dispersion. Both the real-shape and free-space analysis identified dispersion at 2 m. All three techniques identified aggregation at 4 m, but whereas the point and real-shape analysis identified aggregation at 3 m, free-space analysis detected a random distribution.

### **Large individuals**

Notable differences were observed between the real-shape and the point analysis at scales greater than the hard-core distance of 8 m (Figure 6-4c). Both analyses detect a random distribution between 19 and 29 m. However, whereas the real-shape analysis detected aggregation between 9 and 18 m, point analysis detected aggregation at 8 - 11 and 15 m. Disagreement also occurs at scales below the hard-core distance (i.e. a soft-core effect). Whereas the real-shape analysis detected a random distribution at all scales below the hard-core distance, point analysis detected dispersion between 1 and 5 m, a random distribution at 6 – 7 m and aggregation at 8 m.

### 6.3.2 THE EFFECT OF DATA REPRESENTATION ON THE SCALE OF THE STRONGEST AGGREGATION



### **6.3.2 THE EFFECT OF DATA REPRESENTATION ON THE MAGNITUDE AND SCALE OF THE STRONGEST AGGREGATION**

Table 6-1 compares the magnitude and scale of the strongest aggregation detected by the different data representation techniques. Three measures of pattern strength are used. The first identifies the strongest aggregation as occurring at the scale at which  $g(r)$  reaches its maximum value. The second identifies the strongest aggregation as occurring at the scale at which the height of  $g(r)$  above the upper simulation envelope is greatest (the 'difference' technique). The final technique identifies the strongest aggregation as occurring at the scale at which the ratio between the height of  $g(r)$  above/below the simulation envelope and the width of the simulation envelope is greatest (the 'ratio' technique).

When analysed as point data, all three measures of pattern strength detect the strongest aggregation at similar scales, although the ratio technique detected the strongest aggregation at slightly larger scales when analysing small and medium-sized individuals. When used in real-shape and free-space analysis, the difference and the ratio techniques typically identified the strongest aggregation at the same scales. When using the maximum value of  $g(r)$  however, the strongest aggregation was always detected at the smallest scale (1 m). Regardless of the pattern strength measure used, free-space and real-shape analysis detected the strongest aggregation at similar, or the same scale for both the small and medium-sized individuals. When measuring the pattern strength as the maximum  $g(r)$ , the scale of the dominant pattern is larger when individuals are represented as points. However, when measuring pattern strength using the difference or the ratio technique, the scale of the strongest aggregation is shorter when medium-sized and large individuals are represented as points. With the exception of the maximum  $g(r)$  measure of pattern strength, analyses using points identified stronger aggregation than real-shape or free-space when analysing small and medium-sized individuals. When analysing the pattern of large individuals, real-shape analysis identified the strongest aggregation regardless of the pattern strength measure.

Table 6-1 Comparing the magnitude and scale of the strongest aggregation detected for small, medium-sized and large individuals when using point, real-shape, and free-space analysis.

Size class	Point						Real-shape						Free-space						
	Max $g_{pp}(r)$	Scale	(Max $g_{pp}(r)$ )	Difference	Scale	(Difference)	Ratio	Scale	(ratio)	Max $g_{rs}(r)$	Scale	(Max $g_{rs}(r)$ )	Difference	Scale	(Difference)	Ratio	Scale	(ratio)	
Small	6.98	2	5.50	2	6.50	3		24.86	1	4.84	2	4.50	2	23.39	1	4.69	2	4.41	3
Medium	1.76	4	0.67	4	3.89	5		9.18	1	0.26	6	2.05	6	8.79	1	0.25	6	2.04	6
Large	1.39	9	0.12	9	0.24	9		17.49	1	0.13	13	0.46	13	-	-	-	-	-	-

Three measures of pattern strength are used. The first measures the strongest aggregation as occurring at the scale where the magnitude of  $g(r)$  is greatest ('Max  $g_{pp}(r)$ '). The second technique measures the strongest aggregation as occurring where the height of  $g(r)$  above the upper simulation envelope is greatest ('Difference'). The final technique measures the strongest aggregation as occurring where the ratio between the height of  $g(r)$  above the upper simulation envelope and the width of the simulation envelope is greatest ('Ratio'). For all measures the scale of dominant pattern is given in the subsequent column. Free-space analysis was not conducted for the large individuals.

### 6.3.3 THE EFFECT OF DATA REPRESENTATION ON THE INTERPRETATION OF ECOLOGICAL PROCESSES

When analysed as points the difference between the empirical  $g(r)$  indicated that aggregation was strongest among the small individuals (Figure 6-5c). The medium-sized individuals continue to show a weak aggregative signature at a slightly larger scale than the small individuals. The distribution of the large individuals is only slightly more aggregated than a completely spatially random distribution (i.e.,  $g(r) = 1$ ).

Interpretation of the free-space analysis is limited as it was only performed on the small and medium-sized individuals (Figure 6-5b). These analyses concur with the point analysis by detecting weaker aggregation among the medium-sized individuals than among the small individuals.

As with the point analysis, the real-shape analysis detected stronger aggregation among the small individuals than among medium-sized individuals (Figure 6-5a). However, unlike in the point analysis, the  $g_{rs}(r)$  of large individuals in the real-shape analysis is quite similar to the  $g_{rs}(r)$  of the small individuals, especially at scales less than  $r = 12$  m.



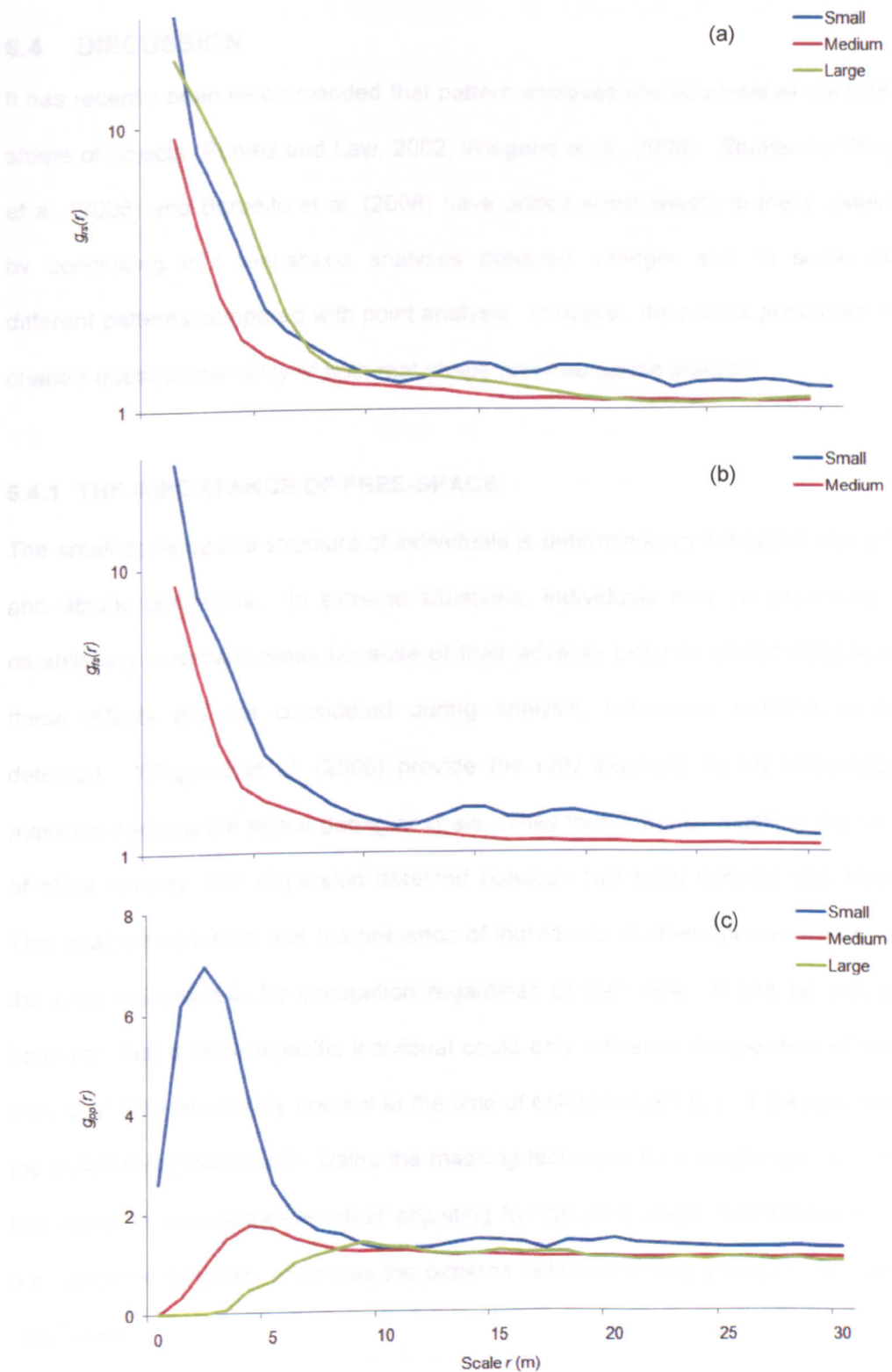


Figure 6-5 Comparing the pattern detected in each size class when using (a) real-shape, (b) free-space and, (c) point analysis. Free-space analysis was not conducted for large individuals. Note that graphs (a) and (b) are plotted on a log scale.

## **6.4 DISCUSSION**

It has recently been recommended that pattern analyses should preserve the size and shape of objects (Purves and Law, 2002; Wiegand et al., 2006). Studies by Wiegand et al. (2006) and Barbeito et al. (2008) have added some weight to these assertions by concluding that real-shape analyses detected stronger and, in some cases, different patterns compared with point analysis. However, the results presented in this chapter question the utility of both real-shape and free-space analysis.

### **6.4.1 THE IMPORTANCE OF FREE-SPACE**

The small-scale spatial structure of individuals is determined by biological interactions and abiotic conditions. In extreme situations, individuals may be prevented from establishing in certain areas because of their adverse biotic or abiotic conditions. If these effects are not considered during analysis, inaccurate patterns could be detected. Wiegand et al. (2006) provide the only example, to my knowledge, of masking inaccessible space during analysis. They found that by masking the location of other species, the dispersion detected between two focal species was stronger. This analysis assumed that the presence of individuals of other species would render the area inaccessible for occupation regardless of their age. It can be presumed, however, that a heterospecific individual could only influence the position of another individual if it was already present at the time of establishment (i.e., if it was older than the establishing individual). Using the masking technique for a single-species pattern, this research investigates whether adjusting for the area made inaccessible by older (i.e., larger) individuals influences the patterns detected among younger (i.e., smaller) individuals.

Compared with real-shape analysis, masking the space considered to be inaccessible had no notable effect on the scale-dependent patterns detected in either the small or medium-sized individuals (Figure 6-4). Furthermore, both real-shape and free-space analysis found the strongest aggregation to occur at the same scale (Table 6-1). The only notable difference was that free-space analysis detected slightly weaker

aggregation than real-shape analysis (Table 6-1, Figure 6-1f), although the difference was minimal. In the current system, therefore, it seems that controlling for the location of larger individuals has little effect on the spatial structures, and the inferred ecological processes, of small or medium-sized individuals. This is despite the fact that many of the large individuals are considerably greater in canopy area than the small individuals.

It is suggested that the *S. supranubius* population is so sparse that controlling for the effect of large individuals on the establishment of small does not affect the patterns detected at ecologically meaningful scales. The effect of older individuals' locations on the distribution of young individuals may be greater in denser communities. However, the justification and utility of the free-space analysis is questioned, especially when used to define biotically inaccessible locations. The most obvious theoretical restriction of this technique is the use of the contemporary distribution of (presumed) inaccessible space which, because of the senescence and growth of individuals, will be temporally dynamic. Free-space analysis may have greater utility when locations can be unequivocally determined as abiotically inaccessible. This too, however, would require much information on the abiotic requirements of the focal species, and is probably of limited utility given the plot extents commonly used in the literature (Figure 1-3).

#### **6.4.2 REAL-SHAPE VERSUS POINT APPROXIMATION**

##### **Interpreting real-shape analysis is difficult at small scales**

Approximating the size and shape of an individual as a group of adjacent points makes it harder to interpret the pair-correlation function at small scales. The distance between two objects is no longer measured as a single metric, but as a distribution of distance separating every point within one individual from every point in a neighbouring individual (Figure 6-3b). More importantly, however, the distribution of distances separating points within the same individual are also included in the calculation of the pair-correlation function. Consequently, there is a high frequency of

inter-point distances up to the scale of the canopy diameter which may mask any real interaction effects occurring in this range (Figure 6-4). This effect should be reduced in communities of non-overlapping individuals, as the spatial structure resulting from interactions should occur beyond the scale of the canopy diameter. It is suggested, however, that in denser communities, where interactions may occur at scales smaller than the canopy diameter, real-shape analysis may lose information at the scale of the individual and real interaction effects may be masked. In these situations it is recommended that the average canopy diameter is acknowledged and the strength of the pair-correlation function below this distance is investigated to ensure no interaction effects are being overlooked.

### **Detecting the strength of aggregation in real-shape analysis**

One technique of assessing pattern strength is to define the scale at which the pair-correlation function obtains its highest value as the scale at which the strongest spatial structure occurs (e.g., Getzin et al., 2008; Barbeito et al., 2009). This technique appears to be suitable when individuals are represented as points. However, because calculation of the pair-correlation function in real-shape analysis incorporates the distances separating the cells belonging to the same individual, the function is inflated at small scales (Figure 6-4). Therefore, measuring pattern strength using  $g_{max}(r)$  can result in an under-estimation of the scale of the strongest pattern (Table 6-1). When analysing individuals as objects composed of multiple points, evaluation of pattern strength should assess either the height of the empirical  $g(r)$  above the upper simulation envelope or the ratio between this measure and the width of the simulation envelope. The height of the empirical  $g(r)$  above the simulation envelope could be misleading as two patterns with (apparently) equal strength may have very different simulation envelope widths. Conversely, the ratio technique would enable the researcher to consider the likelihood of the pattern detected.

### **The difference between real-shape and point approximation depends on the size of the individuals**

As anticipated, the difference in the pattern detected by the real-shape and point analysis was greatest at scales less than the hard-core distance. At these scales point analysis detected a greater dispersion of both medium-sized and large individuals than the real shape analysis, as would be expected from a soft-core effect (Wiegand et al., 2006; Figure 6-1a,b; Figure 6-4, Section 6.1). Differences between real-shape and point analysis, at scales both below and above the hard-core distance, increased as the size of the individuals investigated increased. Both analyses detected the same spatial structure when analysing the distribution of small individuals (Figure 6-4 1a and 2a). This is easily explained as the small size of the individuals ( $1 - 2.8 \text{ m}^2$ ) and the small grid cell size ( $1 \text{ m}^2$ ) would mean that once converted to points, the real-shape pattern would essentially replicate the point pattern. When analysing the medium-sized individuals, the most notable difference between real-shape and point analysis was in the scale of the strongest pattern, which was estimated at a higher scale by real-shape analysis. Real-shape analysis also detected the strongest pattern of large individuals at a greater scale than point analysis. Therefore, when individuals are represented as a group of adjacent points, real-shape analysis identifies a greater modal plant–plant distance than point analysis. The following explanation is provided. Unlike point analysis, the distance between two objects in real-shape analysis is not a single, discrete value but a distribution of distances measured between multiple cells (Figure 6-3b). Despite this, however, the average distance should equal the distance separating the centroids of those objects and should therefore be equivalent to point analysis. However, because real-shape analysis calculates multiple distances for every pair of individuals, pairs that consist of one or more large individual (i.e., larger than average) will contribute more distances to the estimation of the function than pairs that consist of average or below-average sized individuals (Nuske et al., 2009). Competition theory predicts that large individuals will be separated from neighbouring individuals by greater distances (Getzin et al., 2006). Consequently, a few object pairs containing large individuals

separated by great distances may overpower the effect of a larger number of smaller, less widely spaced individuals. This weighting effect, therefore, may explain the increased modal distance observed when using real-shape analysis compared to point analysis and indicates that when analysing the spatial structure of a population of individuals of varying size, real-shape may overestimate the scale of the strongest pattern.

In addition to differences in the scale of the strongest pattern, real-shape analysis of large individuals detected aggregation over a larger spatial range than corresponding point analysis, agreeing with Nuske et al.'s (2009) analysis of simulated data. This may again be explained by the use of multiple distances between pairs of individuals, perhaps causing real-shape analysis to detect aggregation at scales ranging from the separation of near canopy edges to the separation of far canopy edges (Figure 6-3c).

Both Wiegand et al. (2006) and Barbeito et al. (2008) concluded that point approximation detected weaker effects than real-shape analysis, whereas Nuske et al.'s (2009) study of simulated data found that point approximation detected stronger effects than real-shape analysis. In the present research, the patterns identified by real-shape analysis were weaker than identified by point approximation when analysing the small and medium-sized individuals, but slightly stronger than point approximation when analysing the large individuals. This is, again, attributed to the use of multiple distances separating individuals which dampens the signature at the scale where distances are more frequent (Nuske et al., 2009). This seemingly size-dependent effect has not been acknowledged in previous studies.

### **Interpretative consequences of real-shape versus point analysis**

Implications for the interpretation of ecological processes arise from the analysis of medium-sized and large individuals where the differences between real-shape and point analysis were greatest. Most notably, the aggregation detected by real-shape analysis was more spatially extensive and the strongest pattern occurred at a larger scale than under corresponding point approximation. Both these effects are believed

to be consequences of the calculation of the real-shape pair-correlation function from multiple points within the same individual, leading to the overestimation of modal plant–plant spacing and overestimation of the importance of biological interactions in structuring a population.

Further assessment of the ecological consequences of data representation type was made by comparing the empirical functions between size classes. When analysed as points, the aggregation of individuals decreased as size increased. This is usually considered to be consistent with the operation of density-dependent thinning (Phillips and MacMahon, 1981; Lepš and Kindlmann, 1987). When analysed as real-shapes, there was a decrease in the pair-correlation function between small and medium-sized individuals, but the function for large individuals was similar to that of the small individuals. This could be considered as evidence for a shift towards facilitative interactions as a cohort ages (Bruno et al., 2003). However, the similarity between the functions for large and small individuals may also be evidence for the operation of clonal reproduction whereby clumps of clonally reproduced ramets are of a similar size and shape as large individuals and follow a similar distribution (Figure 6-6). Such an interpretation is only possible with *a priori* knowledge of the study system. However, although real-shape analysis may be capable of detecting patterns consistent with clonal reproduction, it does not give information about the relative spacing of ramets within clumps as point approximation does. When individuals are very closely distributed, real-shape analyses may lose information at the individual scale. Thus it seems that real shape analyses may be useful if there are specific hypotheses to be tested, and using both real-shape and point analysis in tandem may have merits when supported by knowledge of the study system.

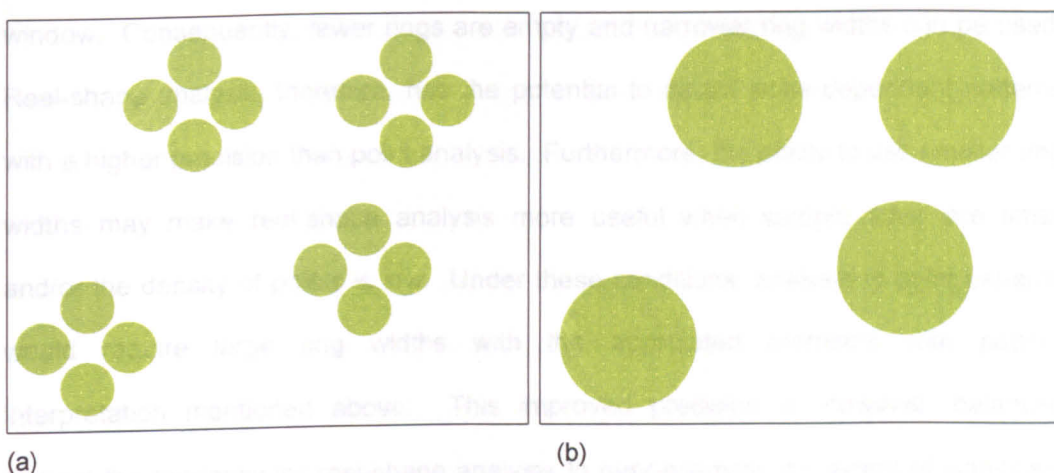


Figure 6-6 Diagram explaining the potential use of real-shape analysis to examine clonal reproduction. The clonal dynamics of *S. supranubius* may explain why the real-shape pair-correlation function for small individuals was similar to the function calculated for large individuals (Figure 6-5). Figure (a) shows the expected distribution of recently produced ramets whereas figure (b) shows the expected distribution of the maternal shrubs.

#### 6.4.3 THE POTENTIAL UTILITY OF REAL-SHAPE ANALYSIS

##### **Real-shape analysis may reduce error and improve the precision of scale-dependent pattern detection**

Real-shape analysis has several potential advantages. Firstly, it removes the potential for error associated with measurements of precise point locations, which Freeman and Ford (2002) demonstrated can affect the significance and scale of identified patterns. The errors associated with point approximation could be considerable when data are obtained from remote sensing. In these situations the best approximation of the rooting point is usually the centre of the canopy, although effects such as wind-throw, asymmetric canopies or the underlying topography will increase the error of such approximations.

When calculating the pair-correlation function the selection of ring widths is important as it determines the resolution with which patterns are detected. With point data there are often many empty rings, producing a jagged function, which can be difficult to interpret (Wiegand and Moloney, 2004; Illian et al., 2008). Increasing the ring width, however, reduces the ability of the pair-correlation function to isolate specific distance classes. Real-shape analysis increases the occupied proportion of the sample



window. Consequently, fewer rings are empty and narrower ring widths can be used. Real-shape analysis, therefore, has the potential to detect scale-dependent patterns with a higher precision than point analysis. Furthermore, the ability to use smaller ring widths may make real-shape analysis more useful when sample sizes are small and/or the density of points is low. Under these conditions, analysis of point patterns would require large ring widths with the associated problems with pattern interpretation mentioned above. This improved precision is, however, balanced against the tendency for real-shape analysis to over-estimate the extent of significant spatial structure and the scale of the strongest pattern, as discussed above.

### **Identifying the hard-core distance and inferring density-dependent competition**

Much of the criticism of point pattern analysis comes from the inability of distinguish small-scale dispersion from soft-core effects. However, identifying the presence and scale of soft-core effects could be useful. When applied in a sparse community, the scale of convergence of the pair-correlation function produced by real-shape and point analysis reveals the hard-core distance below which soft-core effects operate. Understanding how the scale of the strongest pattern relates to the hard-core distance and how this difference changes with the size of individuals could have important implications for determining the presence of density-dependent competition. An increase in both the hard-core distance and the scale of maximum aggregation with size would provide strong evidence for density-dependent competition. In other words, as shrubs increase in size (i.e. increase in the hard-core distance) they become separated by larger distances (i.e. increase in the scale of maximum aggregation) as previous neighbours are outcompeted. Furthermore, the distance between the hard-core distance and the scale of maximum aggregation should increase as competition increases in strength. This technique is investigated further in Chapter 7.

### **The application of real-shape analysis in studies of biomass**

Real-shape analyses may have important applications where the distribution of biomass, rather than rooting points, is of interest. For instance, Barbeito et al. (2008) used a combination of real-shape and point analyses to investigate the influence of tree crown pattern on the spatial pattern of seedlings and saplings. Maheu-Giroux and de Blois (2007) suggested real-shape techniques may be important when analysing clonal species whose population growth and spread is predominantly by the contiguous expansion of existing patches rather than the establishment of new patches or ramets, as in *S. supranubius*. Real-shape analyses could also be useful in studies of the distribution of gaps in biomass, such as forest-gap dynamics and patch-gap dynamics in arid systems.

## **6.5 CONCLUSIONS**

The utility of real-shape analyses that approximate the size and shape of individuals as multiple adjacent points is questioned. Most of the limitations of the technique can be attributed to the fact that the distance between two neighbouring objects is not a single value as in point analysis, but a distribution of distances separating multiple points. This is believed to have many effects, including overestimating the spatial extent of significant structure, overestimating the scale of the strongest pattern and underestimating the magnitude of the strongest pattern. These effects are exacerbated as the size of the individuals being analysed increases.

As expected, substantial differences between real-shape and point analysis occurred below the hard-core distance. These are attributed to soft-core effects making point analysis incorrectly detect dispersion at small scales. If only point data are available, caution should be practised when interpreting spatial structures at small scales. Ideally interpretation should be supported by consideration of the hard-core distance which could be estimated from field measurements of the maximum canopy diameter. However, although real-shape analyses do not incorrectly detect small-scale dispersion as readily, they can also be difficult to interpret at small scales. This is

especially so when neighbouring individuals are located close together. In these situations information at the scale of the individual may be masked by the calculation of multiple distances both within and between neighbouring objects.

Real-shape analysis does, however, have some benefits. In addition to allowing more precise assessment of scale-dependent pattern, the technique may be useful where patterns of biomass, rather than individuals, are of interest. It is suggested, however, that the benefits of real-shape analysis may be greatest when it is applied in tandem with point analysis, and supported by knowledge of the study system. Clearly, mapping the real shape of plants is a lot more time consuming than mapping point coordinates, which may explain the lack of interest in the real-shape approach until recently. However, with the increasing availability of remotely sensed data it should become increasingly feasible to quickly and accurately map the size and shape of tree and shrub crowns (see Chapter 4; Getzin et al., 2008).

The recommendations and comments in this chapter apply predominantly to analyses of populations of non-overlapping individuals.

## CHAPTER 7: PATTERN AND PROCESS IN *S. SUPRANUBIUS* AND THE EFFECT OF HETEROGENEITY

---

### 7.1 INTRODUCTION

One of the central aims of ecology is to investigate and understand the processes driving patterns of ecological phenomena (Levin, 1992; Tuda, 2007). Experimental techniques for studying plant population dynamics cannot generally be applied to long-lived species in low productivity systems, as direct measurements of processes are not feasible within realistic time-frames (see Section 1.3.2). Consequently, this chapter uses detailed spatial analyses to investigate the fine-scale characteristics of the distribution of *S. supranubius*, and to infer the operation of abiotic and biotic processes.

Despite much research, we have a poor understanding of the factors determining the spatial pattern of vegetation (Bestelmeyer et al., 2006). Process inference can be hampered by the presence of spatial environmental heterogeneity. Almost all natural environments are patchy (Hewitt et al., 2007). Topography, microclimate and resource availability all vary in space, producing a mosaic of habitat quality. Many studies have demonstrated the preferential location of individuals within, for example, certain habitats (Pueyo and Alados, 2007) or topographic regimes (Klausmeier, 1999). Thus spatial environmental heterogeneity can affect the broad-scale distribution of a species, known as first-order effects (see Table 3-1). Cases where the spatial pattern of individuals can be unequivocally shown to be a consequence of biological interactions between individuals alone are uncommon (Rohani et al., 1997; Perfecto and Vandermeer, 2008; Rietkerk and van de Koppel, 2008). The pattern induced by exogenous abiotic controls can be mistaken for spatial correlation that is due to demographic processes (Wagner and Fortin, 2005), or may mask true demographic effects. Consequently, much work in contemporary ecology attempts to separate the abiotic and biotic controls on species' patterns and dynamics. To achieve this, many studies are performed in what are presumed to be environmentally homogeneous

areas (Wagner and Fortin, 2005; Perfecto and Vandermeer, 2008) to try to remove the effect of environmental heterogeneity. These, and the majority of other studies, including simulation studies of species self-organisation (e.g., Barbier et al., 2008; Pueyo et al., 2008), are based on the assumption that the second-order spatial structure of a species (which is believed to result from ecological interactions between individuals) occurs independently of environmental heterogeneity. This is despite the fact that almost all natural communities are embedded in a landscape of numerous environmental heterogeneities, both spatial and temporal (Hewitt et al., 2007; see Chapter 5).

While theoretical studies have considered how spatial environmental heterogeneity can influence co-existence mechanisms (e.g., Chesson, 2000a; Amarasekare, 2003), few studies, either theoretical or analytical, have considered how spatial environmental heterogeneity can affect the secondary spatial structure of individual species. Those studies that have, used artificial research designs (Hartgerink and Bazzaz, 1984; Neatrour et al., 2007; Roiloa and Returto, 2006, 2007; Roiloa et al., 2007) and therefore cannot test the importance of natural heterogeneity, especially at large scales. Therefore, there is surprisingly little theory or understanding of how natural spatial environmental heterogeneity may interact with biological processes to determine population dynamics (Wagner and Fortin, 2005; Murrell, 2009). One exception is a study by Getzin et al. (2008), which demonstrated that the biological processes operating in Douglas fir (*Pseudotsuga menziesii*) forests interact with spatial environmental heterogeneity to produce qualitatively and quantitatively different population spatial structures in different areas. The interaction of abiotic and biotic processes operating at different scales poses a formidable challenge to ecological researchers (Shimatani and Kubota, 2004; Wagner and Fortin, 2005). Thus, in addition to investigating the spatial structure of *S. supranubius*, this chapter also considers whether, and how, spatial environmental heterogeneity interacts with biological processes to determine the spatial structure of *S. supranubius*.

### ***A priori* predicted versus observed spatial patterns**

When investigating spatial point patterns most ecological studies apply Ripley's  $K(r)$  or the pair-correlation function (or their variants) to mapped distributions of points representing individuals of the species of interest. Despite the ubiquity of articles using these techniques (Figure 1-2), numerous studies have asserted that the analysis of pattern alone is not enough to infer underlying processes (Mahdi and Law, 1987; Borcard et al., 2004; Schurr et al., 2004). However, it is widely accepted that non-random processes frequently result in highly structured, distinctive patterns (McIntire and Fajardo, 2009). Biological organisation exists and, although the link between pattern and process may be imperfectly understood, patterns of ecological phenomena continue to provide important opportunities for enhancing our understanding of population dynamics and structure. To account for the inferential gap between pattern and process, increasing emphasis is being placed on deductive reasoning rather than inductive description of pattern (Schurr et al., 2004; Fajardo et al., 2008; McIntire and Fajardo, 2009). Instead of attempting to assign processes to observed patterns, deductive reasoning uses ecological theory and knowledge of the focal system and species to formulate precise *a priori* hypotheses of the likely abiotic and biotic processes of importance and their expected spatial signatures. Support for these hypotheses can then be tested by analysing the observed pattern of individuals and assessing how closely they fit the predictions.

## 7.2 AIMS AND OBJECTIVES

This chapter addresses two main hypotheses:

*Hypothesis 1:* The five focal substrates will exhibit differing levels of large-scale spatial environmental heterogeneity.

*Hypothesis 2:* Large-scale environmental heterogeneity will not influence the spatial structure of, and (by inference) the biological processes structuring, the *S. supranubius* populations.

*A priori* hypotheses of the likely biological processes are formulated and their support tested using detailed spatial analysis of the observed patterns of *S. supranubius* individuals. A summary of the hypothesised processes and their associated spatial predictions is provided in Table 7-1.

Table 7-1 Hypothesised processes influencing the spatial structure of *S. supranubius* and the related spatial pattern predictions.

Hypothesised process		Spatial expectations	
Biological processes	Vegetative reproduction by branch layering.	Increased frequency of inter-shrub distances between 0 and 10 m among small individuals causing aggregation at small scales.	Hypothesised Process 1
	Intra-specific competition.	Reduced strength of aggregation as cohort ages.	Hypothesised Process 2
Biotic-abiotic interaction	There will be no difference in the presence of clonal reproduction between substrates, regardless of any heterogeneity.	Increased frequency of inter-shrub distances between 0 and 10 m among small individuals causing aggregation at small scales on all substrates. Modal inter-shrub distances (among small individuals) show little difference between substrates.	Hypothesised Process 3
Large-scale heterogeneity	The signature of environmental heterogeneity will vary between substrates.	Under environmental heterogeneity the pattern of largest individuals ( $\geq 30\text{ m}^2$ ) will deviate from complete spatial randomness (CSR) generating aggregation at large scales ( $r > 20\text{ m}$ ). Stronger heterogeneity will result in more extensive and larger deviations.	Hypothesised Process 4



### **Hypothesised Process 1: vegetation reproduction by branch layering**

*S. supranubius* is capable of both sexual and vegetative reproduction (Kyncl et al., 2006). Many studies have suggested that vegetative spread is more important than sexual reproduction for maintaining population growth rates in clonal species (Mandujano et al., 2001; Clark-Tapia et al., 2005; Mandujano et al., 2007). Vegetative spread by branch layering is therefore expected to be an important process driving the spatial structure of *S. supranubius*. The production of clonal offspring via branch layering primarily occurs when individuals attain large sizes (McAuliffe et al., 2007). Field observations indicate that *S. supranubius* individuals typically reproduce clonally once individuals reach a diameter of c. 10 m. The rooting of lateral branches produces independent ramets around the periphery of the senescing maternal shrub. This is expected to produce small-scale aggregations of young individuals at scales of less than 10 m.

### **Hypothesised Process 2: intra-specific competition**

Limited water availability invokes strong competitive interactions between arid plants (Briones et al., 1998; Gebauer et al., 2002). Density-dependent thinning should cause cohorts of shrubs to become increasingly dispersed over time (Metsaranta and Lieffers, 2008). Field observations suggest that interconnections between ramets are lost shortly after ramet establishment and senescence of the maternal shrub. Thus, it is expected that density-dependent thinning will influence the distribution of *S. supranubius* individuals shortly after ramet establishment. This process should be evident in the decreasing strength of aggregation with cohort age (i.e., a decrease in the magnitude of the strongest pattern as cohorts age).

### **Hypothesised Process 3: clonal reproduction will occur regardless of environmental heterogeneity**

Landscape-scale heterogeneity, perhaps associated with broad habitat types, has been shown to influence the dynamics of clonal plants, often by driving spatial variation in reproductive dynamics (Mandujano et al., 2001, 2007). However, all focal *S. supranubius* populations come from the same habitat type and it is therefore

predicted that *S. supranubius* will portray a spatially consistent signature of clonal growth, regardless of the degree of environmental heterogeneity present. Thus, where environmentally heterogeneity is present, individuals will be clustered at large scales but within these patches, the spatial pattern of young individuals should follow the signature of clonal reproduction (see Hypothesis 1). This should be evident in similar scales of aggregation of young individuals on all five substrates once the effects of large-scale heterogeneity have been removed from the species distribution.

**Hypothesised Process 4: the magnitude of spatial environmental heterogeneity will vary between substrates**

The five substrates considered are of different age and vary considerably in their surface geomorphological characteristics (pers. obs.; Kyncl et al., 2006; see Table 2-1). Assuming that the geomorphology of the substrates affects the broad-scale distribution of *S. supranubius* it is hypothesised that the substrates will display differing levels of biologically relevant environmental heterogeneity.

## **7.3 METHODS**

### **7.3.1 DATA COLLECTION**

The canopy areas and locations of 17,551 *S. supranubius* individuals were mapped over 162 hectares on five substrates, using spectral one-class classification (Chapter 4). The locations of *S. supranubius* individuals were represented by points located at the centre of the canopies, defined by co-ordinates ( $x, y$ ) (see Section 4.4). Sample windows on each substrate were rectangular. However, because of variation in the shape and extent of the five substrates, the sample window size had to vary between substrates (see Section 4.4). For each substrate the data were divided into two equally sized, spatially contiguous datasets. One dataset from each substrate was randomly selected for analysis in the present chapter (Table 7-2). The remaining data have been retained for the validation of a simulation model that is currently in development.

Table 7-2 Dimensions of the sample windows used in Chapter 7. The dimensions of the sample windows are given in Universal Transverse Mercator (UTM). ULX/ULY and LRX/LRY provide the x and y coordinates of the upper left, and lower right corners of the sample window respectively (metres).

Substrate	ULX / ULY	LRX / LRY	Area (ha)
1*	340150 / 3124050	340700 / 3123700	19.25
2	340400 / 3125100	340750 / 3124400	24.50
3	341710 / 3126780	342110 / 3126455	13.00
4	342600 / 3123950	343300 / 3123600	24.50
5	342785 / 3124995	342985 / 3124745	5.00

\* Substrate 1a in Table 4-6.

### 7.3.2 ANALYSES

All pattern analyses were performed in Programita (Wiegand and Moloney, 2004).

#### Analysis 1: environmental heterogeneity

This analysis determined the level of spatial environmental heterogeneity present on each of the five substrates (Hypothesis 1; Hypothesised Process 4). Separating the true effects of environmental heterogeneity from the effects of biological processes is not a clear-cut task, and is the subject of current statistical research (Law et al., 2009).

Heterogeneity is defined as spatially structured variability in a property of interest (Wagner and Fortin, 2005). Because of differences in resource requirements and life history attributes, species will differ in their response to the spatial distribution of particular resources. Thus, measures of heterogeneity should capture the distribution of resources important to the species of interest. To ensure this, *S. supranubius* individuals were used as a biological indicator of habitat suitability and thus environmental heterogeneity (following Stoyan and Penttinen, 2000; Getzin et al., 2008, Barbeito et al., 2009; Zhu et al., 2010). It is assumed that very large *S. supranubius* individuals are either very old, or have a very high growth rate. Either way, the presence of very large individuals implies the local habitat quality is superior relative to surrounding locations. Thus, the location of the largest individuals is

assumed to provide a reasonable proxy for biologically relevant environmental heterogeneity.

The small-scale structure of plant communities is typically attributed to plant–plant interactions, whereas deviation from spatial randomness at larger scales is attributed to environmental heterogeneity (Stoyan and Penttinen, 2000; Wiegand et al., 2007b). Environmental heterogeneity should cause broad-scale patches of elevated plant density. This should be represented by deviations of  $g(r)$  and  $L(r)$  from the CSR expectation (i.e.,  $g(r) = 1$ ,  $L(r) = 0$ ) at scales exceeding the distance at which shrub–shrub interactions are believed to be important (Wiegand and Moloney, 2004). Biotic interactions between *S. supranubius* individuals are not expected to extend beyond c. 22 m (see page 188). Because  $L(r)$  is a cumulative function small-scale spatial structure can influence the function values at larger scales. This may cause a steady increase in  $L(r)$  as scale increases independent of any heterogeneity effects. Therefore,  $L(r)$  was only considered to represent heterogeneity if statistically notable pattern (i.e., location of the empirical function outside the simulation envelopes) followed an increase in the gradient of the function. To ensure any heterogeneity indicated by the pattern analyses is genuine, it is important to support the analyses with knowledge of the heterogeneity processes that may be important (Law et al., 2009).

On each substrate the spatial pattern of the largest individuals ( $\geq 30 \text{ m}^2$ ) was compared to the null model of CSR using both  $g(r)$  and  $L(r)$ . This excluded all but the largest mature adults and thus should be a good indicator of environmentally driven habitat quality. Sample sizes for this analysis were 431, 131, 47, 141 and 15 on Substrates 1 to 5 respectively. A ring width of 5 m was used in analyses using  $g(r)$  as it produced relatively smooth functions.

## **Analysis 2: size–abundance distribution**

Several methods were used to describe the size–abundance distribution of *S. supranubius* (Hypothesis 2). For each substrate the density of individuals, the average canopy area and the standard deviation in canopy area was calculated. Size

inequality within substrates was characterised by the coefficient of variation (CV) in canopy area (Coomes and Allen, 2007). Larger values of CV indicate greater size inequality.

Mathematical distribution functions can be used to model the size–abundance distribution of a population (i.e., the number of individuals that fall within each size class). Over recent years there has been an increasing emphasis on using size–abundance distributions to help understand underlying demographic processes (Wang et al., 2009). Under demographic equilibrium, the size–abundance distribution of plants can be understood as the consequence of size-dependent variation in growth and mortality (Coomes et al., 2003). Both mortality and growth reduce the number of individuals in a size class; mortality decreases the number of individuals in the present and subsequent size classes, whereas growth decreases the number of individuals in the present size class and moves individuals into the next larger size class (Muller-Landau et al., 2006; Figure 7-1a). Depending on how growth and mortality scale with size, different size–abundance distributions will be created. Muller-Landau et al. (2006) used relationships between size, growth and mortality to derive three analytical predictions of size–abundance distributions: power law, exponential and Weibull functions. Using the predictions of Muller-Landau et al. (2006) and knowledge of the focal species the size-abundance distribution of *S. supranubius* could be predicted.

In addition to the effects of mortality and growth, clonal reproduction is expected to have a strong influence on the *S. supranubius* size-abundance distribution (Hypothesised Process 1). Because of clonal reproduction, large individuals are not only lost from the system through mortality (as in Muller-Landau et al.'s models) but may re-enter the model as multiple small individuals (Figure 7-1b). New ramets are expected to be clustered and consequently, competitive interactions among the smallest individuals are expected to be intense. Therefore, mortality in the smallest size-classes is predicted to be a function of individual size, decreasing as individuals increase in size and become more dispersed (following the hypothesis of density-dependent thinning). These processes are predicted to generate a steep size–

abundance distribution among small *S. supranubius* individuals, which should be most closely fit by a scaling function (Coomes et al., 2003) or the Weibull distribution (Muller-Landau et al., 2006). In the intermediate size-classes competition is expected to be less intense and exogenous disturbance may be a major source of mortality (Coomes et al., 2003). The mortality rate should therefore be constant across the intermediate size classes, corresponding with a Weibull distribution (Muller-Landau et al., 2006). However, in the largest size classes clonal reproduction is again predicted to have a strong influence on mortality dynamics. In many demographic models, the mortality rate is often expected to remain constant in the largest size classes. Among *S. supranubius* individuals, however, the mortality rate is expected to increase with size as large individuals collapse and produce clonal ramets. Following the predicted effects of demographic processes on size-abundance distributions described Coomes et al. (2003) and Muller-Landau et al. (2006) it is hypothesised that the size–abundance distribution of *S. supranubius* should be most accurately described by either a scaling function or the Weibull distribution.

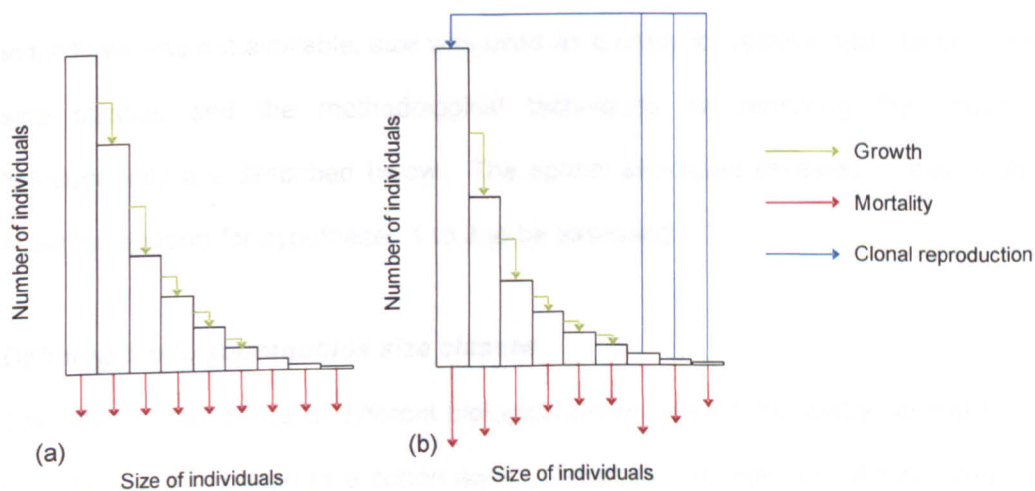


Figure 7-1 Size abundance distributions as conceptualised by (a) Muller-Landau et al. (2006) and, (b) and with the addition of clonal reproduction. Clonal reproduction is expected to produce a steeper size–abundance relationship than when only growth and mortality determine its shape. Furthermore, whereas some demographic models predict mortality rate to be constant with size (e.g., the power-law and exponential distribution [Muller-Landau et al., 2006, a]), mortality rate is expected to scale with size among the smallest and largest individuals (b) (see text for explanation).

Exponential, power, two-parameter Weibull, and lognormal functions were fit to the observed *S. supranubius* size–abundance data using maximum likelihood methods. The exponential, Weibull and lognormal functions were fit using the *fit.distr* tool in the *MASS* package of R (Ripley, 2009). The power law distribution was fit using the *pareto.fit* tool. Akaike's Information Criterion (AIC) (Akaike, 1974) was used to compare the fit of the different functions.

### **Analysis 3: the spatial structure of *S. supranubius***

This analysis aimed to compare the spatial structure of individuals between substrates, without the confounding effects of environmental heterogeneity. This comparison would reveal whether there were any differences in the structure of *S. supranubius* individuals on different substrates that could be attributed to environmental heterogeneity (Hypothesis 2). The biological processes driving the dynamics of a cohort of individuals are likely to change with the age of the cohort. It is therefore common to analyse the spatial structure of individuals of different age (e.g., Barbeito et al., 2009; Zhang et al., 2009). As information on the age of *S. supranubius* individuals was not available, size was used as a proxy for relative age. Details of the size classes and the methodological techniques for removing the effects of heterogeneity are described below. The spatial structures revealed in this analysis allow the support for hypotheses 1 to 3 to be assessed.

### **Defining the *S. supranubius* size classes**

The relative importance of different biological processes for the spatial structure of *S. supranubius* will change as a cohort ages. Therefore, the definition of size class must consider the processes being investigated. There is, however, a theoretical–statistical trade-off when defining size classes. On the one side it is important not to dilute spatial signatures by including individuals that are too large or too small for the predicted process. For instance, the signature of clonal reproduction is expected to be strongest among the youngest individuals before processes such as competitive thinning modify the pattern. On the other hand, narrow size classes may produce small sample sizes, potentially affecting the reliable identification of spatial pattern

(Chapter 5). There is, however, no formal definition of a minimum sample size in spatial pattern analyses. Published recommendations range from 15 individuals (Rossi et al., 2009) to 70 individuals (Wiegand et al., 2007a).

Both clonal reproduction by branch layering, and intra-specific competition are expected to be major processes driving the spatial structure of *S. supranubius*. The spatial signature of clonal reproduction should be most pronounced among recently produced ramets (Hypothesised Process 1). Kyncl et al. (2006) assumed that recent *S. supranubius* clonal offspring would have canopy diameters of less than 1 m (i.e., 0.79 m<sup>2</sup> canopy area). However, as individuals with a canopy area less than 1 m<sup>2</sup> were excluded from the dataset (see Section 4.4), replicating Kyncl et al.'s (2006) size class definition was not possible. Conversely, because of the slow dynamics of arid shrubs, the univariate signature of intra-specific competition may not be evident until cohorts are very old (Hypothesised Process 2). Therefore, the investigation of these processes requires size classes that allow the investigation of individuals that are reliably very young and very old.

For subsequent analysis, *S. supranubius* individuals were divided into three canopy-size classes representing small, medium-sized and large individuals. For various reasons, size may be an inadequate representation of age. However, because of the slow growth of *S. supranubius*, it is assumed that a broad division based upon canopy area provides a suitable correlate of age. In general the log-normal distribution was found to fit the size–abundance distribution of *S. supranubius* most accurately (see Section 7.4.1). The canopy-area data were logged to give normally distributed data. Medium-sized shrubs were defined as those with (log)canopy area within one standard deviation of the mean. Individuals with smaller canopy-areas were defined as small, whereas individuals with larger canopy-areas were defined as large. This corresponded to canopy-area ( $x$ ) ranges of  $1 \leq x < 2.8 \text{ m}^2$ ,  $2.8 \leq x < 20.6 \text{ m}^2$  and  $20.6 \leq x \text{ m}^2$  for small, medium-sized and large individuals respectively. Table 7-3 shows the number of individuals within each size class for each sample window. The definition of the large size class produced relatively small sample sizes on Substrates



3 and 5 (Table 7-3). It was decided, however, that reducing the lower size limit to increase sample sizes would risk diluting any spatial signature of intra-specific competition. For the purpose of describing large-scale environmental heterogeneity an extra size class was defined containing all individuals with a canopy area  $x \geq 30 \text{ m}^2$ . This class excluded all but the largest mature adults and thus should be a good indicator of environmentally driven habitat quality. Small samples sizes, however, were produced on Substrates 3 and 5 (Table 7-3).

Table 7-3 The number of *S. supranubius* individuals within each size class in each of the five sample windows used in Chapter 7. The number of individuals used in the analysis of heterogeneity (Het.) is also given.

Substrate	Number of individuals per size class			Het.
	Small	Medium	Large	
1	342	988	619	431
2	432	1711	329	131
3	224	423	83	47
4	345	2333	382	141
5	219	264	43	15

***Removing the effect of heterogeneity: the inhomogeneous  $g(r)$***

The inhomogeneous  $g(r)$  compares the distribution of individuals to a heterogeneous Poisson null model. The heterogeneous Poisson process displaces the original location of shrubs in accordance with a user-defined intensity function  $\lambda(x,y)$  that represents first-order effects (i.e., environmental heterogeneity). This destroys the small-scale spatial structure (driven by biological interactions) whilst maintaining the large-scale pattern (driven by environmental heterogeneity). The intensity function was constructed from the distribution of the largest individuals ( $\geq 30 \text{ m}^2$ ) using a circular moving window of radius  $h$ . The intensity  $\lambda(x,y)$  is weighted by an edge-corrected Epanečnikov kernel (Stoyan and Stoyan, 1994; Wiegand et al., 2007a). This technique produces a spatially explicit intensity function that is notably smoother than more traditional moving window approaches. The Epanečnikov kernel is defined as:

$$e_h(d) = \begin{cases} \frac{3}{4h} \left( 1 - \frac{d^2}{h^2} \right) \\ 0 \end{cases}$$

if  $-h \leq d \leq h$ , and zero otherwise, where  $d$  is the distance from the focal point and  $h$  is the bandwidth (Cousens et al., 2008; Zhu et al., 2010).

The above approach is based on the assumption of the separation of scales (Wiegand et al., 2007a); that is, environmental heterogeneity will influence the distribution of shrubs at large scales, whereas shrub–shrubs interactions will take place at small scales. Therefore, to account for environmental heterogeneity, the kernel bandwidth ( $h$ ) should be greater than the scale of second-order effects, but smaller than the scale of environmental controls (Thorsten Wiegand, personal communication). In northern forests second-order effects are widely believed to extend to a maximum of 15 m (Stephan Getzin, personal communication). However, no comparable studies are available for arid shrubs. It is generally accepted that ecological interactions in arid systems are primarily conducted below ground for water resources (Noy-Meir, 1973) via laterally extended root systems, which can stretch well beyond shrub canopy limits (Hartle et al., 2006). Without excavation, which is not permitted in the study site or for the study species, scales of below-ground interaction can only be estimated from above-ground morphology. Although some studies have attempted to map the below-ground morphology of arid shrubs, and in some cases its relationship to canopy morphology (e.g., Kummerow et al., 1977; Palacio and Montserrat-Marti, 2007), results are species specific and based on the excavation of relatively few individuals, making it difficult to generalise to equivalent species. Barbier et al. (2008) estimated that arid shrub root systems extended horizontally beyond the canopy by a minimum 25% of the canopy radius. Assuming isotropic root distributions and canopy circularity, *S. supranubius* individuals in the study system have a maximum canopy radius of c. 9 m. Using the above metric, this corresponds to a maximum horizontal root extent of 11 m, and thus a maximum interaction scale of 22 m. Conversely, Caldwell et al. (2008) found that the influence of *Larrea tridentata* and *Lycium pallidum* Miers. (Solanaceae) canopies on the structure of soils and their hydraulic properties

extended to 1.4 times the canopy radius. In the present system, this would represent a maximum scale of interaction of c. 25 m. Consequently, a conservative value of  $h = 30$  m was selected for the Epanečnikov kernel. This incorporated both Barbier et al.'s and Caldwell et al.'s metrics, but also allowed for slightly longer-distance interactions to occur.

To control for the effects of heterogeneity, the second-order structure of *S. supranubius* was analysed using the homogeneous  $g(r)$  and the inhomogeneous  $g(r)$  on environmentally homogeneous and heterogeneous substrates respectively. The  $g(r)$  for each size class on each of the substrates was compared. If analyses revealed similar spatial structures on the homogeneous and heterogeneous substrates it could be concluded that no interaction occurred between environmental heterogeneity and *S. supranubius* demographic processes (e.g., Hypothesised Process 3). Conversely, strong differences in the patterns detected in homogeneous and heterogeneous substrates would provide evidence consistent with feedback effects between heterogeneity and demographic processes (Hypothesis 2).

Because of differences in data density, different ring widths ( $\delta r$ ) were used to construct  $g(r)$  in each size class. Small ring widths can produce noisy function estimates, often producing spurious and meaningless spikes, especially when sample size is small (Wiegand and Moloney, 2004; Illian et al., 2008). However, large ring widths lose fine-scale information (see Figure 3-2e and f). Analysis of the small, medium and large size classes used ring widths of 3, 2 and 4 m respectively, as these produced relatively smooth  $g(r)$ -functions.

When individuals in a population do not overlap, approximating their location as dimensionless points can lead to the  $g(r)$  incorrectly detecting dispersion at small scales (Wiegand et al., 2006; see Section 6.1). To estimate the hard-core distance below which  $g(r)$  may be influenced by the shape and size of individuals (i.e. soft-core effects), the analyses were performed once with the data estimated as points and again with the size and shape of the shrubs maintained (i.e., real-shape analysis; see Section 6.2.5). The scale at which these two functions have equivalent  $g(r)$  values

estimates the maximum canopy diameter (see Section 6.2.2). The convergence of the point and real-shape  $g(r)$  is interpreted as the hard-core distance below which soft-core effects may affect the pattern detected by  $g(r)$ . For each analysis the maximum value of  $g(r)$ , the scale at which it occurred, and the hard-core distance were recorded. Analysis of spatial patterns beyond the hard-core distance used the results of the point pattern analysis.

### ***Comparing pattern strength***

Studies of population patterns commonly use simulation envelopes generated from Monte Carlo simulations to generate qualitative interpretations of pattern aggregation, randomness or dispersion (see Section 1.3.4). Such interpretations can prevent meaningful comparisons being made between different populations or scenarios (Fajardo et al., 2006); especially as the width of the simulation envelope is closely related to the number of individuals mapped (see Sections 5.3.4 and 5.4.3). As these analyses used point patterns and the null model of complete spatial randomness (CSR) (expected values of  $g(r) = 1$  and  $L(r) = 0$  under CSR), differences in pattern strength between size classes and substrates were assessed by directly comparing the empirical  $g(r)$  curves (following Getzin et al., 2008; Barbeito et al., 2009; Meador et al., 2009; see Section 6.4.2 page 166).

## 7.4 RESULTS

### 7.4.1 ENVIRONMENTAL HETEROGENEITY

To determine the presence of large-scale spatial environmental heterogeneity, the pattern of the largest individuals ( $\geq 30 \text{ m}^2$ ) on each substrate was compared with the null model of CSR (Figure 7-2). A strong increase in  $L(r)$  at large scales ( $> 20 \text{ m}$ ) was observed on Substrate 4. Substrate 2 showed a weak increase in  $L(r)$  at large scales but as there was no obvious change in the gradient of the curve it was not considered to represent heterogeneity. The  $L(r)$  on Substrate 1 did not indicate the presence of heterogeneity. Analysis by  $g(r)$  largely confirmed the results of  $L(r)$ , indicating homogeneity on Substrates 1 and 2, but a deviation from homogeneity at large scales on Substrate 4.

The sample sizes used on Substrates 3 and 5 are considered to be small (i.e.,  $< 70$ ; Wiegand et al., 2007a). The results of both the  $g(r)$  and  $L(r)$  may therefore be unreliable, especially on Substrate 5 where only 15 individuals were used. At such small sample sizes the simulation envelopes can become wide, increasing the risk of Type II error. As such, interpretation focuses only on the main patterning trends revealed by the functions, and whether these indicate any underlying heterogeneity (i.e., the 'significance' of the pattern is not considered as important as the presence of strong increases or decreases in the function values). Analysis by  $g(r)$  on Substrate 3 does not exceed the simulation envelopes at any scale. The function does, however, increase sharply after 35 m and approaches significance at the largest scales. The  $L(r)$  function on Substrate 3 oscillates around the upper simulation envelope at all scales. These results suggest that the substrate is environmentally heterogeneous. Observations in the field revealed strong ridge–trough topography on this substrate which adds support to the apparent presence of heterogeneity. The  $L(r)$  function on Substrate 5 does not indicate the presence of heterogeneity, but the  $g(r)$  function shows considerable fluctuations. However, because of the small sample size (exacerbated by the non cumulative nature of the  $g(r)$  function) fluctuations in  $g(r)$  were expected. Furthermore, field observations did not reveal any obvious

heterogeneity forces. Analysis of the same data as real-shapes using  $g(r)$  detected dispersion ( $g(r) < 1$ ) of the largest individuals between 21 and 30 m (see Appendix E). However, the real-shape  $g(r)$  remained within the simulation envelope at all scales. Therefore, along with Substrates 1 and 2, Substrate 5 was considered to be homogeneous in subsequent analyses. Conversely, both Substrates 3 and 4 appear to be environmentally heterogeneous (supporting Hypothesised Process 4). Knowledge of the different substrates supports this conclusion as both Substrates 3 and 4 are pahoehoe flows and thus have pronounced ridge–trough topography. Substrate 5 is treated as homogeneous in the following analyses, but it is acknowledged that some heterogeneity may be present.

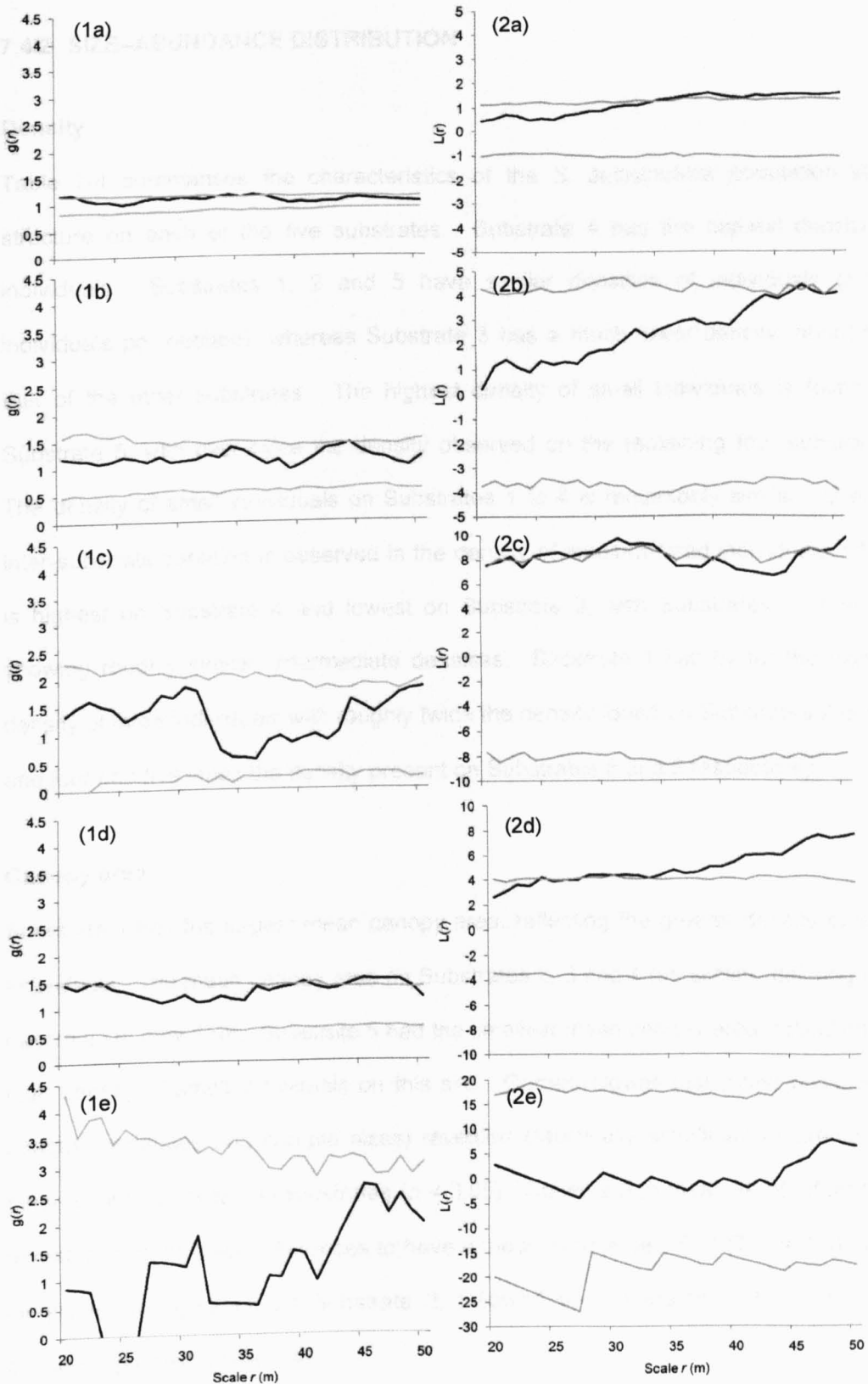


Figure 7-2 Assessing the presence of environmental heterogeneity on the five substrates. The pattern of large individuals ( $\geq 30 \text{ m}^2$ , black lines) is analysed using the pair-correlation function ( $g(r)$ ) (1) and the  $L(r)$ -function (2) on Substrates 1 to 5 (graphs (a) to (e) respectively). Approximately 99% simulation envelopes generated from the 5<sup>th</sup>-highest and 5<sup>th</sup>-lowest value of 999 simulations of the CSR null model (grey lines). Note the changes in the y-axis scales.

#### 7.4.2 SIZE-ABUNDANCE DISTRIBUTION

##### Density

Table 7-4 summarises the characteristics of the *S. supranubius* population size-structure on each of the five substrates. Substrate 4 has the highest density of individuals. Substrates 1, 2 and 5 have similar densities of individuals (c.100 individuals per hectare), whereas Substrate 3 has a much lower density, about half that of the other substrates. The highest density of small individuals is found on Substrate 5, with over twice the density observed on the remaining four substrates. The density of small individuals on Substrates 1 to 4 is remarkably similar. Greater inter-substrate variation is observed in the density of medium-sized individuals, which is highest on Substrate 4 and lowest on Substrate 3, with Substrates 1, 2 and 5 showing roughly similar, intermediate densities. Substrate 1 has by far the highest density of large individuals with roughly twice the density found on Substrates 2 and 4, and four and five times the density present on Substrates 5 and 3 respectively.

##### Canopy area

Substrate 1 has the largest mean canopy area, reflecting the greater density of large individuals. The mean canopy area on Substrates 2, 3 and 4 are similar, differing by a maximum of only 1 m<sup>2</sup>. Substrate 5 had the smallest mean canopy area, reflecting the high density of small individuals on this site. Games–Howell test (used because of unequal variances and sample sizes) revealed statistically significant differences in mean canopy between all substrates ( $p < 0.05$ ). Cohen's  $d$  (a measure of effect size) revealed most of these differences to have a medium or large effect (Table 7-5). Size inequality was greatest on Substrate 3, followed by Substrates 1, 5, 2 and 4 in descending order (Table 7-4).



Table 7-4 Size-structure statistics for the *S. supranubius* population on each of the five substrates. All statistics are computed on a reduced dataset (N) where individuals on the border of the sample window have been removed to avoid underestimation of the mean canopy area and inaccurate statistical estimates. The column 'CV' contains the coefficient of variation (standard deviation / mean) in canopy area.

Substrate	N	Density (individuals /ha)			Mean canopy area (m <sup>2</sup> ) (s.e.)		Standard deviation of canopy area (m <sup>2</sup> )	CV
		All	Small	Medium	Large			
1	1892	98	17	50	31	21.3 (0.6)	27.6	1.300
2	2418	97	17	68	13	10.6 (0.2)	10.4	0.985
3	715	55	17	32	6	9.9 (0.7)	17.6	1.781
4	3012	123	14	94	15	10.9 (0.2)	9.6	0.886
5	515	103	43	51	8	7.3 (0.4)	8.1	1.110

Table 7-5 Effect sizes of pair-wise comparisons of the mean *S. supranubius* canopy area on different substrates, measured using Cohen's *d*. Cohen's *d* is a measure of effect size (i.e., a measure of the strength of the relationship between two variables). A value between 0.2 and 0.5 equates to a 'small' effect, a value between 0.5 and 0.8 equates to a 'medium' effect whereas a value of 0.8 or greater indicates a 'large' effect (Cohen, 1988).

Substrate	1	2	3	4	5
1	-				
2	0.764	-			
3	1.250	0.729	-		
4	0.589	0.323	1.061	-	
5	1.607	1.217	0.412	1.608	-

### **Size distribution**

On all substrates the size distribution of *S. supranubius* was highly positively skewed, with abundance decreasing with plant size (Figure 7-3). On all substrates, therefore, the majority of individuals have a canopy area that is smaller than the mean. Larger size classes have relatively few individuals, but some individuals attain very large sizes (especially on Substrate 1). The log-normal function provided the closest fit to the *S. supranubius* size–abundance data for all individuals on Substrate 1 to 4 (Table 7-6). Substrate 5 appears to have a different size–abundance distribution to the other four substrates, being most accurately described by a power-law distribution. Compared to the best-fitting function, there was little empirical support for the competing functions ( $\Delta AIC > 10$ ; Burnham and Anderson, 2002). The Weibull distribution provided the second closest fit on Substrates 1, 2 and 4. On Substrate 3, the second closest fit was provided by the power-law distribution. The worst description of the size–abundance distribution on Substrate 5 was given by the Weibull distribution.

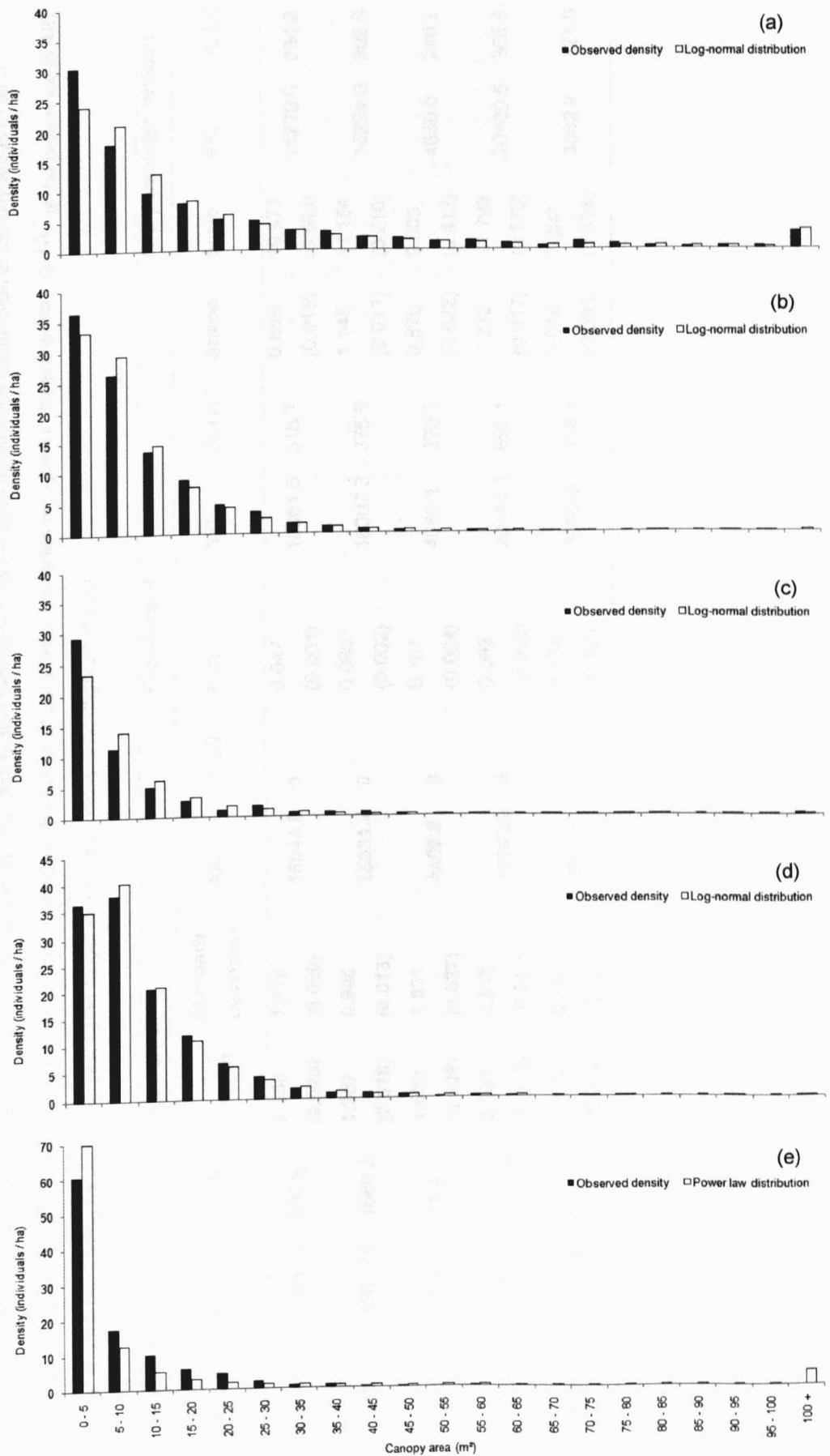


Figure 7-3 The observed (black columns) and best-fitting (white columns) canopy area distribution of *S. supranubius* individuals on Substrates 1 to 5 (graphs (a) to (e)).

Table 7-6 Parameter estimates and Akaike Information Criterion (AIC) values for four distribution functions fit to the size–abundance distribution of canopy areas of all *S. supranubius* individuals (sample size 'n') on each of the five substrates. Standard deviations of parameter estimates are given in parentheses.  $\Delta AIC$  is the difference in AIC between each model and the best-fitting model. The best fitting model in each size class is highlighted in bold.

Substrate	Power-law				Log-normal			Exponential			Two-parameter Weibull				
	n	Exponent	AIC	$\Delta AIC$	Mean	Standard Deviation	AIC	$\Delta AIC$	Rate	AIC	$\Delta AIC$	Shape	Scale	AIC	$\Delta AIC$
1	1893	1.424	15951.1	906.3	2.356 (0.028)	1.219 (0.020)	15044.8	0	0.047 (0.001)	15361.5	316.7	0.859 (0.015)	19.473 (0.553)	15279.0	234.2
2	2428	1.510	17617.6	1686.2	1.957 (0.018)	0.908 (0.013)	15931.4	0	0.095 (0.002)	16307.3	375.9	1.145 (0.017)	11.154 (0.210)	16234.0	302.6
3	711	1.603	4502.6	92.8	1.659 (0.038)	1.021 (0.027)	4409.8	0	0.101 (0.004)	4680.1	270.3	0.870 (0.022)	9.005 (0.413)	4650.0	240.2
4	3015	1.485	22829.9	3067.9	2.062 (0.015)	0.815 (0.010)	19762.0	0	0.092 (0.002)	20414.1	652.1	1.272 (0.017)	11.799 (0.179)	20130.5	368.5
5	516	1.677	2961.9	0	1.478 (0.044)	0.988 (0.031)	2981.0	19.1	0.138 (0.006)	3080.0	118.1	1.010 (0.033)	7.302 (0.338)	3082.9	121.0

### 7.4.3 STRUCTURE OF DIFFERENT SIZE CLASSES

Following the analysis of environmental heterogeneity, the homogeneous  $g(r)$  was used to analyse the spatial pattern of *S. supranubius* on Substrates 1, 2 and 5, whereas the inhomogeneous  $g(r)$  was used on Substrates 3 and 4 (Figure 7-4). All analyses, with the exception of large individuals on Substrates 3 and 5, exceeded the minimum sample size of 70 individuals recommended by Wiegand et al. (2007a). The hard-core distance increased with the size classes reflecting the increase in canopy extent. At scales exceeding the hard-core distance, the use of real-shape or point analysis had little effect types of pattern identified. The only exceptions were in the analysis of large individuals on Substrates 3 and 5 where the function produced by analysis of points was more jagged than that produced by real-shape analysis. This was probably a result of the relatively low sample sizes in these size classes (Table 7-3; Wiegand and Moloney, 2004).

Interpretation of the *S. supranubius* spatial structure uses the results of the point analysis. Any dispersion or decreases in  $g(r)$  detected below the hard-core distance are not discussed, as it is not possible to separate real dispersion from the effects of canopy extent at these scales. Increases or peaks in  $g(r)$  below and above the hard-core distance are, however, discussed as these represent real aggregation rather than an artefact of the canopy extent.

The distribution of small individuals on all five substrates was aggregative (corresponding with Hypothesised Process 1; Figure 7-3). Aggregation of small individuals was strongest on Substrate 2 followed by Substrates 1, 4, 3 and 5 in descending order (Table 7-7). On all substrates the modal inter-shrub distance between small individuals occurred at distances of between 1 and 3 m (corresponding with hypotheses 1 and 3). On Substrates 2 and 3 the scale of maximum aggregation coincided with the hard-core distance (Table 7-7). On Substrate 1 and 4 maximum aggregation occurred at scales below the hard-core distance, whereas the scale of maximum aggregation exceeded the hard-core distance on Substrate 5.

Table 7-7 The hard-core distance, maximum magnitude of  $g(r)$  ( $g_{max}(r)$ ) and scale ( $r$ ) of  $g_{max}(r)$  for each *S. supranubius* size class on each of the five substrates. Both the hard-core distance and the scale ( $r$ ) are measured in metres.

Substrate	Small			Medium-sized			Large		
	Hard-core distance	$g_{max}(r)$	$r$	Hard-core distance	$g_{max}(r)$	$r$	Hard-core distance	$g_{max}(r)$	$r$
1	3	4.90	2	4	2.73	4	8	1.40	10
2	2	6.47	2	4	1.82	4	8	1.42	11
3	3	3.91	3	4	2.98	4	10	3.27	10
4	3	4.15	1	4	1.42	4	8	1.61	9
5	2	3.60	3	5	1.62	3	7	1.86	8

The distribution of medium-sized individuals was also aggregative on all five substrates (Figure 7-4). On all substrates the aggregation among medium-sized individuals was weaker than the aggregation among small individuals (corresponding with Hypothesised Process 2; Table 7-7). The strongest aggregation of medium-sized individuals occurred at 4 m, except on Substrate 5 where the strongest aggregation was at 3 m (Figure 7-3; Table 7-7). However, the strength of aggregation in the medium-sized shrubs does not follow the strength of aggregation among small individuals (Table 7-7). The strongest aggregation among medium-sized individuals was observed on Substrate 3, followed by Substrates 1, 2, 5 and 4 (Table 7-7). As with the small individuals, the  $g(r)$  for medium-sized shrubs remained above 1 at all scales on Substrates 1, 2, 3 and 5 (Figure 7-4). On Substrates 1 to 4 the scale of the maximum aggregation coincided with the hard-core distance (Table 7-7). On Substrate 5, however, the strongest aggregation of medium-sized individuals occurred at scales below the hard-core distance.

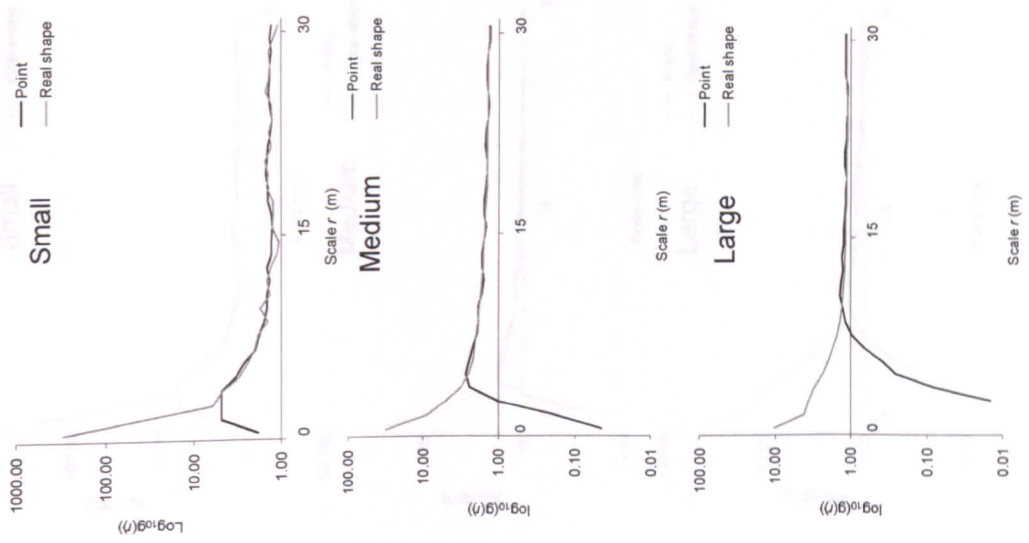
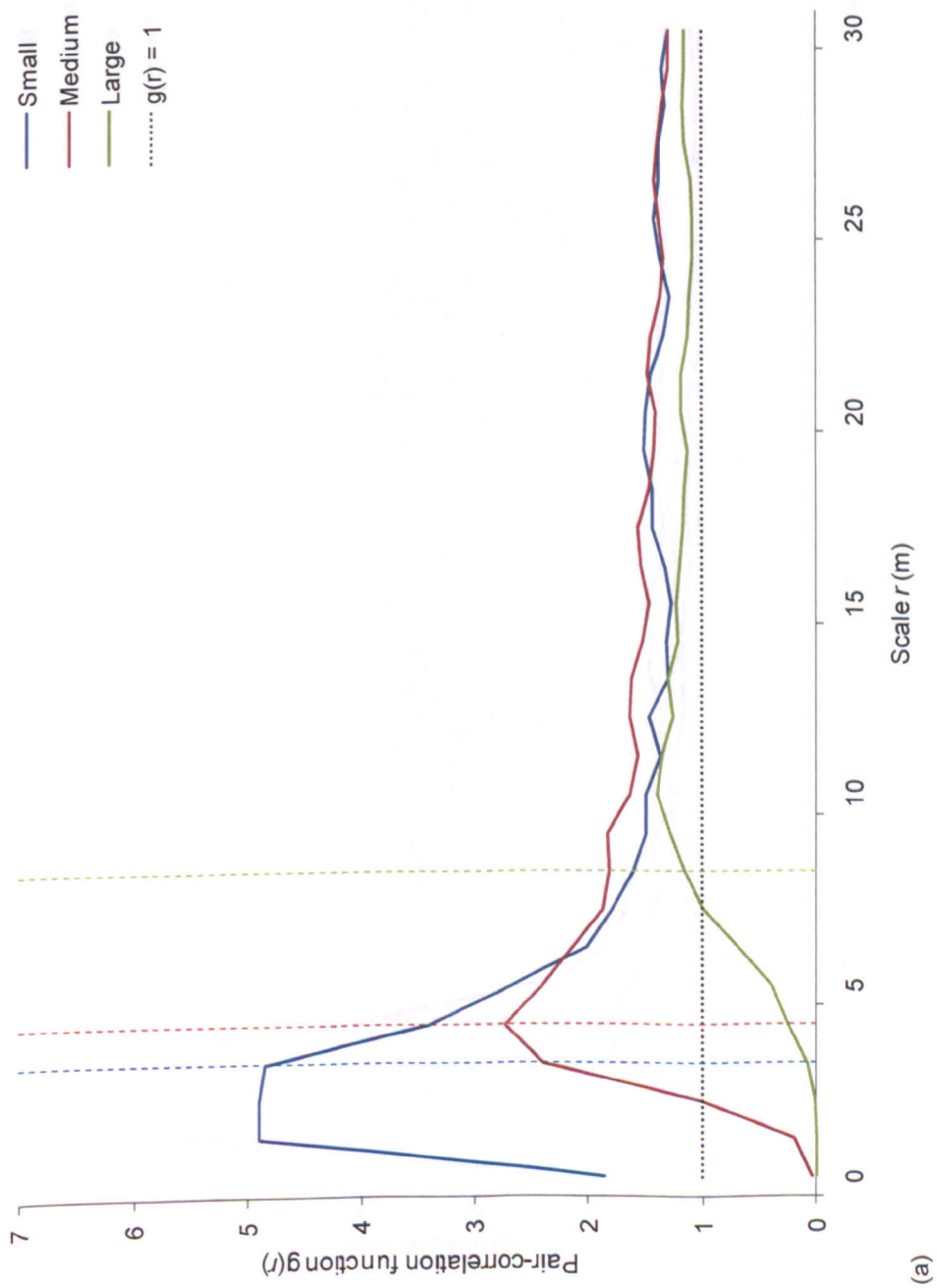
Large individuals were aggregated on all substrates (Figure 7-4). The scale at which maximum aggregation occurred increased to between 8 and 11 m (Table 7-7). On all substrates this either coincided with, or exceeded, the hard-core distance. On Substrates 1 and 2 the aggregation was weak, with the function reaching a maximum height of  $g(r) = 1.4$  at scales of 10 and 11 m respectively (Figure 7-4a, b; Table 7-7).

On both Substrates 1 and 2 this aggregation is weaker than observed in previous size classes (corresponding with Hypothesised Process 2). Similar aggregation strengths were detected on Substrate 4 with maximum aggregation occurring at 9 m (Figure 7-4d; Table 7-7). On this substrate, however, the aggregation is stronger than the aggregation among medium-sized individuals. This increase in aggregative strength is also observed on Substrates 3 and 5 (disagreeing with Hypothesised Process 2; Figure 7-4c, e). Furthermore, the spatial pattern of large individuals on the latter two substrates shows notable structure at larger scales. Large individuals were dispersed between 20 and 21 m on Substrate 5, and there was a dip in  $g(r)$  to randomness between 15 and 18 m on Substrate 3. Re-analysing these data with a real-shape approach (because of small sample sizes) detected slight dispersion of large individuals between 29 and 39 m on Substrate 3, and between 16 and 21 m and 36 and 41 m on Substrate 5 (Appendix F).

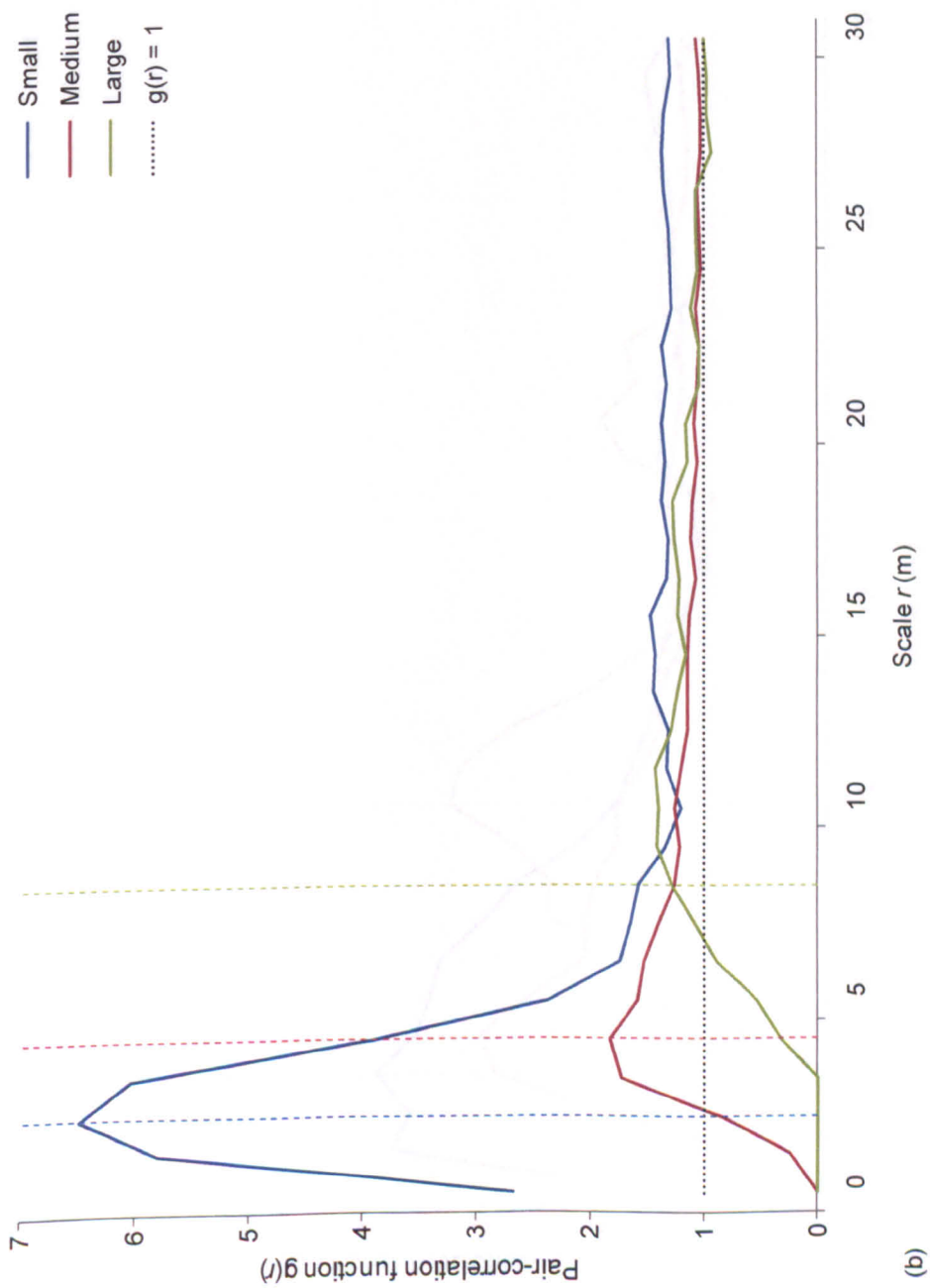
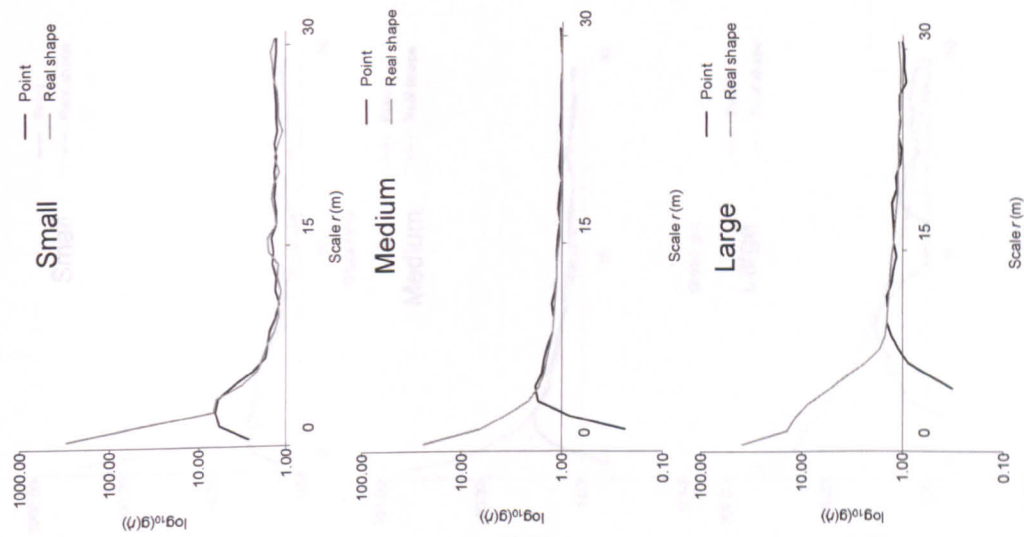
On all substrates the scale of aggregation and the hard-core distance increased with size class (Table 7-7 and Figure 7-4). There is a general trend across all substrates for the scale of maximum aggregation to coincide with, or fall below, the hard-core distance among small and medium-sized individuals. Among large individuals the scale of aggregation exceeds the hard-core distance by up to 3 m.

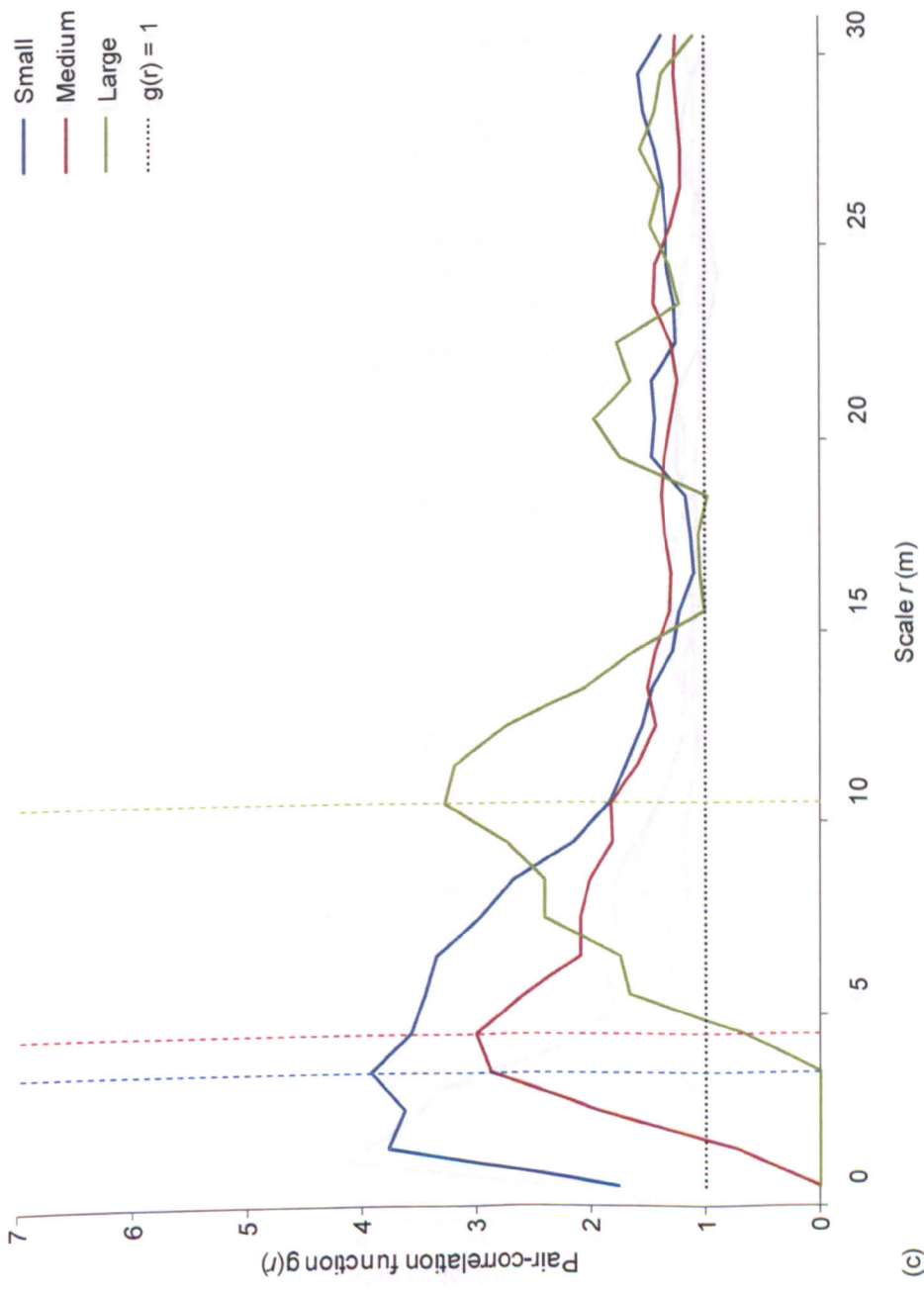
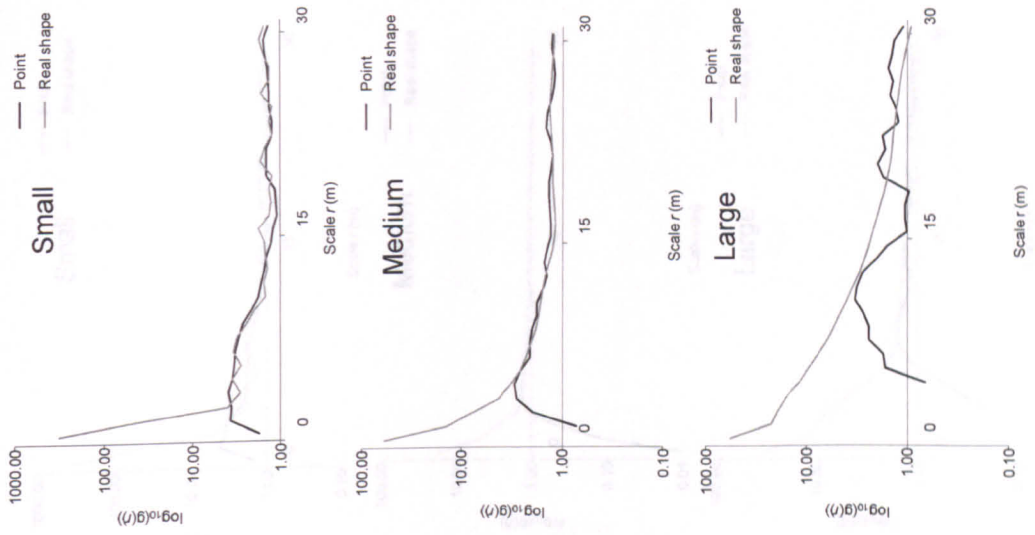
---

Figure 7-4 (pages 203 – 207) The spatial pattern of different size classes of *S. supranubius* on five substrates of contrasting spatial environmental heterogeneity. Substrates 1 to 5 are shown in graphs (a) to (e) respectively. The inset graphs compare the results of analyses using real-shape data (grey lines) and analyses using point data (black lines). The main graph shows the results of analyses using point data. The solid lines show the result for each size class (small, medium and large – blue, red and green lines respectively), whereas the vertical dotted lines show the hard-core distance below which point analysis  $g(r)$  may be influenced by the shape and size of individual shrubs. The black, horizontal dotted line at  $g(r) = 1$  show the expected value under the null model of complete spatial randomness. The homogenous  $g(r)$  was used on substrates 1, 2 and 5, and the inhomogeneous  $g(r)$  was used on substrates 3 and 4. The intensity functions used in the inhomogeneous  $g(r)$  were constructed from positions of the largest individuals ( $\geq 30 \text{ m}^2$ ).

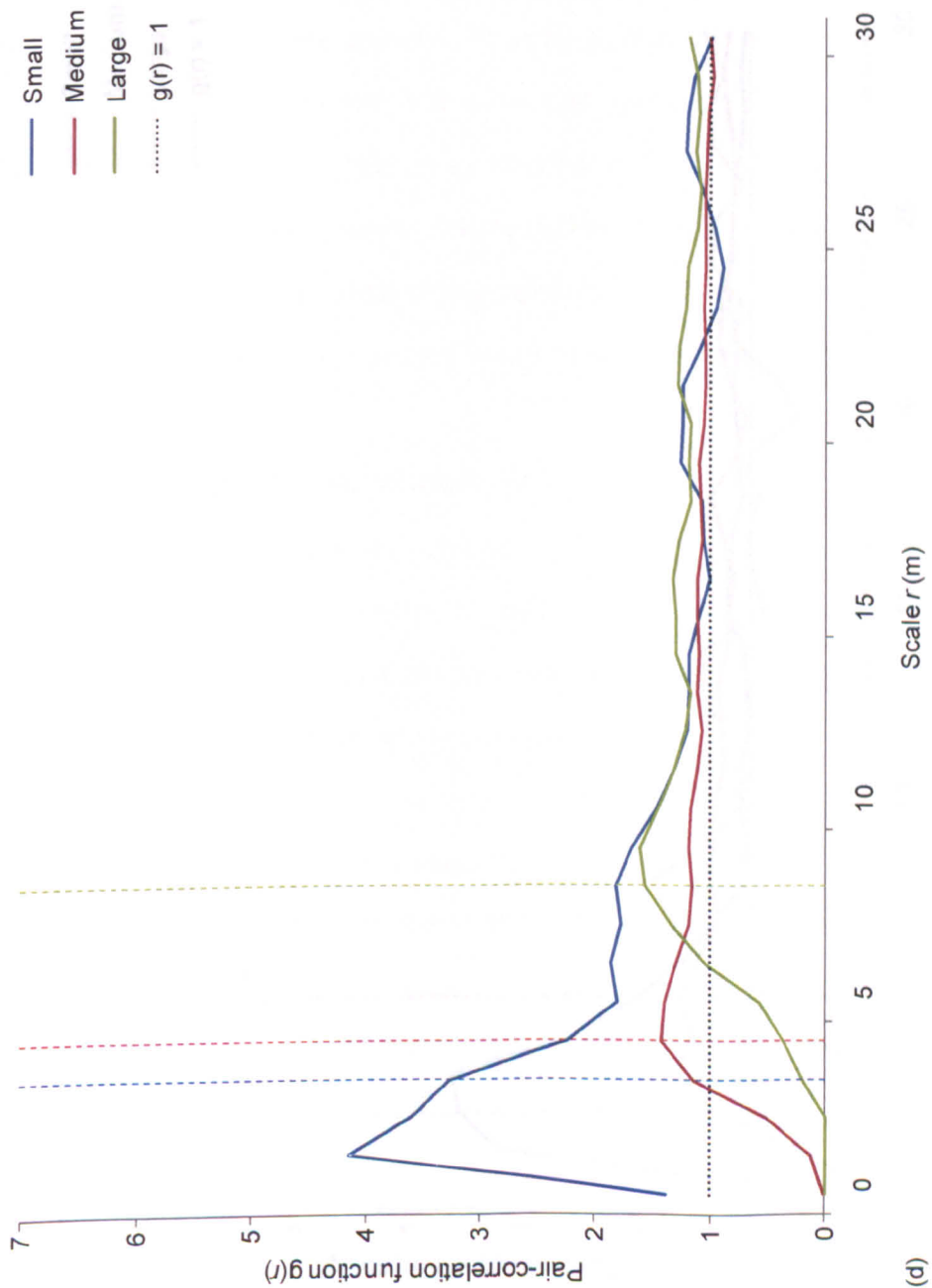
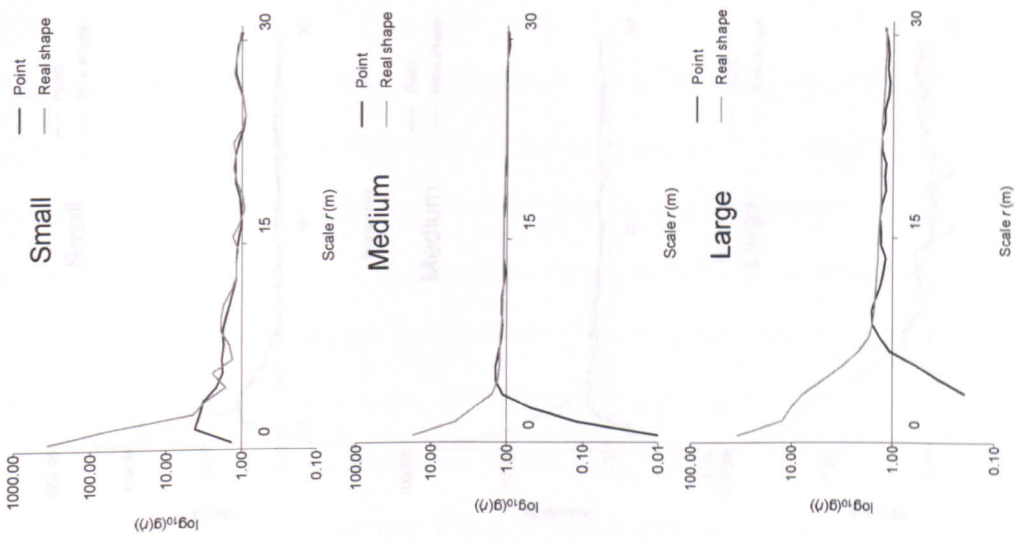




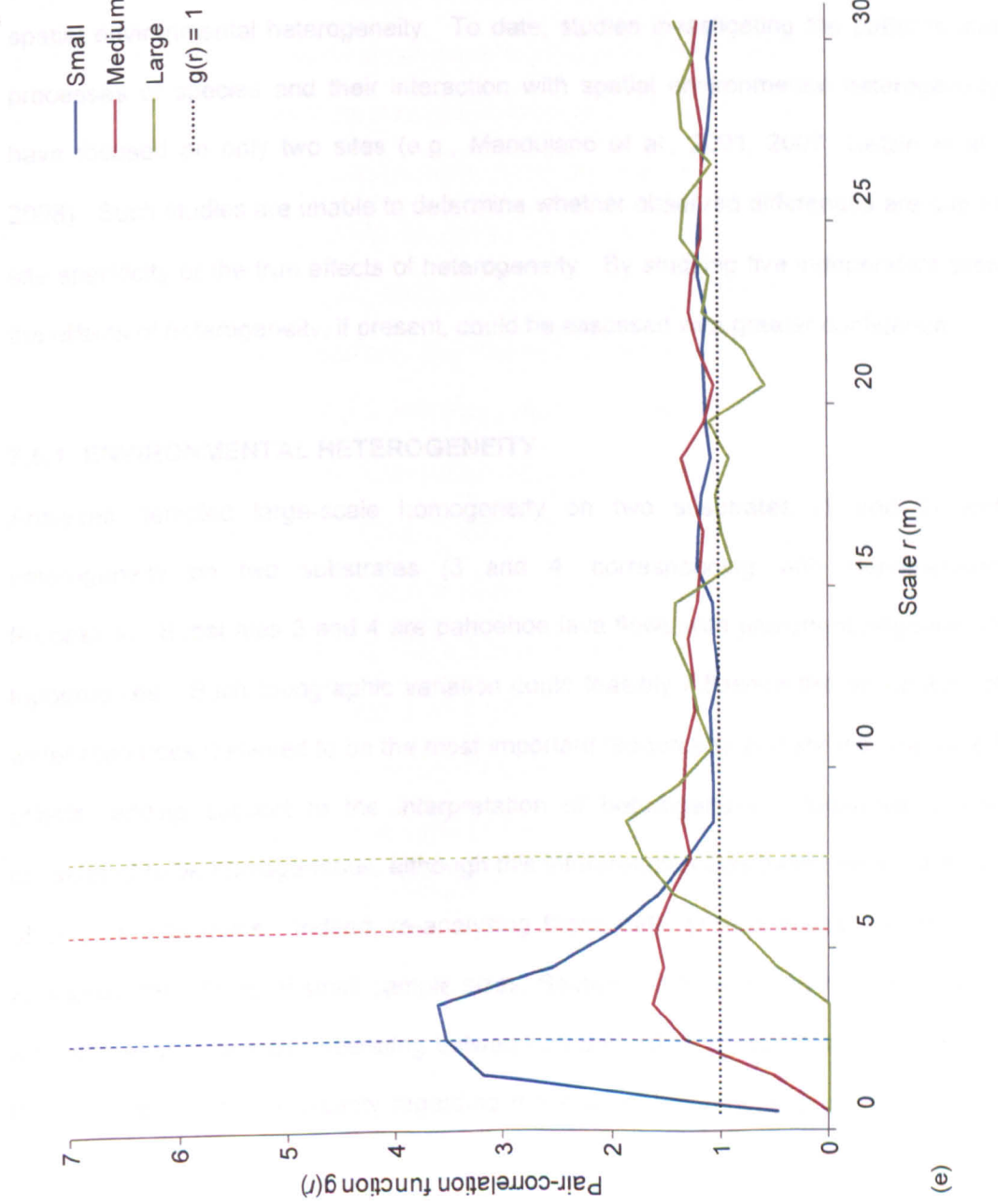
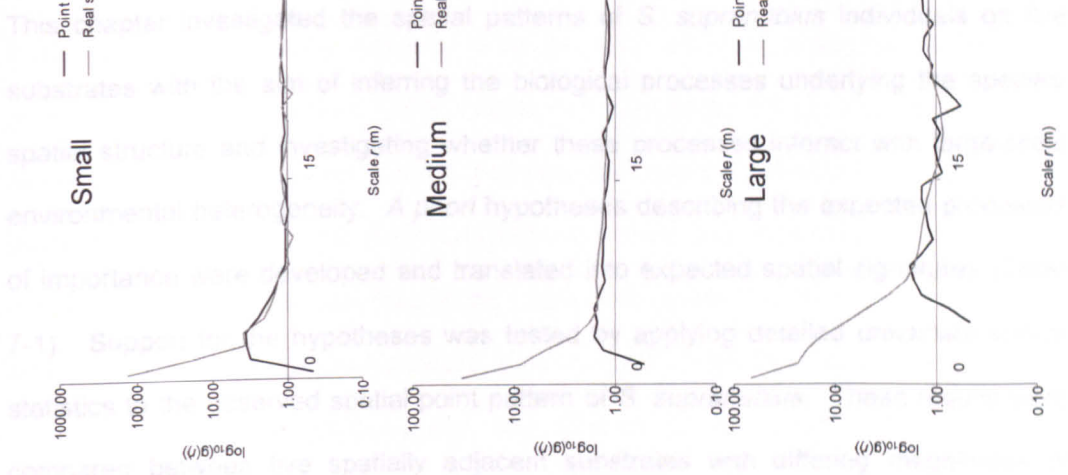




(c)



7.5 DISCUSSION



(e)

## 7.5 DISCUSSION

This chapter investigated the spatial patterns of *S. supranubius* individuals on five substrates with the aim of inferring the biological processes underlying the species' spatial structure and investigating whether these processes interact with large-scale environmental heterogeneity. *A priori* hypotheses describing the expected processes of importance were developed and translated into expected spatial signatures (Table 7-1). Support for the hypotheses was tested by applying detailed univariate spatial statistics to the observed spatial point pattern of *S. supranubius*. These results were compared between five spatially adjacent substrates with differing magnitudes of spatial environmental heterogeneity. To date, studies investigating the patterns and processes of species and their interaction with spatial environmental heterogeneity have focused on only two sites (e.g., Mandujano et al., 2001, 2007; Getzin et al., 2008). Such studies are unable to determine whether observed differences are due to site specificity or the true effects of heterogeneity. By studying five independent sites the effects of heterogeneity, if present, could be assessed with greater confidence.

### 7.5.1 ENVIRONMENTAL HETEROGENEITY

Analyses detected large-scale homogeneity on two substrates (1 and 2) and heterogeneity on two substrates (3 and 4; corresponding with Hypothesised Process 4). Substrates 3 and 4 are pahoehoe lava flows with prominent ridge-trough topographies. Such topographic variation could feasibly influence the distribution of water resources (believed to be the most important resource to arid shrubs) via run-off effects, adding support to the interpretation of heterogeneity. Substrate 5 was considered to be homogeneous, although this interpretation was based on an analysis of only 15 individuals. Indeed, re-analysing these data as 'real shape' (which may ameliorate the effects of small sample sizes; Section 6.4.3) indicated the presence of a heterogeneous process operating between 21 and 30 m on Substrate 5 (Appendix E). Because of the uncertainty regarding the level of heterogeneity on Substrate 5,

this discussion will focus most attention on Substrates 1 to 4, where the heterogeneity results are more conclusive.

Water availability is widely believed to be the main driver of ecological processes in arid environments (Noy-Meir, 1973; Walker and Langridge, 1997). Thus the observed heterogeneity on Substrates 3 and 4 may represent spatial variation in the provision of plant-available water. As inter-substrate distances are relatively small, external precipitation inputs are assumed to be consistent across the study area. Thus the spatial heterogeneity in plant-available water will be determined by the local balance of infiltration and evaporation, and any surface or subsurface flow. Soil and geomorphic conditions can determine the availability of plant-available water via their control on infiltration rates, depth of moisture storage and evaporative losses (Grayson et al., 2006). Broad-scale geomorphologically and edaphically induced variation in moisture regimes in arid environments have been shown to influence vegetation distribution (Bisigato et al., 2009) as well as the spatial patterns (Schenk et al., 2003), abundance (Hamerlynck et al., 2002; Bestelmeyer et al., 2006), physiological activity (Hamerlynck et al., 2000), mortality (Hamerlynck and McAuliffe, 2008) and competitive interactions (Hamerlynck et al., 2002) amongst arid perennials. Similarly, substrate-driven water availability at the local scale can drive the response of individual plants to precipitation events (Pérez, 2000, 2003). The potential for soil and geomorphic properties to vary at several spatial scales means they can theoretically produce highly heterogeneous plant-available water distributions. It is proposed that the ridge–trough topography on Substrates 3 and 4 drives meso-scale spatial variation in plant-available water, and thus the heterogeneous distribution of the largest individuals. Specifically, run-off effects are expected to reduce the quantity of water infiltrating the ridges, resulting in increased plant-available water in the intervening troughs. This is investigated further in Chapter 8.

### **7.5.2 SIZE–ABUNDANCE DISTRIBUTION**

The *S. supranubius* populations on all five substrates followed a similar size hierarchy with monotonically declining abundance with increasing plant size. On Substrate 4

this distribution was interrupted by a lower abundance in the first size class (0 – 5 m<sup>2</sup>; Figure 7-3). This reverse-J distribution has been observed in other semi-arid perennials (Barbour, 1969; Fonteyn and Mahall, 1981; Turner, 1990). The reverse-J distribution is often interpreted as representative of populations with a constant rate of recruitment and time-dependent survivorship and is often observed in clonal species (Mandujano et al., 2007).

Understanding the size–abundance distribution of a population can allow inferences to be made about the processes underlying its structure. Size distributions in nature are typically Gaussian with the average size of an individual dictated by physical and biological constraints (Manor and Shnerb, 2008). However, two recent studies (Kefi et al., 2007; Scanlon et al., 2007) have shown that vegetation patches in arid zones are power-law distributed. Scanlon et al. (2007) concluded that the power law distribution of tree clusters arose from the interaction of resource constraint effects (water availability) and local facilitation. Other studies have suggested that power law distributions can occur when a system shows self-organised criticality (Allen et al., 2008). The power-law distribution, however, did not provide a good description of the *S. supranubius* size–abundance distribution except on Substrate 5. On the remaining substrates the lognormal distribution provided the best description of the *S. supranubius* size–abundance distribution. These results suggest that the population dynamics of *S. supranubius* (i.e., growth and mortality) on Substrate 1, 2, 3, and 4 are approximately equivalent whereas Substrate 5 may have different *S. supranubius* dynamics.

### **7.5.3 SPATIAL PATTERN OF DIFFERENT SIZE CLASSES**

Wheeler and Dickenson (1990) provide the only other known study of *S. supranubius* distribution patterns. Using Clark and Evans' (1954) nearest neighbour technique, they measured the distances separating 100 pairs of *S. supranubius* individuals and concluded that the species was uniformly distributed, a conclusion that this chapter is at variance with. Following germination experiments, Wheeler and Dickenson (1990) suggested intraspecific competition for water and/or herbivory as the mechanisms

behind the distribution of individuals. Although field observations suggest the presence of herbivory, it is not expected to affect the success of clonal reproduction in *S. supranubius* (Kyncl et al., 2006). The current research supports their suggestion of the importance of intraspecific competition (supporting Hypothesised Process 2).

It has previously been assumed that abiotic heterogeneity will affect the broad-scale distribution of plants (i.e., the first-order properties of spatial patterns), but is less likely to affect the distribution and size of individuals relative to one another (i.e., the second-order properties of patterns). By using the inhomogeneous  $g(r)$  to remove the effect of environmental heterogeneity on Substrates 3 and 4, this study found notable differences in the second-order properties of *S. supranubius* patterns between homogeneous and heterogeneous sites, indicating that *S. supranubius* interactions and demographics are affected by habitat characteristics.

### **The spatial structure of small individuals**

Some 70% of all plant species display capability for clonal growth (Klimeš et al., 1997). Despite the ubiquity of clonal reproduction in arid shrubs (Schenk, 1999), no studies, to my knowledge, have considered how this process, and specifically branch-layering, influences the spatial structure of a species. That vegetative reproduction can generate distinct spatial structures has been acknowledged in studies of herbaceous species (Mahdi and Law, 1987; Kenkel, 1993; Oborny and Cain, 1997; Pottier et al., 2007), but there has remained a lack of understanding of the role of clonal propagation in the population spatial structure of arid-zone shrubs (Jiménez-Lobato and Valverde, 2006). The spatial patterns of small individuals were consistent with the operation of clonal reproduction (hypotheses 1 and 3). On all substrates the peak inter-shrub distance between small individuals occurred at small scales (between 1 and 3 m), with high frequencies of inter-shrub distances relative to CSR observed up to 10 m. This indicates that small individuals in clumps are separated from their neighbours by between 1 and 3 m. Furthermore, clonal reproduction appears to have a strong influence on the spatial structure of older cohorts, with small-scale aggregation persisting among medium-sized individuals on all substrates (Figure 7-4).



Clonal reproduction may explain the reverse-J shaped size–abundance distribution and the constant recruitment rate this distribution implies.

As with species of similar physiology (McAuliffe et al., 2007), field observations indicate that *S. supranubius* individuals generate clonal offspring only when they have attained a certain size. It is assumed that individuals are only capable of reaching these sizes in resource-rich/sufficient areas. As such, clonally reproduced ramets are also assumed to be located in resource-rich sites. With no anticipated differences in inter-ramet competition on the different substrates, it was hypothesised that any clonal signature should show little variation between substrates of differing heterogeneity (Golubski et al., 2008; Hypothesised Process 3). However, although clonal reproduction appears to be occurring on all substrates, the relative importance of the process seems to vary. The magnitude of aggregation among small individuals is notably stronger on the two homogeneous substrates (Substrates 1 and 2) compared to the heterogeneous substrates (Substrates 3 and 4). Assuming the observed spatial pattern is indeed a signature of clonal reproduction, the results suggest that clonal reproduction is more prevalent or more successful on homogeneous substrates (disagreeing with Hypothesised Process 3).

Recent studies highlight the importance of environmental heterogeneity in describing plant regeneration dynamics (Barbeito et al., 2009 and references therein). McAuliffe et al. (2007) concluded that the successful growth and clonal reproduction of *L. tridentata*, an arid shrub known to vegetatively reproduce in a similar fashion to *S. supranubius* (Vasek, 1980; McAuliffe et al., 2007), is largely dependent on substrate conditions. McAuliffe et al. (2007) contend that vegetative spread in *L. tridentata* requires fine, continually renewed aeolian deposits which have high rates of infiltration and moisture storage. These conditions, they propose, enhance individual plant performance and prospects for long-term survival, which are necessary for clonal development. The homogeneous substrates have much more extensive coverage of fine surface sediments than the heterogeneous substrates (pers. obs.; Table 2-1). However, large individuals were present on all substrates

indicating that all the substrates provide the resource conditions required for clonal reproduction. Therefore, in addition to the geo-hydrologic relationships proposed by McAuliffe et al. (2007), it is hypothesised that the geomorphologic conditions on the heterogeneous substrates may have a direct effect on reproduction by physically restricting the adventitious rooting of branches. Both Substrates 3 and 4 have prominent rocky ridges, with Substrate 4 also noted for the dominance of large, unsorted surface clasts in the intervening troughs (Table 2-1). Alternatively, the fine surface materials on Substrates 1 and 2 may enable lateral branch rooting and thus extensive clonal reproduction (Illa et al., 2006).

### **The spatial structure of medium-sized and large individuals**

On all substrates the pattern of medium-sized and large individuals was predominantly aggregative. The scale of maximum aggregation among medium-sized individuals was 4 m on all substrates except Substrate 5, perhaps suggesting inter-substrate consistency in the demographic processes operating on medium-sized shrubs. There was no apparent relationship between the magnitude of aggregation among either medium-sized or large individuals, and the heterogeneity of the substrate. However, there was a notable difference in the relative aggregation of medium-sized and large individuals on heterogeneous and homogeneous substrates. On the homogeneous substrates the strength of aggregation decreases in subsequent size classes with the weakest aggregation observed among large individuals, corresponding with the operation of density-dependent competition (e.g., Meyer et al., 2008; Metsaranta and Lieffers, 2008; Gray and He, 2009; Hypothesised Process 2). On both homogeneous substrates the large individuals show only minimal deviation from complete spatial randomness. Conversely, on the heterogeneous substrates the weakest aggregation is among the medium-sized individuals, with aggregative strength increasing among large individuals. As well as exceeding the intra-substrate magnitude of aggregation of medium-sized individuals, the aggregative strength of large individuals on Substrates 3 and 4 exceeds that of large individuals on Substrates 1 and 2. It therefore appears that the patterns of stand development on Substrate 3 and 4 are

more complex than predicted by competition alone (Metsaranta and Leiffers, 2008; Hypothesised Process 2).

It has been proposed that as desert shrubs mature they gain access to less readily depletable, deep-soil moisture reserves, reducing the total impact of both inter- and intra-specific competition (Golluscio et al., 1998; Toft and Fraizer, 2003). Thus, it may be predicted that upon reaching a large size, mortality should become density independent and older cohorts should revert to a random distribution as individuals are lost from the system because of stochastic events. This explanation fits the distribution of large individuals on the homogeneous substrates which show only small deviations from complete spatial randomness. However, on the heterogeneous substrates the increase in aggregation among large individuals suggests that the relative risk of mortality increases with isolation as the shrubs age. Ontogenetic shifts in biological interaction have been previously reported (Miriti, 2006), but these generally describe an increase in competition, not facilitation, as individuals age. It is presumed that competition and facilitation in arid environments are driven by water availability. Thus, one possible explanation is that the geological make-up of the heterogeneous substrates has prevented the formation of deep water reserves, or the distribution of deep water reserves is spatially heterogeneous. Alternatively, the geological make-up of the heterogeneous substrates may prevent the roots of *S. supranubius* from penetrating the substrate to access the deep water reserves. The positive feedback between plant biomass and infiltration is widely recognised (Rietkerk et al., 2004; Ludwig et al., 2005). Increasing aggregation among large individuals may increase infiltration enhancing the survival of individuals in clumps relative to isolated individuals. Alternatively, it is possible that clustering of large individuals has climatic benefits by reducing low-level wind speeds and thus reducing wind-induced desiccation. However, this explanation is deemed less plausible because of the sheltering effects of the caldera walls.

#### 7.5.4 A SIGNATURE OF COMPETITIVE THINNING

On all substrates the decrease in aggregation among medium-sized individuals relative to small individuals is consistent with commonly cited hypotheses of density-dependent mortality (Hypothesised Process 2). Further evidence in support of the operation of competition is provided by the comparison of the hard-core distance to the scale of maximum aggregation. On all substrates except Substrate 5, the scale of maximum aggregation of small and medium-sized individuals equalled or was slightly less than the hard-core distance (Table 7-7). In practice this means that within each cohort, an individual's neighbours are most commonly found immediately adjacent to their canopy, corresponding with the observed modal shrub–shrub distance of between 1 and 3 m. Occasionally the canopies of these individuals overlap (causing the scale of maximum aggregation to fall below the hard-core distance), but mostly they do not (i.e., scale of maximum aggregation equals the hard-core distance; Figure 7-5). There may, however, be unobserved overlap in the root zones. The increase in both the hard-core distance and the scale of maximum aggregation with size is strong evidence for density-dependent competition. On all substrates except Substrate 3 the scale of maximum aggregation among large individuals exceeds the hard-core distance by between 1 to 3 m (Table 7-7). In practice that means that within the large cohort individual canopies do not overlap or even touch (Figure 7-5). Instead they are separated by 1 to 3 m at least. The increase in the spacing of large individuals relative to small and medium-sized individuals suggests that competition may become more spatially extensive as individuals reach large sizes (i.e., zone of influence increases disproportionately with canopy size; Figure 7-5). Furthermore, the distance between the hard-core distance and the scale of maximum aggregation should increase as competition increases in strength. Because the distance between the hard-core distance and the scale of the strongest pattern among large individuals was greatest on Substrates 1 and 2, the results suggest that competition, at least among large individuals, is strongest on the two homogeneous substrates. The decrease in the maximum  $g(r)$  of large individuals, compared to small or medium-sized individuals on Substrates 1 and 2 reflects the loss of previous neighbours whose canopies/root

systems began to interfere with other, more competitive individuals (i.e., density-dependent competition; Figure 7-4a and b; Table 7-7).

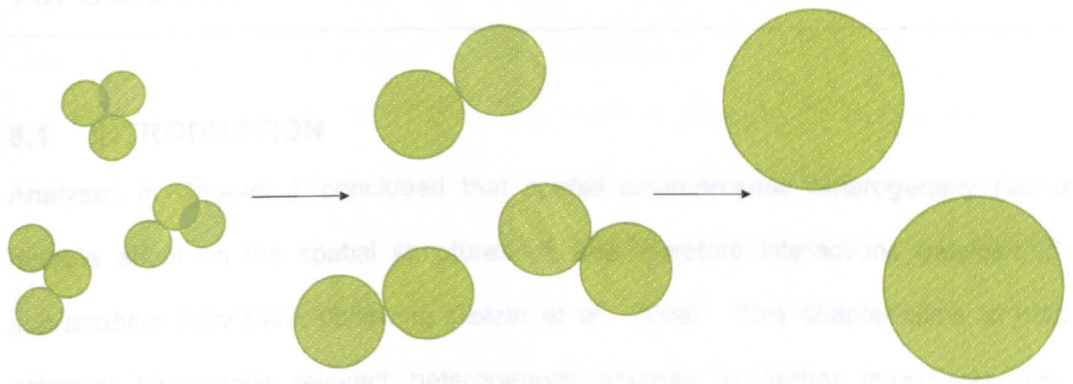


Figure 7-5 A signature of competitive thinning derived from the changes in the difference between the hard-core distance and the maximum scale of aggregation as cohorts age.

## 7.6 CONCLUSIONS

It is generally accepted that while abiotic heterogeneity can affect the first-order structure of a species, it operates on too large a scale to have a quantitative influence on the second-order structure of a population. This study, however, has demonstrated that after removing the effects of environmental heterogeneity, neighbouring lava flows that differ in age and geomorphological structure, but which are not expected to differ significantly in climate, had notably different *S. supranubius* population spatial structures. Differences in the inter-substrate first- and second-order properties of *S. supranubius* spatial pattern are both attributed to substrate geomorphological characteristics. However, whereas differences in the former are attributed to the effect of geomorphology on plant-available water, differences in second-order properties are attributed to the physical effects of geomorphology on the physiology of individuals. This latter effect is hypothesised to limit both the rooting of lateral branches (and thus clonal reproduction), and the access of large shrubs to deep water reserves on the heterogeneous substrates. Thus, in accordance with other studies (Hamerlynck et al., 2000, 2002; Peters et al., 2006) it is proposed that population spatial structures in arid environments cannot be understood without an understanding of how the soil-geomorphic template influences the spatial distribution of plant-available water.

## CHAPTER 8: SPATIAL VARIATION IN THE DENSITY AND LOCAL SPATIAL STRUCTURE OF *S. SUPRANUBIUS*: THE ROLE OF TOPOGRAPHY

---

### 8.1 INTRODUCTION

Analyses in Chapter 7 concluded that spatial environmental heterogeneity had a notable effect on the spatial structures of, and therefore interactions between, *S. supranubius* individuals (following Getzin et al., 2008). This chapter aims to map potential biologically relevant heterogeneity sources to further investigate their influence on the density and the spatial structures of *S. supranubius* populations.

Understanding the relationship between abiotic and biotic processes, and how they influence population dynamics, is a fundamental aim of ecology (Dahlgren and Ehrlén, 2009). Some studies have compared the spatial patterns of plants under different abiotic conditions (most commonly comparing patterns under different fire [Fulé and Covington, 1998; Park, 2003; Yu et al., 2009] or disturbance regimes [Wells and Getis, 1999; Call and Nilsen, 2003; Fajardo and Alaback, 2005; Malkinson and Kadmon, 2007; Appendix A]). However, little attention has been paid to investigating and quantifying the effect of continuous abiotic gradients on plant population spatial structure, perhaps because the plot sizes commonly used are too small to contain noteworthy abiotic gradients (Figure 1-3). Consequently, there is surprisingly little understanding of how continuous spatial environmental heterogeneity may interact with biological processes to determine population dynamics (Wagner and Fortin, 2005; Murrell, 2009). If abiotic factors influence both long-term demographic processes (e.g., establishment and survival) and interactions between individuals, then both the density and second-order structure of the population should change simultaneously with abiotic gradients. Investigation of such phenomena has, to date, been limited to a study by Shimatani and Kubota (2004), in which they quantified changes in the spatial pattern of a coniferous tree by constructing and assessing the fit of a novel inhomogeneous point process model that incorporated spatial variation in both density and point–point interactions.

To investigate the role of spatial environmental heterogeneity, that heterogeneity must be identified and mapped. Soil water availability is widely believed to be the main driver of ecological processes in arid environments (Noy-Meir, 1973; Walker and Langridge, 1997; Grayson et al., 2006; Robertson et al., 2009) and is therefore assumed to be important in the present system. Consequently, understanding the spatial structure of *S. supranubius* may require knowledge of the spatial distribution of water. Water availability at any one location is determined by the balance between the vertical transfer (infiltration and evaporation) and horizontal redistribution (surface and subsurface) of water (Grayson et al., 2006). This chapter assumes that any feature that significantly influences local hydrology is likely to have an effect on overlying vegetation structure and dynamics. Field studies have demonstrated that substrate characteristics in arid regions, such as clast size (Diaz et al. 2005), clast depth (Tejedor et al., 2002) and clast sorting (Pérez, 2000; Tejedor et al., 2003), can influence the local balance of infiltration and evaporation and thus the moisture availability in upper soil layers. This can have subsequent effects on the distribution and dynamics of individual plants (Pérez, 2003; Hamerlynck et al., 2002). However, woody shrubs which remain physiologically active throughout the year, such as *S. supranubius*, are expected to be more affected by the distribution and dynamics of water stored at depth (Gebauer et al., 2002; Schenk and Jackson, 2002). Therefore, the spatial structure and intensity of *S. supranubius* may be expected to correlate with abiotic features that influence the recharge of deep water reserves.

Deep water stores are recharged when the storage capacity of upper soil layers is exceeded (Grayson et al., 2006; Wilcox et al., 2006a). Thus deep water recharge occurs when there are large precipitation pulses (Gebauer et al., 2002), or when the water from small precipitation pulses is concentrated into confined locations. Topography and slope influence the horizontal redistribution of water from precipitation events (Chaplot and Le Bissonnais, 2000; Wilcox et al., 2003; Grayson et al., 2006) and therefore may be important in determining the spatial distribution of deep water recharge. Topographically driven water availability has previously been shown to affect the density and biomass of arid vegetation (Imeson and Prinsen,

2004; Ju et al., 2008; Hamerlynck and McAuliffe, 2008; Svoray et al., 2008; Popp et al., 2009). However, most of our current understanding concerns how topography interacts with hydrological processes to determine vegetation patterning at the patch scale, for example the generation of banded vegetation patterns or 'tiger bush' (Saco et al., 2007; McDonald et al., 2009). Our understanding of how topography influences vegetation dynamics at the scale of individuals is still poor.

This chapter investigates whether topography influences the density and spatial structure of the *S. supranubius* population on Substrate 3. Substrate 3 is the youngest of the focal sites and has prominent ridge-trough topography (see Table 2-1). Chapter 7 concluded that intra-specific competition may be an important process structuring *S. supranubius* populations, and that the importance of competition may be influenced by spatial environmental heterogeneity. Consequently, this chapter investigates the presence of spatial structures that are consistent with the operation of intra-specific competition as an important organising force. The results are interpreted in relation to the following conceptual model of topographically driven water redistribution.

### **8.1.1 A MODEL OF WATER REDISTRIBUTION ON SUBSTRATE 3**

Despite common acceptance of the importance of water availability in driving biological processes in arid systems, there are very few datasets of sufficient duration and spatial extent to quantify aspects of water availability that are relevant to arid shrub dynamics (Breshears et al., 2009). Given the large inter-annual variability in arid precipitation events (Snyder and Tartowski, 2006), the typically short duration of precipitation events, and the slow demographic responses of arid shrubs (Cody, 2000; Bowers, 2005), many years of continuously collected data would be required to empirically investigate the relationship between water availability and shrub dynamics. Neither time nor financial resources allowed such measurements to be made. Consequently, *a priori* knowledge on the focal sites and the influence of geomorphological characteristics on the spatial partitioning of rainfall events (e.g., Monger and Bestelmeyer, 2006) was used to develop the following conceptual model.



Substrate 3 is composed of rocky ridges with intervening troughs of pumice and erosional deposits. It is hypothesised that the distribution of the ridges and the slope of the terrain will produce linearised spatial variation in water availability. The model predicts that there will be four zones of alternating high and low water availability between each ridge and the centre of the neighbouring trough (Figure 8-1).

#### **Zone 1: low water availability**

Zone 1 is located on the rocky ridges where the slope of the terrain is relatively steep. Runoff magnitude increases with increasing slope, reducing the chance of infiltration (Chaplot and Le Bissonnais, 2000; Wilcox et al., 2003; Monger and Bestelmeyer, 2006). Furthermore, the presence of large rocks will reduce the volume available for moisture storage to the areas of soil/humus trapped between rocks. Consequently, it is hypothesised that the majority of the precipitation in this zone will be lost via run-off.

#### **Zone 2: high water availability**

Zone 2 occurs where the pumice troughs adjoin the rocky ridges. The zone represents a change in substrate characteristics and a sudden decrease in the slope of the terrain. Sandy and coarse textured soils have a high infiltration rate and enable deep drainage (Wilcox et al., 2006a; Popp et al., 2009). Much infiltration of run-off from zone 1 is expected to occur and recharge deep water reserves in zone 2.

#### **Zone 3: low water availability**

Much of the run-off from zone 1 has already been absorbed in zone 2. Therefore, zone 3 largely relies on precipitation inputs. It is hypothesised that much of the precipitation will be lost to evaporation before it can reach deep storage.

#### **Zone 4: high water availability**

Sub-surface topography is expected to result in the sub-surface flow of the water infiltrated in zone 2 through zone 3 to accumulate in zone 4. The subsurface movement of water is hypothesised to generate a reservoir of deep water in the centre of troughs.

8.1.2 AIMS AND OBJECTIVES

Two main hypotheses are outlined

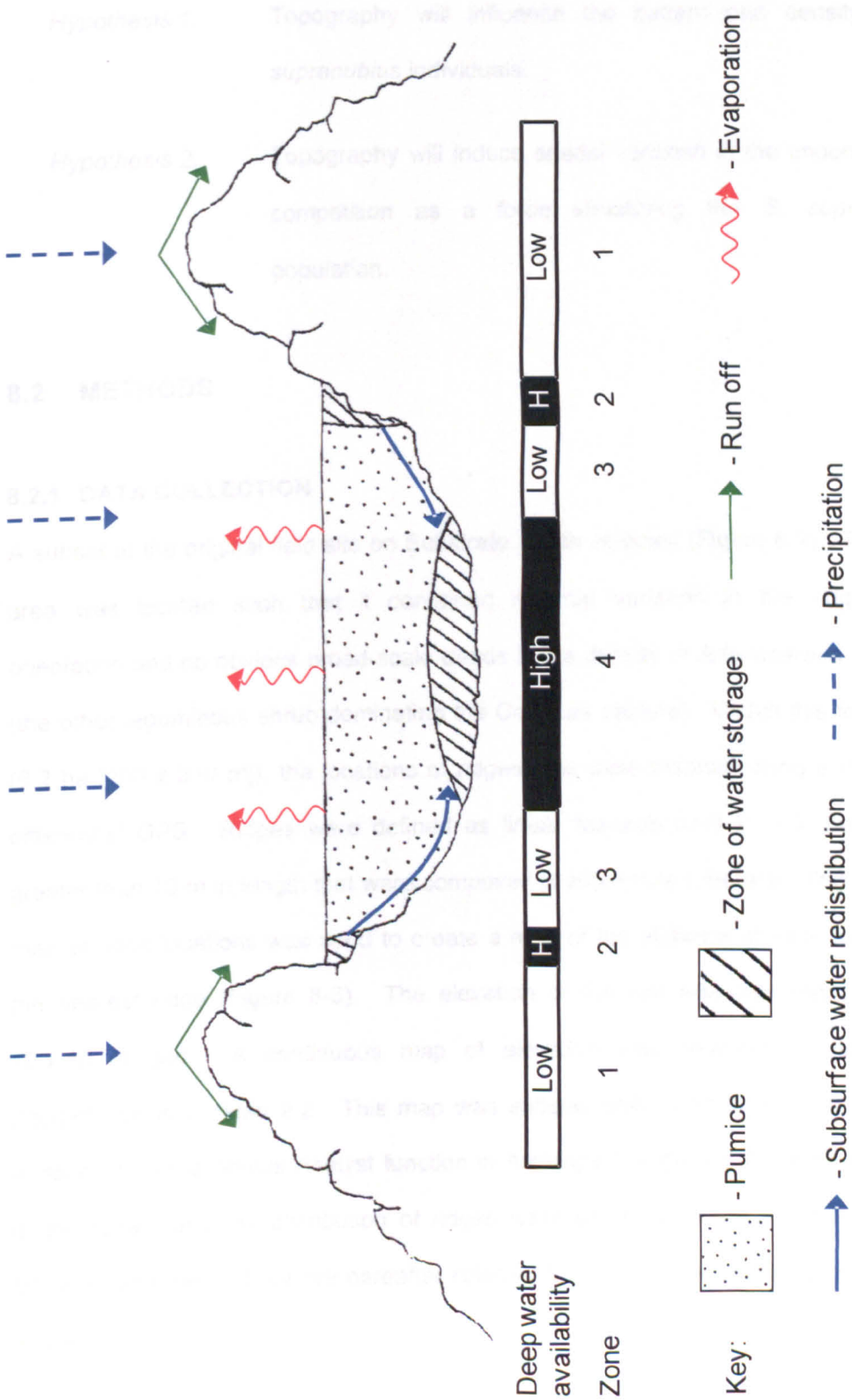


Figure 8-1 Conceptual model of the spatial variation in deep water availability on Substrate 3.

### 8.1.2 AIMS AND OBJECTIVES

Two main hypotheses are outlined:

*Hypothesis 1:* Topography will influence the pattern and density of *S. supranubius* individuals.

*Hypothesis 2:* Topography will induce spatial variation in the importance of competition as a force structuring the *S. supranubius* population.

## 8.2 METHODS

### 8.2.1 DATA COLLECTION

A subset of the original field site on Substrate 3 was selected (Figure 8-2). The focal area was located such that it contained minimal variation in the ridge-trough orientation and no obvious broad-scale trends in the density of *Adenocarpus viscosus* (the other leguminous shrub dominating the Cañadas caldera). Within this focal area (6.2 ha [200 x 310 m]), the locations of ridges tops were mapped using a Promark3 differential GPS. Ridges were defined as linear features over 2 m in height and greater than 10 m in length that were composed of more than one rock. The resulting map of ridge locations was used to create a map of the distance of each location to the nearest ridge (Figure 8-3). The elevation of the site was also mapped on a 10 x 10 m grid. A continuous map of elevation was produced using kriging interpolation in ArcMap 9.2. This map was subsequently used to generate a slope surface using the Spatial Analyst function in ArcMap9.2 (Figure 8-4). Both the slope of the terrain and the distribution of ridges were used as spatial covariates in the following analyses. They are hereafter referred to as the slope and ridge covariates respectively.

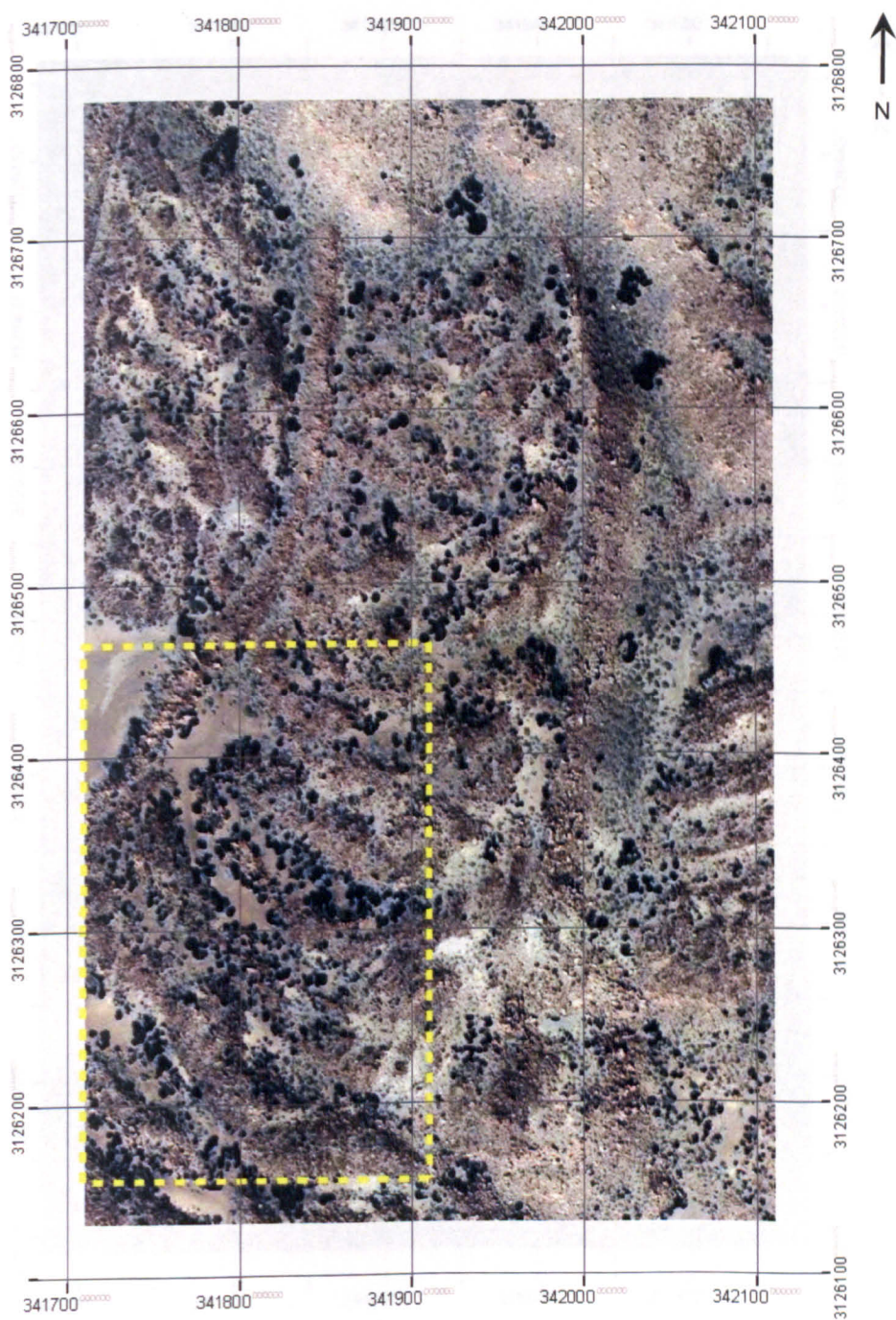


Figure 8-2 The focal site for Chapter 8 (highlighted in the yellow box) – a subset of Substrate 3.



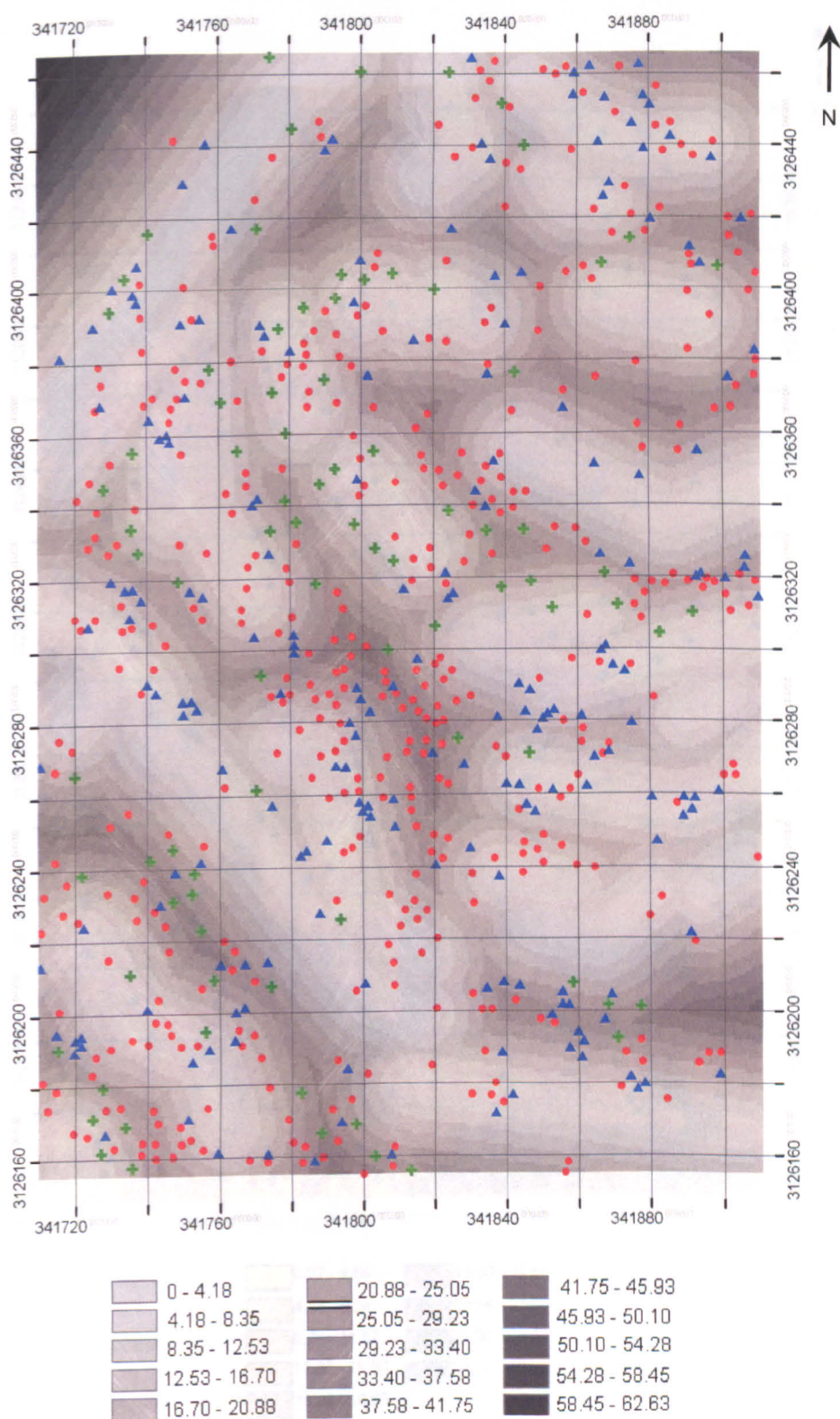


Figure 8-3 Raster image (resolution 1 m) showing the distance to the nearest ridge top (m). Superimposed on the image are the locations of *S. supranubius* individuals (small = blue, medium-sized = red, large = green). NB: although shown as a categorical map a continuous surface was used during analysis.

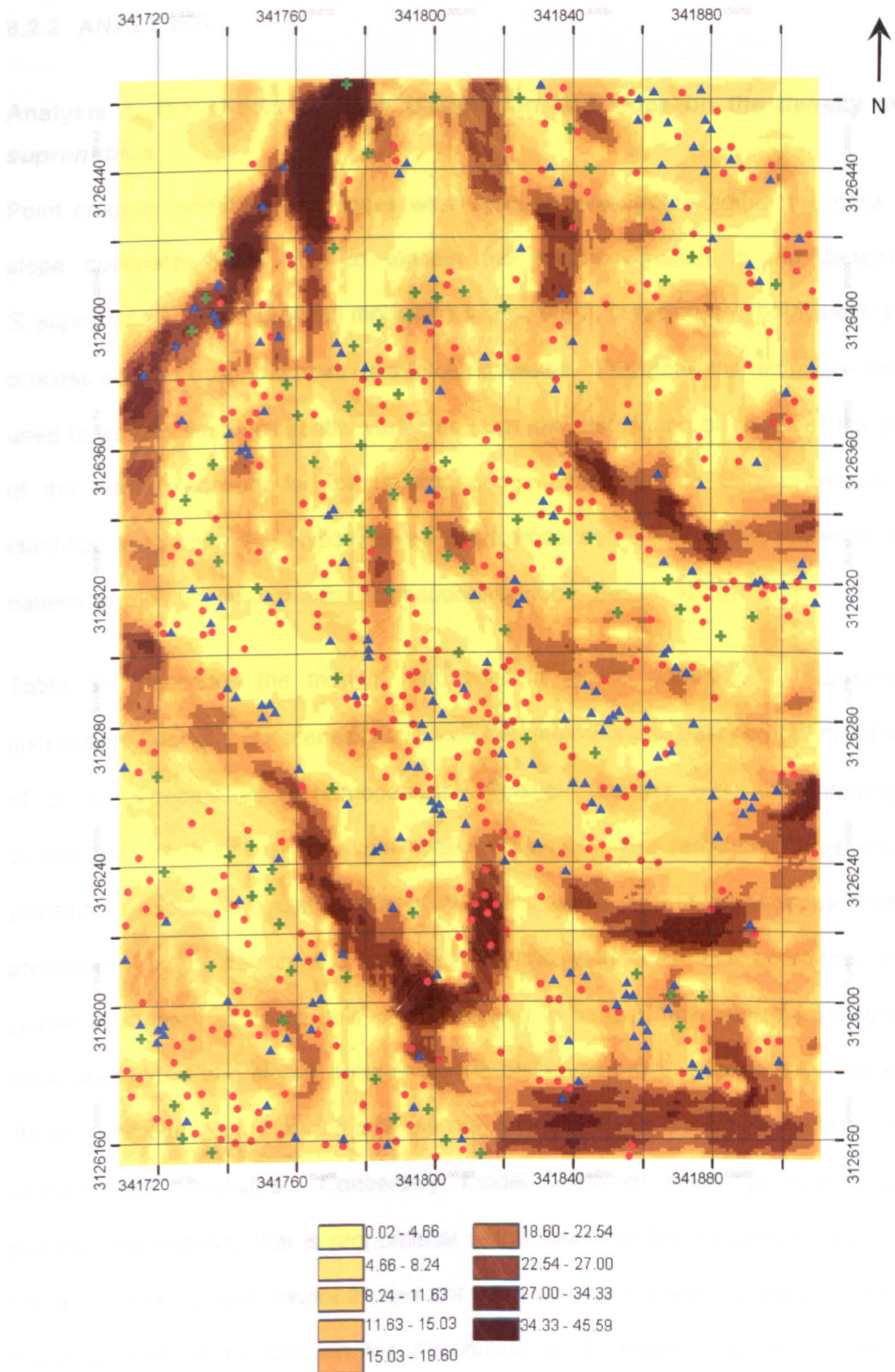


Figure 8-4 Raster image showing the slope (degrees) of the terrain. Superimposed on the image are the locations of *S. supranubius* individuals (small = blue, medium-sized = red, large = green). NB: although shown as a categorical map a continuous surface was used during analysis.

## 8.2.2 ANALYSES

### **Analysis 1: the effects of ridge distribution and slope on the density of *S. supranubius***

Point process modelling techniques were used to investigate whether the ridge and slope covariates were able to explain the spatial variation in the density of *S. supranubius*. Point process modelling allows the fit between a hypothesised point process and the observed point data to be assessed. Therefore, the technique can be used to assess how well spatial variables (e.g., the distribution of ridges or the slope of the terrain) account for the heterogeneous distribution of *S. supranubius* (as identified in Chapter 7). Hypothesised point processes are fit to the observed point pattern by the method of maximum pseudo-likelihood (Besag, 1975).

Table 8-1 describes the models that were fit to the observed *S. supranubius* distribution. Model 1 is essentially a null model, attempting to describe the distribution of *S. supranubius* as a homogeneous Poisson process (i.e., the intensity of *S. supranubius* does not vary with spatial location). The remaining three models attempt to explain the distribution of *S. supranubius* as inhomogeneous Poisson processes (i.e., the density of *S. supranubius* varies with location in accordance with a spatial covariate). Two models are developed to assess how well the topography covariates (ridge and slope) account for the distribution of *S. supranubius*. Model 3 fits an inhomogeneous Poisson process with intensity that is a loglinear function of the topographical covariates. Conversely, model 4 fits an inhomogeneous Poisson process with intensity that is proportional to the values of the topography covariates. Unlike models 3 and 4, model 2 does not use measured spatial covariates. Instead, model 2 attempts to describe the distribution of *S. supranubius* using Cartesian effects. The identification of Cartesian effects that account for the distribution of *S. supranubius* may imply the presence of gradients in abiotic conditions that have not been directly measured.

Table 8-1 The four point process models fit to the observed point pattern of *S. supranubius*.

Model	Description	Intensity term
1	A homogeneous Poisson process where the density $\lambda(u)$ of shrubs is spatially uniform.	$\lambda(u) = a$
2	An inhomogeneous Poisson process where the density $\lambda(u)$ of shrubs is log-linear in the Cartesian coordinates.	$\lambda(u) = \exp\{(a + bx + cy)^2\}$
3	An inhomogeneous Poisson process where the density of shrubs $\lambda(u)$ is a log-linear function of the covariate $Z(u)$ .	$\lambda(u) = \exp(a + bZ(u))$
4	An inhomogeneous Poisson process where the density of shrubs $\lambda(u)$ is proportional to the covariate $Z(u)$ .	$\lambda(u) = aZ(u)$

In the intensity term,  $a$ ,  $b$  and  $c$  are parameters to be estimated from the fitted model, and  $Z(u)$  is the value of the covariate (i.e., the distance to the nearest ridge or the slope of the terrain) at location  $u$ . The ridge and slope covariates are derived from Figure 8-3 and Figure 8-4.

Akaike information criterion (AIC) was used to assess the relative fit of competing hypothetical models to the observed point pattern. The intensity term in model 2 (Table 8-1) was simplified by removing terms to identify the individual Cartesian effect (e.g.  $x$ ,  $xy$ ) that best described the density of *S. supranubius*. The AIC values of the models were compared. The model producing the lowest AIC was assumed to provide the best available description of *S. supranubius* density. The difference between the AIC of each model and the best-fitting model was calculated ( $\Delta AIC$ ). Interpretation of  $\Delta AIC$  follows Burnham and Anderson (2002). Models with  $\Delta AIC > 10$  provide a poor explanation of the variation in the data relative to the best-fitting model (i.e., little empirical support; Burnham and Anderson, 2002). Models with  $4 \leq \Delta AIC \leq 7$  have 'considerably less' support than the best-fitting model (Burnham and Anderson, 2002: p.70), whereas models with  $0 \leq \Delta AIC \leq 2$  have substantial support. If either of the models containing the topographical covariates (i.e., model 3 or 4) provided the lowest AIC value it implies that topographical gradients are important in determining the distribution of *S. supranubius* individuals. However, if both models 3 and 4 produced higher AIC values than models 1 and 2, then two



possible interpretations existed. It could be that topography has little influence on *S. supranubius* density compared to other, unmeasured environmental gradients (cf. model 2). Alternatively, it could be that the relationship between topography and the density of *S. supranubius* is non-proportional (cf. model 4). The hypothesised water redistribution model (Section 8.1.1) suggests that the effects of topography and geomorphology will generate distinct zones of vegetation response rather than a continuous surface, therefore a non-proportional relationship between *S. supranubius* density and the topography covariates was considered to be more likely. Therefore, if models 3 and 4 provided weak explanations of *S. supranubius* density, the topography covariate was adapted to test for discrete heterogeneity effects by dividing the original continuous image into a series of binary images where the value separating the two binary classes took increasing values. Each binary image was used as a covariate in model 3. The binary covariate that produced the lowest AIC value indicates the topographical position (i.e., the slope or the distance from a ridge) which has the greatest effect on the density of *S. supranubius*.

The point process models described in Table 8-1 were fit to the observed pattern of *S. supranubius* in each of the three size classes (Section 7.3.2). Models 3 and 4 (and any binary models) were applied separately using the ridge (Analysis 1a) and the slope covariate (Analysis 1b). Up to this point the models fitted to the *S. supranubius* distribution have contained only one term (i.e., either slope, ridge or Cartesian). In Analysis 1c, models incorporating a combination of these three terms are fit to the *S. supranubius* distribution. To try to achieve the best fit possible, these combined models used the best-fitting binary topography covariates and the best-fitting Cartesian effects as identified in Analyses 1a and 1b. All analyses were performed using the *spatstat* package (v.1.17-2, Baddeley and Turner, 2005) in R (v.2.10.0, R Core Development Team, 2009).

## **Analysis 2: the effect of ridge distribution and slope on the local spatial structure of *S. supranubius***

This analysis investigates the presence of local spatial structures that are consistent with the importance of intra-specific competition as an organising force.

### ***Size-distance correlation as a signal of competition***

Competition may control both the size and the local density of shrubs. Nearest neighbour techniques provide a simple and intuitive approach to exploring local interactions (Perry et al., 2009). Under the presence of competition, the size of an individual is expected to be a function of the size and distance of all neighbouring individuals that fall within its zone of influence. A positive correlation between the size of an individual and the distance separating it from its nearest neighbouring individuals is expected if competition reduces growth (Getzin et al., 2006). Therefore, analysing the correlation between the sum of the area of several nearest neighbours and the sum of distances separating them from the focal shrub provides a good indication of the importance of competition as an organising force at the local scale. This technique has been employed in several recent articles (Schenk et al., 2003; Getzin et al., 2006; Getzin and Wiegand, 2007; Meyer et al., 2008; Gray and He, 2009). Steeper regression slopes between the two variables suggest the presence of local spatial structures that are indicative of an increased importance of competition.

### ***Quantile regression***

The relationship between the sum of the canopy area of the five nearest neighbours (following Gray and He, 2009) plus the canopy area of the focal shrub, and the sum of the distances separating them from the focal shrub, was analysed using quantile regression (following Meyer et al., 2008; Lawes et al., 2008) to reveal whether there was more than a single slope (rate of change) describing the relationship between area and distance (Cade and Noon, 2003). The presence of multiple rates of change would imply inconsistency in local spatial structure and the presence of a factor(s) that interacts with nearest neighbour distances to increase the heterogeneity of shrub size (i.e., an interactive factor that is influencing local spatial structure; Cade and Noon,

2003; Meyer et al., 2005). Further analysis could then be conducted to determine whether either the ridge or the slope covariate could be the interactive factor.

Quantile regression can be used to fit linear (or non-linear) trends to quantile surfaces within the data (Koenker, 2005). Thus, whereas ordinary least-squares regression models the relationship between a variable  $X$  and the mean of the response variable  $Y$ , quantile regression models the relationship between  $X$  and the quantiles of  $Y$ , such as the 75<sup>th</sup> percentile. For example, the  $\tau$ <sup>th</sup> quantile regression function  $Q(\tau)$  describes a linear (or non-linear) fit through the data so that  $\tau$  proportion of the data are located below the regression line ( $Q(\tau)$ ), and  $1 - \tau$  proportion of the data are located above the regression line ( $Q(\tau)$ ). Thus, instead of just modelling the mean effect corresponding to a set of  $x$ s, multiple properties of the distribution are modelled (Guisan et al., 2006). Linear regressions were fit to the 0.95, 0.9, 0.75, 0.5, 0.25, 0.1 and 0.05 quantiles. The slope of the regression line for each quantile was extracted. The estimates of rates of change in quantile regression are semi-parametric in the sense that no parametric distributional form (e.g. normal, Poisson) is assumed for the error (i.e. residuals) of the model (Cade and Noon, 2003). Confidence intervals (90%) were constructed for the slope estimates (following Meyer et al., 2008). Where the lower 90% confidence limit was greater than zero, a significant positive relationship between the combined nearest neighbour distance and the combined area was deemed to exist. Differences in the regression slopes imply that factors other than nearest neighbour distances are having an (interactive) effect on shrub size (Meyer et al., 2005, 2008). Analysis of deviance techniques were used to test the equality of the quantile regression slopes.

Further analyses were performed to investigate whether either the ridge or slope covariate could explain any observed inconsistency in local spatial structure. As the effects of the covariates are anticipated to be non-linear (Section 8.1.1) the covariates could not simply be incorporated as interaction terms in linear regressions between size and distance. Instead, the data were divided into seven subsets of similar sample size such that each subset contained covariate values higher than the previous

subset, but lower than the next subset. Seven subsets were chosen as this provided a large sample size in each subset ( $n \approx 100$ ) but did not reduce the covariate into too few spatial categories. This was done once for the ridge and once for the slope covariate. Within each subset a bivariate linear regression between nearest neighbour distance and combined canopy area was performed and the coefficient of determination ( $R^2$ ) calculated. Differences in  $R^2$  between the data subsets suggest that the covariate influences the local spatial structure. Higher values of  $R^2$  indicate spatial structures that are consistent with competition as an important organising force; lower  $R^2$  values indicate factors other than competition may be driving local spatial structure. In addition to reporting the values of  $R^2$ , the statistical significance of the model and the slopes of the regression lines are also reported. This provides a more robust comparison of the relationship between nearest neighbour distance and combined canopy area in different locations within the terrain, and allows more confidence in making assessments of the likelihood that any observed differences in  $R^2$  may have occurred by chance. Nearest neighbour calculations were performed using Hawth's tools in ArcMap 9.2, quantile regression was performed using *R* and the *quantreg* package (Koenker, 2009).

Shackleton et al. (2002) provide one of the only studies, to my knowledge, that uses the regression between nearest neighbour distances and the sum of shrub sizes to investigate how site factors (e.g. aspect, slope, landscape position) correlate with the presence/absence and relative importance of intraspecific competition between woody savanna species (as measured by  $R^2$  values). They concluded that slope position influenced the presence of competition. To maintain comparability with Shackleton et al.'s (2002) study the regressions discussed above used canopy area ( $m^2$ ) as a measure of shrub size, instead of more commonly used measurements such as canopy diameter (e.g. Meyer et al., 2008).

8.3 RESULTS

8.3.1 ANALYSIS 1A: THE EFFECT OF RIDGES ON THE DENSITY OF *S. SUPRANUBIUS*

Table 8-2 shows the AIC values when the four point process models were fitted to the *S. supranubius* point pattern in each size class. Model 2 (Cartesian trends) provided the best explanation of the density of *S. supranubius* in all three size classes. Relative to the competing models, model 4 (which assumed that shrub density was proportional to the distance to the nearest ridge) has no empirical support. Compared to model 2 there was considerably less support for either model 1 or 3 when attempting to explain the density of medium-sized and large individuals. When fit to the density of small individuals, however, the explanatory power of models 1 – 3 was largely indistinguishable. Therefore, if the distribution of ridges does affect the density of shrubs, the effect is probably non-linear.

Table 8-2 The AIC values (1 d.p.) for each of the four point process models described in Table 8-1 when fitted to the pattern of *S. supranubius* individuals in each of the three size classes. Models 3 and 4 use the ridge covariate. Bold text indicates the model with the lowest AIC.  $\Delta AIC$  calculates the difference in AIC between each model and the best-fitting model. The Cartesian trends providing the best spatial fit are shown in parentheses.

Size class	Model 1		Model 2(best fitting Cartesian model)		Model 3		Model 4	
	AIC	$\Delta AIC$	AIC	$\Delta AIC$	AIC	$\Delta AIC$	AIC	$\Delta AIC$
Small	2708.1	0.8	<b>2707.3</b> (x*y)	0	2709.0	1.7	2855.8	148.5
Medium-sized	5087.4	4.9	<b>5082.5</b> (y)	0	5087.7	5.2	5322.1	239.6
Large	1305.9	6.4	<b>1299.5</b> (x)	0	1304.9	5.4	1311.1	11.6

Figure 8-5 plots the AIC values of models using binary distance classes. For each size class the binary model providing the greatest improvement in fit over the Cartesian effects (model 2) is described.

### **Small individuals**

Compared to the Cartesian model, the binary distance classes did not improve the explanation of the observed point pattern until large distance classes were separated out (Figure 8-5a). The best model divided the ridge distribution covariate into two classes, separated at a distance of 28 m (AIC = 2698.3). The model identified a decrease in density from 33 individuals at distances less than 28 m from a ridge, to 0 individuals per hectare at greater distances. Compared to this model, the best fitting Cartesian model had  $\Delta AIC = 9$ , and thus provided a considerably worse explanation of the data.

### **Medium-sized individuals**

The only model which had a better fit than the Cartesian model divided the ridge distribution covariate at 12 m (AIC = 5082.2; Figure 8-5b). This model identified a density of 61 individuals per hectare in locations close (< 12 m) to a ridge compared with a density of 79 individuals per hectare at greater distances. However, compared to this model, the best-fitting Cartesian model had  $\Delta AIC = 0.3$  suggesting that there was little difference in the explanatory power of the two models.

### **Large individuals**

When applied to large individuals, the binary ridge covariate provided a good fit when divided at distances of 5 m (AIC=1288.6; Figure 8-5c). This model identified a low density of individuals in areas less than 5 m from a ridge (3 individuals per hectare) and a higher density of individuals in locations more than 5 m from a ridge (17 individuals per hectare). Compared to this model, the best-fitting Cartesian model had  $\Delta AIC = 10.9$  suggesting that it had relatively little explanatory power.

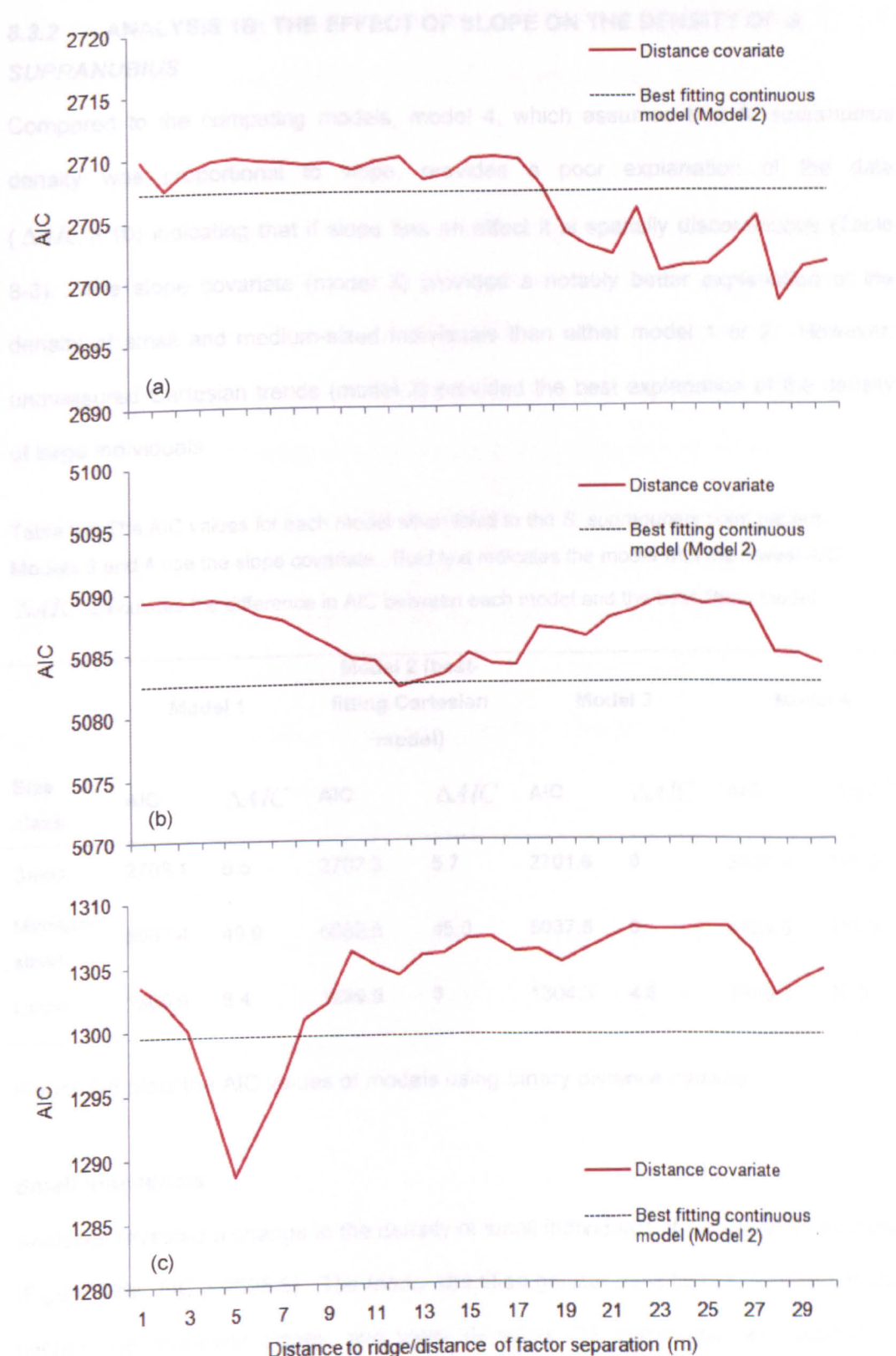


Figure 8-5 The AIC of models using the binary ridge distribution covariates to explain the density of (a) small, (b) medium-sized and, (c) large *S. supranubius* individuals. The red line shows the change in AIC as spatial covariates using different distance classes are fitted to the data. The black dashed line shows the AIC value of the best-fitting continuous model (see Table 8-2).

**8.3.2 ANALYSIS 1B: THE EFFECT OF SLOPE ON THE DENSITY OF *S. SUPRANUBIUS***

Compared to the competing models, model 4, which assumed that *S. supranubius* density was proportional to slope, provides a poor explanation of the data ( $\Delta AIC > 10$ ) indicating that if slope has an effect it is spatially discontinuous (Table 8-3). The slope covariate (model 3) provided a notably better explanation of the density of small and medium-sized individuals than either model 1 or 2. However, unmeasured Cartesian trends (model 2) provided the best explanation of the density of large individuals.

Table 8-3 The AIC values for each model when fitted to the *S. supranubius* point pattern. Models 3 and 4 use the slope covariate. Bold text indicates the model with the lowest AIC.  $\Delta AIC$  calculates the difference in AIC between each model and the best-fitting model.

Size class	Model 1		Model 2 (best-fitting Cartesian model)		Model 3		Model 4	
	AIC	$\Delta AIC$	AIC	$\Delta AIC$	AIC	$\Delta AIC$	AIC	$\Delta AIC$
Small	2708.1	6.5	2707.3	5.7	<b>2701.6</b>	<b>0</b>	2837.4	135.8
Medium-sized	5087.4	49.9	5082.5	45.0	<b>5037.5</b>	<b>0</b>	5454.8	417.3
Large	1305.9	6.4	<b>1299.5</b>	<b>0</b>	1304.3	4.8	1369.5	70.0

Figure 8-6 plots the AIC values of models using binary distance classes.

**Small individuals**

Analyses revealed a change in the density of small individuals at slopes of 16 degrees (Figure 8-6a; AIC = 2696.6). The model identified greater densities (37 individuals per hectare) on shallower slopes, and lower densities (18 individuals per hectare) on slopes of 16 degrees and steeper. This model, however, was practically indistinguishable from the 11 degrees binary model ( $\Delta AIC = 1.2$ ) which also identified greater densities (39 individuals per hectare) on shallower slopes, and lower densities (23 individuals per hectare) on steeper slopes. Compared to these models, the best fitting Cartesian model had little empirical support ( $\Delta AIC = 10.6$  and  $9.4$  respectively),



and the best-fitting continuous slope model (Model 3 [see Table 8-3]) had considerably less support ( $\Delta AIC = 3.8$  and  $5.0$  respectively).

### **Medium-sized individuals**

Binary covariates using a break in slope between 8 and 23 degrees all provided a notably better explanation of the density of medium-sized individuals than the best-fitting Cartesian model (Figure 8-6b). The difference in AIC between these models and the best-fitting Cartesian model was consistently greater than 10 ( $\Delta AIC > 10$ ). The largest  $\Delta AIC$  was obtained when the best-fitting Cartesian model was compared to the covariate identifying a break in slope at 18 degrees ( $\Delta AIC = 56.8$ ). This model identified greater densities on slopes of less than 18 degrees (79 individuals per hectare) and lower densities on slopes steeper than 18 degrees (19 individuals per hectare). Compared to this model, there is very little empirical support for either the Cartesian model ( $\Delta AIC = 56.8$ ) or the best-fitting continuous slope model (Model 3 [see Table 8-3],  $\Delta AIC = 11.9$ ).

### **Large individuals**

All binary slope covariates provided a poorer description of *S. supranubius* density than the best-fitting Cartesian model (Figure 8-6c). The strongest effect of slope was observed at 19 degrees (AIC = 1301.8).

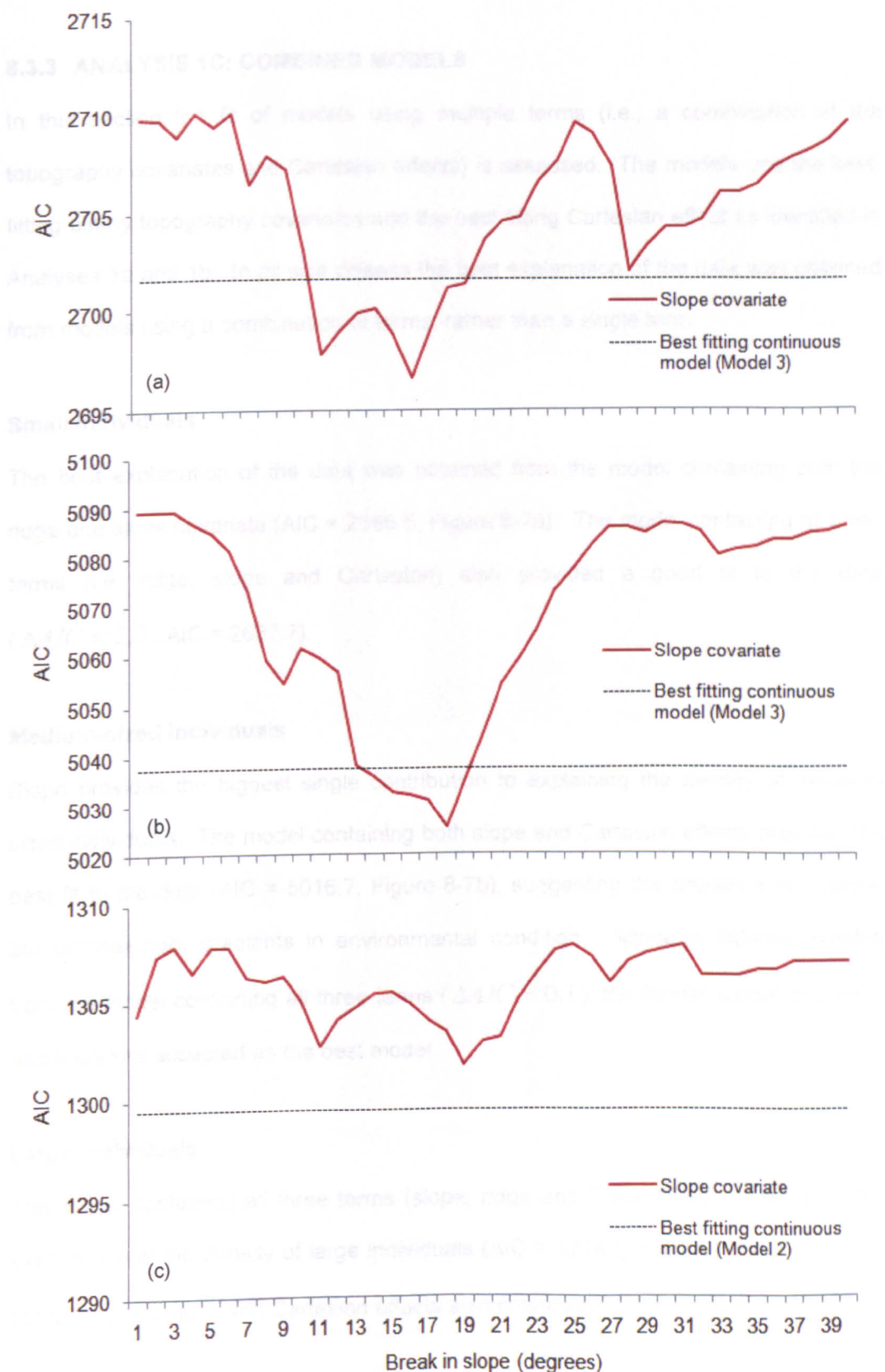


Figure 8-6 The AIC values of models using the binary slope covariates to explain the density of (a) small, (b) medium-sized and, (c) large *S. supranubius* individuals. The red line shows the change in AIC as spatial covariates using different distance classes are fit. The black dashed line shows the AIC value of the best-fitting continuous model (see Table 8-3).

### 8.3.3 ANALYSIS 1C: COMBINED MODELS

In this section the fit of models using multiple terms (i.e., a combination of the topography covariates and Cartesian effects) is assessed. The models use the best-fitting binary topography covariates and the best-fitting Cartesian effect as identified in Analyses 1a and 1b. In all size classes the best explanation of the data was obtained from models using a combination of terms, rather than a single term.

#### Small individuals

The best explanation of the data was obtained from the model containing both the ridge and slope covariate ( $AIC = 2685.5$ ; Figure 8-7a). The model containing all three terms (i.e. ridge, slope and Cartesian) also provided a good fit to the data ( $\Delta AIC = 2.2$ ,  $AIC = 2687.7$ ).

#### Medium-sized individuals

Slope provides the biggest single contribution to explaining the density of medium-sized individuals. The model containing both slope and Cartesian effects provided the best fit to the data ( $AIC = 5016.7$ ; Figure 8-7b), suggesting the presence of notable, but unmeasured, gradients in environmental condition. Although indistinguishable from the model containing all three terms ( $\Delta AIC = 0.1$ ), the former model is simpler and therefore accepted as the best model.

#### Large individuals

The model containing all three terms (slope, ridge and Cartesian) provided the best explanation of the density of large individuals ( $AIC = 1274.3$ ; Figure 8-7). The model containing only ridge and Cartesian effects also provided a good fit ( $\Delta AIC = 1.7$ ).

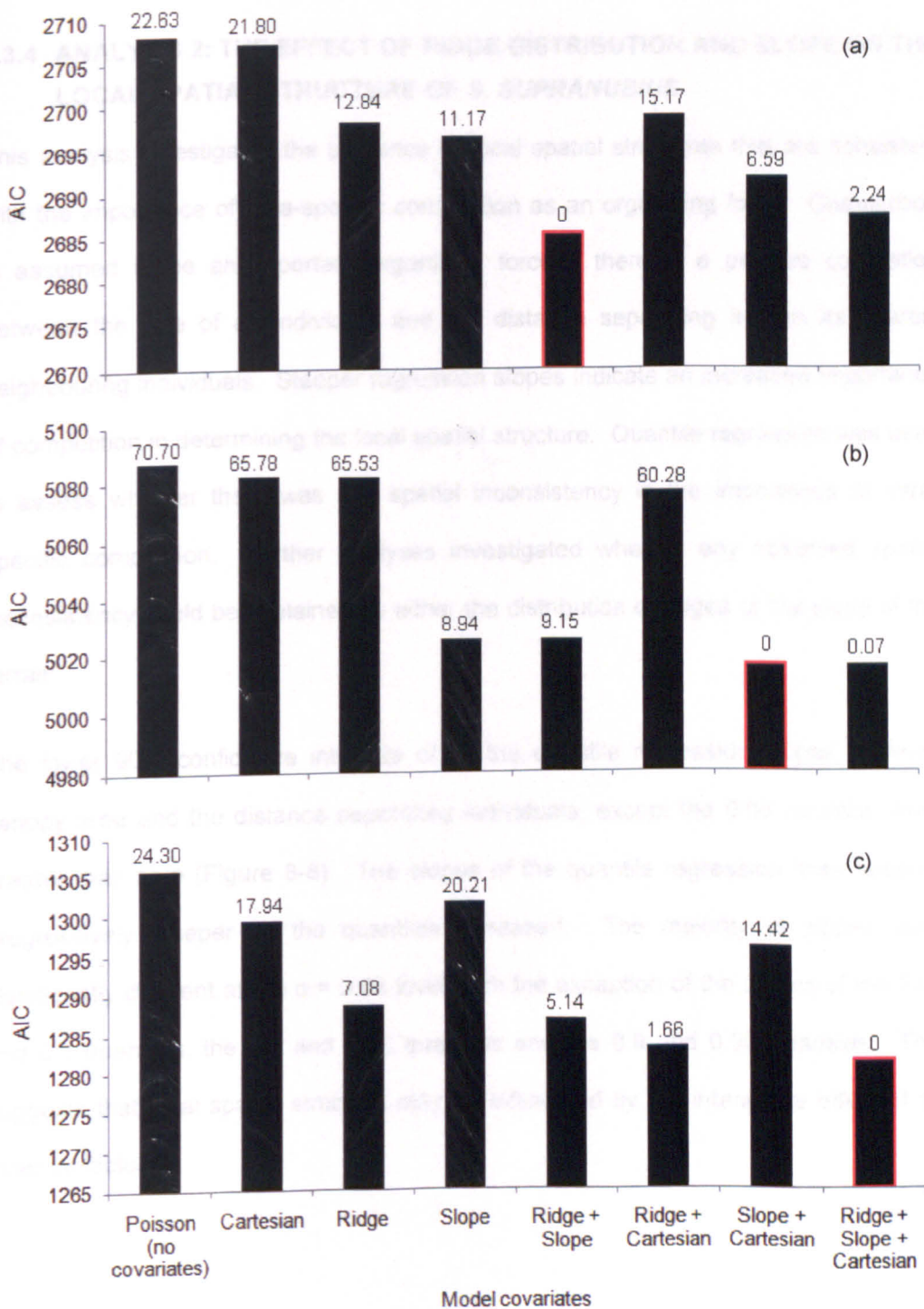


Figure 8-7 The AIC values of models containing the optimum Cartesian, ridge and slope effects (and combinations thereof) when fit to the distribution of (a) small individuals, (b) medium-sized individuals and, (c) large individuals. The difference in AIC values between each model and the best-fitting model (highlighted in red) is shown above each bar.

#### **8.3.4 ANALYSIS 2: THE EFFECT OF RIDGE DISTRIBUTION AND SLOPE ON THE LOCAL SPATIAL STRUCTURE OF *S. SUPRANUBIUS***

This analysis investigated the presence of local spatial structures that are consistent with the importance of intra-specific competition as an organising force. Competition is assumed to be an important organising force if there is a positive correlation between the size of an individual and the distance separating it from its nearest neighbouring individuals. Steeper regression slopes indicate an increased importance of competition in determining the local spatial structure. Quantile regression was used to assess whether there was any spatial inconsistency in the importance of intra-specific competition. Further analyses investigated whether any observed spatial inconsistency could be explained by either the distribution of ridges or the slope of the terrain.

The lower 90% confidence intervals of all the quantile regression slopes between canopy area and the distance separating individuals, except the 0.05 quantile, were greater than zero (Figure 8-8). The slopes of the quantile regression lines became progressively steeper as the quantiles increased. The majority of slopes were significantly different at the  $\alpha = 0.05$  level, with the exception of the slopes of the 0.05 and 0.1 quantiles, the 0.5 and 0.75 quantiles and the 0.9 and 0.95 quantiles. This suggests that local spatial structure may be influenced by the interactive effect of an external factor.

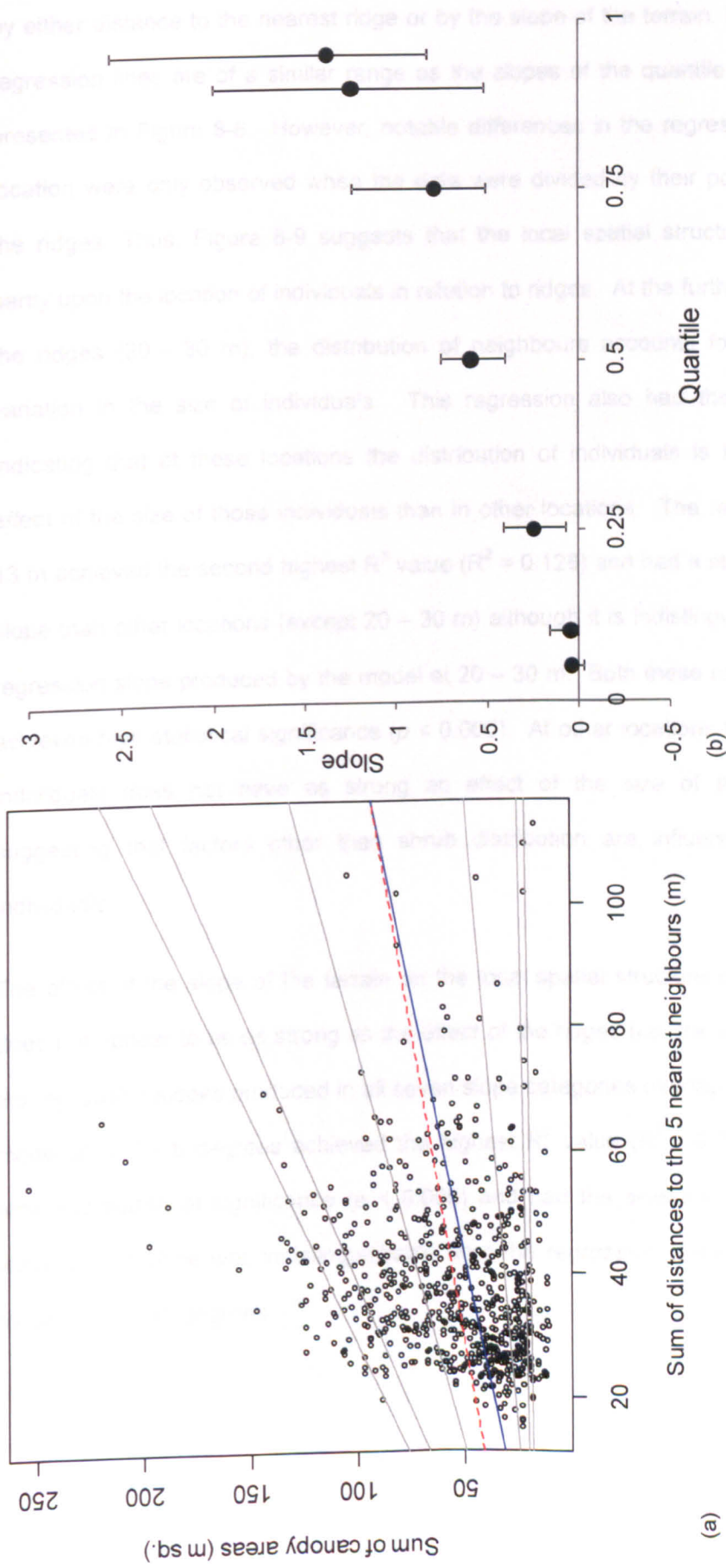


Figure 8-8 Relationship between the sum of the canopy areas of the focal shrub and the five nearest neighbours and the sum of the distances to the five nearest neighbours:

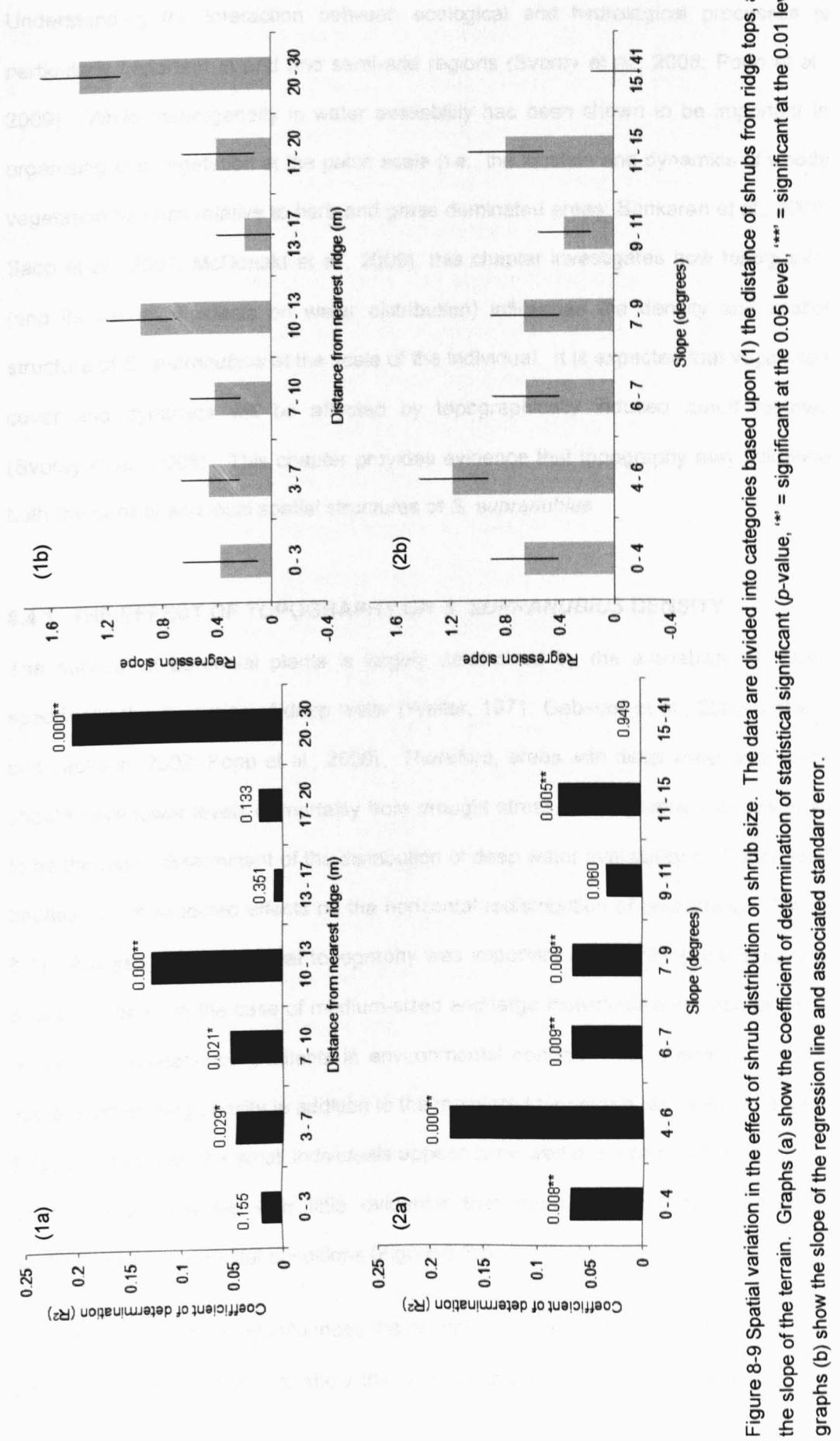
(a) Quantile regression estimates. The blue line shows a standard linear regression (least squares estimate of the mean effect), whereas the red line shows the position of the linear regression calculated at the median (the 0.5 quantile). The grey lines show the regression lines at the remaining quantiles (0.95, 0.9, 0.75, 0.25, 0.1, 0.05). As quantiles decrease the regression slopes get progressively shallower. (b) slopes of the regression lines in (a) (solid circles) with upper and lower 90% confidence intervals.

Figure 8-9 shows the results of the regressions between nearest neighbour distance and canopy size in different locations within the terrain. When the data are subdivided by either distance to the nearest ridge or by the slope of the terrain, the slopes of the regression lines are of a similar range as the slopes of the quantile regression lines presented in Figure 8-8. However, notable differences in the regression slopes with location were only observed when the data were divided by their position relative to the ridges. Thus, Figure 8-9 suggests that the local spatial structure may depend partly upon the location of individuals in relation to ridges. At the furthest location from the ridges (20 – 30 m), the distribution of neighbours accounts for c. 20% of the variation in the size of individuals. This regression also had the steepest slope indicating that at these locations the distribution of individuals is having a greater effect of the size of those individuals than in other locations. The regression at 10 – 13 m achieved the second highest  $R^2$  value ( $R^2 = 0.128$ ) and had a steeper regression slope than other locations (except 20 – 30 m) although it is indistinguishable from the regression slope produced by the model at 20 – 30 m. Both these regression models achieved high statistical significance ( $p < 0.005$ ). At other locations the distribution of individuals does not have as strong an effect of the size of those individuals, suggesting that factors other than shrub distribution are influencing the size of individuals.

The effect of the slope of the terrain on the local spatial structure of *S. supranubius* does not appear to be as strong as the effect of the ridges (i.e. the standard errors of the regression slopes produced in all seven slope categories overlap). The regression model at c. 4 - 6 degrees achieved the highest  $R^2$  value ( $R^2 = 0.183$ ). This model achieved statistical significance ( $p < 0.005$ ) and had the steepest regression slope, although this slope was indistinguishable from the regression slope produced by the model at 17 – 20 degrees.



## 8.4 DISCUSSION





## 8.4 DISCUSSION

Understanding the interaction between ecological and hydrological processes is particularly important in arid and semi-arid regions (Svoray et al., 2008; Popp et al., 2009). While heterogeneity in water availability has been shown to be important in organising arid vegetation at the patch scale (i.e., the location and dynamics of woody vegetation patches relative to herb and grass dominated areas; Sankaran et al., 2005; Saco et al., 2007; McDonald et al., 2009), this chapter investigates how topography (and its assumed effects on water distribution) influences the density and spatial structure of *S. supranubius* at the scale of the individual. It is expected that vegetation cover and dynamics will be affected by topographically induced runoff regimes (Svoray et al., 2008). This chapter provides evidence that topography may influence both the density and local spatial structures of *S. supranubius*.

### 8.4.1 THE EFFECT OF TOPOGRAPHY ON *S. SUPRANUBIUS* DENSITY

The survival of perennial plants is largely determined by the availability of water, specifically the dynamics of deep water (Walter, 1971; Gebauer et al., 2002; Schenk and Jackson, 2002; Popp et al., 2009). Therefore, areas with deep water availability should have lower levels of mortality from drought stress. Topography was assumed to be the major determinant of the distribution of deep water availability on Substrate 3 because of its expected effects on the horizontal redistribution of precipitation (Figure 8-1). Analysis 1 revealed that topography was important in describing the density of *S. supranubius*. In the case of medium-sized and large individuals there appear to be additional, unmeasured gradients in environmental condition (i.e., Cartesian effects) that are influencing density in addition to the measured topographical variables (Figure 8-7b, c). However, the small individuals appear to be well described by the measured topographical variables with little evidence that there are additional, important gradients in environmental conditions (Figure 8-7a).

Assuming that topography influences the distribution of deep water, this research may support previous studies that show that topographically driven water availability can

affect the density and biomass of arid vegetation (Imeson and Prinsen, 2004; Ju et al., 2008; Hamerlynck and McAuliffe, 2008; Svoray et al., 2008; Popp et al., 2009). The effect of the topographical variables on the density of *S. supranubius* individuals of all size classes is spatially non-linear, with abrupt changes in density associated with certain positions within the landscape. However, the importance of the distance class furthest from the ridges ( $> 28$  m) is questioned as only c. 3 % of the study site falls into this category. The density of both the medium-sized and large individuals was lower at locations close to the ridges, and increased at greater distances. This interpretation is supported by the results for the slope covariate which suggest that the density of both the small and medium-sized individuals was greater on shallow slopes ( $< 16$  and  $18$  degrees respectively). A reduction in shrub abundance on steep slopes corresponds with the outcomes of simulation models applied by Popp et al. (2009). This could correspond with the expectations of the water redistribution model (Section 8.1.1) which assumes that the steep slope and rocky composition of the ridges will reduce the chance of infiltration and thus reduce water availability in this location, resulting in a lower density of *S. supranubius*. It is also noted, although not studied, that the physical composition of the ridges could have an effect on *S. supranubius* individuals. Studies have shown that the presence of isolated large rocks may facilitate the establishment and growth of cacti and other desert plants by concentrating moisture (Peters et al., 2008). In the current focal system, however, rocks do not usually exist as discrete entities but as part of a solid ridge complex. Limited soil availability in these locations (pers. obs.) could reduce the successful establishment of *S. supranubius* and thus reduce their density in this area.

Interestingly, the influence of the two topographical variables considered seems to vary according to the size of the *S. supranubius* individuals. Whereas the density of both small and medium-sized individuals was influenced by the slope of the terrain, the distribution of ridges had a greater influence on the density of large individuals (Analysis 1c). These results suggest that the dynamics of *S. supranubius* populations may be influenced by multiple abiotic factors, and that the relative importance of these factors may vary with life stage.

#### **8.4.2 THE EFFECT OF TOPOGRAPHY ON *S. SUPRANUBIUS* SPATIAL STRUCTURE**

A positive relationship between the size of individuals and the distances separating them is consistent with the operation of competition as a structuring force. Therefore, the positive regression slopes identified in Analysis 2 (Figure 8-8) provide further evidence of the importance of competition in driving *S. supranubius* spatial structure, as suggested in Chapter 7. However, differences in the quantile regression slopes suggest that the importance of competition is not consistent throughout the population. The steep slopes in the upper quantiles (e.g., the 95<sup>th</sup> and 90<sup>th</sup> percentile) indicate that competition is an important structuring force in some locations. However, the less steep regression slopes in the lower quantiles suggests that in other locations additional factors may be over-riding the influence of competition as a structuring force. This corresponds with previous work by Shackleton (2002), who used nearest-neighbour techniques to show that the importance of both intra- and inter-specific competition between woody vegetation in the African savannah varied between sites. Shackleton (2002) observed that sites lacking evidence of intra-specific competition were at lower slope positions.

Analysis 2 detected patterns that are consistent with spatial variation in the effect of the relative distribution of individuals on the canopy size of those individuals (Figure 8-9). In locations that are between 20 and 30 m, and 10 – 13 m away from a ridge the distance separating individuals has a stronger effect on the size of those individuals than observed in other locations. These patterns are consistent with stronger competitive interactions in these locations. Although competition may still be occurring in the other locations (i.e., the regression slope is still positive), the effect of the distances separating individuals on the size of those individuals is not as strong indicating that factors other than competition may be influencing the local size-distribution. The implication of this interpretation is that while spatial environmental variation may not influence or alter the biological processes operating, it may influence their importance and thus their impact upon population spatial structure.

The results of Analysis 2 are consistent with the presence of terrain-induced alternating zones of competition importance. This interpretation would coincide with the predictions of the water redistribution model presented in Section 8.1.1. The stress-gradient hypothesis (Bertness and Callaway, 1994) predicts that under severe abiotic conditions, the importance of competition should decrease and facilitation should become the dominant structuring force. Based upon this hypothesis it is possible that the apparent spatial variation in the structuring force of competition indicated by Analysis 2 may be driven by topographically-induced spatial variation in resource (i.e. water) availability. It is important to note that this assertion is speculative and would require the direct measurement, or modelling, of water distribution and availability to substantiate the claims. It is also important to acknowledge the possible alternative explanations. For instance, Analysis 2 provided little evidence for the structuring force of competition close to the ridges. Under Bertness and Callaway's (1994) hypothesis this could be an area experiencing severe abiotic conditions (i.e. low water availability). However, it could also be explained by the physical properties of the ridges which may prevent the root systems of neighbouring individuals from overlapping. Consequently, competition for below-ground water resources will have a minimal effect on the local structure of the population.

Despite the above interpretation, it is noted that in all locations the ability of the distribution of individuals to explain the size of those individuals was low ( $R^2 < 0.25$ ; Analysis 2). This is at best equivalent to, or lower than, the results observed when applying similar techniques to woody vegetation in South African Savannas (Briones, et al., 1996; Shackleton, 2002) and temperate forests (Getzin et al., 2006; Getzin and Wiegand, 2007). This suggests that intra-specific competition is not the only force driving the structure of the *S. supranubius* population. Although the results presented in Figure 8-8 Graph b and Figure 8-9 Graph 1b are consistent with topographically-induced spatial variation in the importance of competition, it is important to consider other possible alternatives for the observed heterogeneity in the effect of shrub distribution on the size of those shrubs (Figure 8-8). Two alternative explanations are

provided which may help account for the increased heterogeneity of canopy sizes when individuals are widely spaced (i.e. the right hand side of Figure 8-8a). (1) Large distances between *S. supranubius* individuals may be caused by the presence of other species. If so, and these other species are competing with *S. supranubius* individuals, the size of those *S. supranubius* individuals may be smaller than expected given their separation. (2) If a large *S. supranubius* individual, separated from neighbouring individuals by large distances, undergoes vegetative reproduction to produce few (i.e. one or two) ramets, the local population will remain similarly distributed (i.e. wide inter-shrub distances) but the average canopy size will have decreased.

Ecological systems are typically complex, and it is likely that the spatial structure and dynamics of *S. supranubius* populations are influenced by a range of processes. Determining the relative importance of alternative processes will require further field observations and/or the application of mathematical simulation models. The results presented in this chapter are consistent with abiotic variation over short gradients influencing the spatial structure of arid shrub populations. The implication of this is that even when investigating spatial pattern and process in small plots (as are commonly used; Chapter 1, Appendix A), studies should consider the potential influence of short gradients in environmental variation on the patterns and processes being investigated.

### **The importance of topographically driven water availability**

The research presented in this chapter suggests that topography may affect the dynamics of woody vegetation perhaps through its effects on water availability. However, it is noted that additional, unobserved abiotic factors may be influencing the dynamics of *S. supranubius*. A review of empirical studies by Grayson et al. (2006) concluded that terrain properties rarely account for more than 50% of the variation in soil water availability. Similarly, Wilcox et al. (2006b) found that overland runoff only contributed to a very small part of the water budget in an arid shrubland, and only occurred during extraordinary precipitation events. It is likely, therefore, that factors

other than topography are important in driving the spatial distribution of water availability in the focal site. One potential factor that may influence the spatial distribution of water availability is the feedback between vegetation patches and runoff. Following a review of their own data, and data from other studies, Ludwig et al. (2005) concluded that vegetation patches in arid systems could obstruct and store more water than inter-patch areas. It is feasible that large *S. supranubius* individuals in particular could interact with topographically induced runoff to generate complex spatial variation in the availability of water.

To date, almost all studies of arid vegetation pattern and its relation to abiotic effects have been conducted in warm deserts. In contrast to many deserts, a large proportion of the precipitation in the Las Cañadas caldera falls as snow. Combined with the reduced potential evapotranspiration during winter months, snow melt can be more effective at recharging soil water than rainfall, per unit precipitation (Loik et al., 2004). A greater understanding of the dynamics and horizontal redistribution of snow melt may be required to understand the dynamics of vegetation in cold deserts. For example, if snow melt is slow, the majority of water will be infiltrated *in situ* and there may be limited spatial variation in deep water recharge. Under these conditions differences in water availability may not be as important in influencing vegetation biomass in cold deserts, increasing the relative importance of other effects, such as the physical effects of the substrate.

## **8.5 CONCLUSIONS**

In this chapter spatial variation in the density and spatial structure of an *S. supranubius* population was investigated and related to topographical features. The density of individuals was partly determined by a spatially non-linear response to both the slope of the terrain and the distribution of the ridges. The results suggest that multiple abiotic factors may influence the population dynamics of arid shrubs, and that the relative importance of abiotic factors may depend upon the life stage of the individual.

Competition is commonly observed in arid shrub communities. This chapter provides evidence that the importance of intra-specific competition as an organising force may vary spatially. In this chapter it is recommended that, even when investigating spatial pattern and process in small plots, studies of density-dependent effects should account for spatial environmental variation over short gradients. It is also recommended that further attention is paid to the horizontal redistribution of precipitation events (especially when precipitation is frozen) and its effects of vegetation dynamics.

## CHAPTER 9: DISCUSSION AND CONCLUSIONS

---

This thesis has addressed a number of methodological and ecological hypotheses. The main aim of the research was to investigate the potential methodological constraints of spatial point pattern analysis, as currently applied in the literature, and how, with the support of remotely sensed data, their application could be improved to help understand the biotic and abiotic processes structuring populations of *Spartocytisus supranubius*. As the results of each chapter have been evaluated in their specific context, this concluding chapter highlights only the key findings and major implications. Figure 9-1 provides a summary of the key findings, the implications of the research, methodological recommendations for the application of spatial pattern analyses, and future research questions. Section 9.1 discusses the utility of remote sensing in arid shrub ecology. The following sections discuss the conclusions of the research in relation to the research hypotheses detailed in Chapter 1 (Figure 1-1):

*Section 9.2: Spatial pattern analysis: reproofs and recommendations  
(Hypotheses 1 and 2)*

*Section 9.3: The biotic processes driving arid shrub population dynamics  
(Hypothesis 3)*

*Section 9.4: The abiotic processes driving arid shrub population dynamics  
(Hypothesis 4)*

The final section of the thesis (Section 9.5) makes a critical evaluation of the thesis. This is subdivided into two parts. The first part (Section 9.5.1) critically evaluates the assumptions made throughout the thesis and the potential implications these have for the reported results. The second part (Section 9.5.2) considers the potential role of other factors (i.e. alternative biotic and abiotic processes) on the observed *S. supranubius* spatial structure.



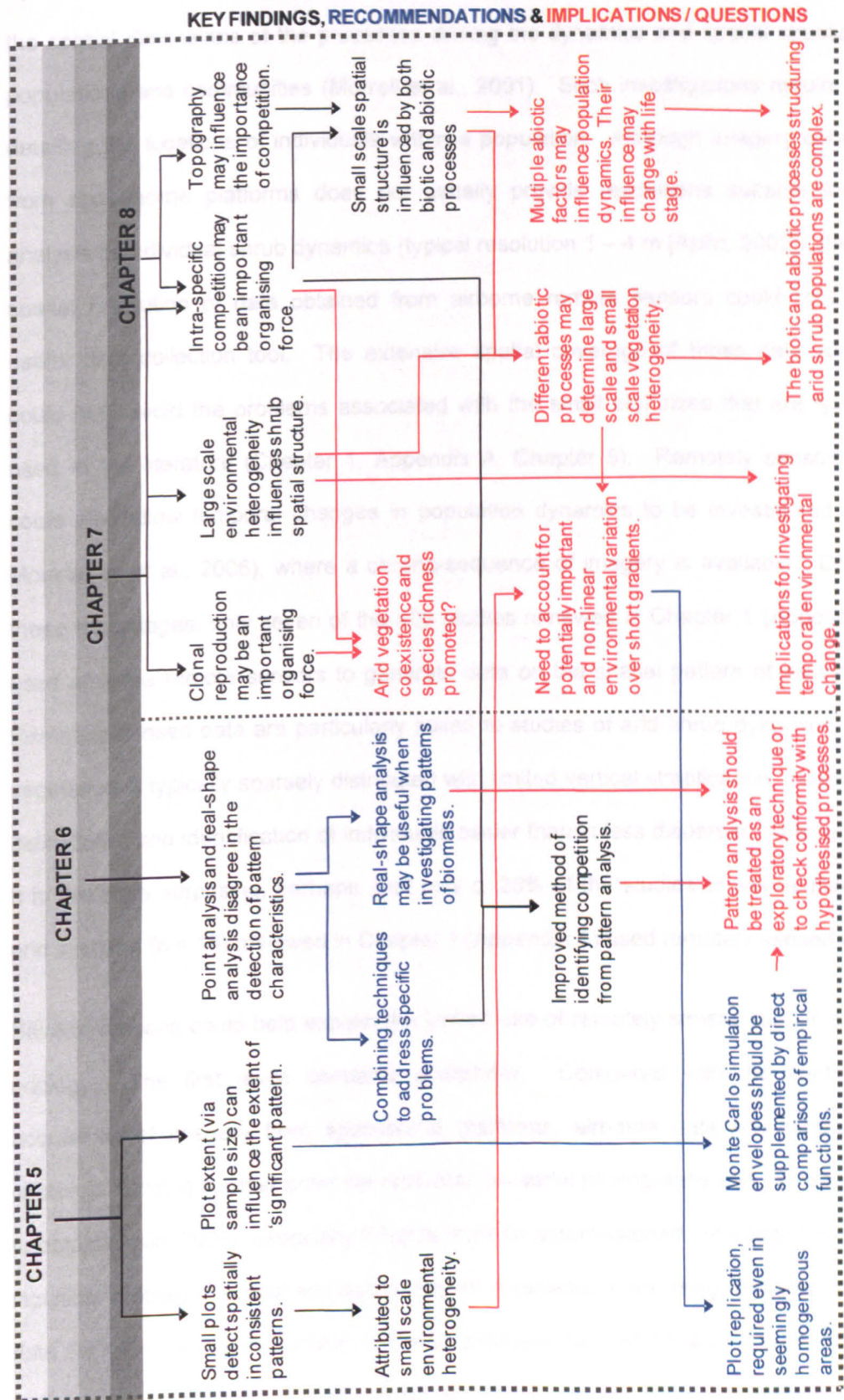


Figure 9-1 Summary of the key findings of the research, methodological recommendations, wider implications and resulting research questions

## **9.1 THE UTILITY OF REMOTE SENSING IN ARID SHRUB ECOLOGY**

Spatial ecology is a specialisation of geography and ecology that aims to understand the spatial dimensions of the processes driving the dynamics and spatial structure of populations and communities (Murrell et al., 2001). Such investigations require data detailing the locations of individuals within a population. Although imagery collected from spaceborne platforms does not usually provide resolutions suitable for the analysis of individual shrub dynamics (typical resolution 1 – 4 m [Aplin, 2005]), the fine spatial resolution of data obtained from airborne remote sensors could provide a useful data collection tool. The extensive spatial coverage of these data sources could help avoid the problems associated with the small plot sizes that are typically used in the literature (Chapter 1, Appendix A, Chapter 5). Remotely sensed data could also allow temporal changes in population dynamics to be investigated (e.g., Moustakas et al., 2006), where a chrono-sequence of imagery is available. Despite these advantages, only seven of the 109 studies reviewed in Chapter 1 (Appendix A) used airborne remote sensors to generate data on the spatial pattern of individuals. Remotely sensed data are particularly suited to studies of arid shrub dynamics. Arid vegetation is typically sparsely distributed with limited vertical stratification, making the delimitation and identification of individuals easier than in less dispersed communities. It is therefore surprising, perhaps, that only c. 20% of the studies of woody plants in arid systems (n = 18) reviewed in Chapter 1 (Appendix A) used remotely sensed data.

Several reasons could help explain the limited use of remotely sensed data in spatial ecology. The first issue concerns availability. Compared with the continuous acquisition of imagery from spaceborne platforms, airborne data are infrequently captured. Cost is another potential restriction as aerial photography can be expensive to obtain (Aplin, 2005), especially if flights must be commissioned. In addition to these logistical problems, spatial ecologists may be dissuaded from using remotely sensed data for other reasons. Unlike manually collected data which are typically of high accuracy, remotely sensed data can contain errors. A common error when mapping individual trees or shrubs is to incorrectly classify two or more neighbouring individuals

with overlapping or adjacent canopies as a single individual (Moustakas et al., 2008). Remotely sensed data can also suffer from errors of omission (e.g., failing to identify an *S. supranubius* individual) and commission (e.g., incorrectly identifying a different species or object as an *S. supranubius* individual). Typically, the primary objective is to minimise errors of commission (i.e., the set of objects identified as *S. supranubius* individuals has a high probability of being *S. supranubius* individuals; Atkinson et al., 2007). Although errors of commission should be infrequent in sparse, arid communities, they have the potential to distort the detection and interpretation of pattern (Freeman and Ford, 2002). Using simulated data, Atkinson et al. (2007) showed that commission errors could affect the strength of the pattern detected by  $g(r)$  (i.e., the position of the empirical  $g(r)$  in relation to simulation envelopes). Commission errors in this thesis ranged from 8% (on Substrate 3) to 16% (on Substrate 5), and could therefore affect the pattern detected when using Monte Carlo simulation envelopes (Atkinson et al., 2007). To avoid inaccurate interpretations of pattern it is recommended that attention is also paid to the magnitude of the empirical functions as a measure of pattern strength (Chapter 7, Section 9.2.2). Further research is needed to determine how greatly errors of commission and omission influence the quantification of pattern from real (i.e., not simulated) data, and whether this effect could be minimised by increases in plot extent. If large plot extents can successfully average the effects of occasional and random classification errors on pattern detection, then remote sensing may provide a realistic alternative to manual data collection.

Errors in remotely sensed data may also occur from the necessary approximation of rooting points as the centre of the canopy, although such errors may be avoided by preserving the size and shape of individuals during analysis (Chapter 6). Another problem that may deter spatial ecologists from using remote sensing data is the restriction of subsequent analyses of pattern and process to individuals above a certain size (usually dictated by the pixel resolution of the image). This may prevent studies that use remotely sensed data from investigating the dynamics of dispersal and establishment.

Despite these potential problems, remotely sensed data offer a valuable, yet underused, resource in studies of arid shrub dynamics. Hyperspectral and LiDAR data could provide information on the characteristics and spatial variation in abiotic conditions such as the texture and moisture properties of soil (Anderson and Croft, 2009). These data could be used to investigate the relationship between abiotic conditions and biotic patterns and processes. A closer collaboration between remote sensors and spatial ecologists may provide exciting opportunities for research (Newton et al., 2009), specifically in our understanding of the interaction between biotic processes and abiotic conditions and their spatio-temporal dynamics.

## **9.2 SPATIAL POINT PATTERN ANALYSIS: REPROOFS AND RECOMMENDATIONS**

The spatial pattern of individuals within a population may provide valuable insights into the biotic and abiotic processes driving its dynamics. However, some authors have contended that the analysis of pattern alone is not enough to infer underlying processes (Mahdi and Law, 1987; Borcard et al., 2004; Schurr et al., 2004), whilst others have condemned the lack of empirical verification of the spatio-temporal theory upon which pattern–process inference is based (Murrell et al., 2001; Perry et al., 2006). However, when the demographics of the focal species are very slow, such as arid shrubs, it may not be feasible to empirically assess the operation of biological processes, especially when the focus of the study is on population-level dynamics (see Section 1.3.2). Until datasets of sufficient spatial and temporal coverage are generated, pattern–process inference may be one of the few techniques available to investigate arid shrub population dynamics. Therefore, we must ensure the inferential link between pattern and process is as strong as possible. The wide availability of spatial point pattern analysis techniques has led to a sharp increase in their application over the last 15 years (Chapter 1, Appendix A). However, our ability to infer processes from patterns is being impeded by the methodological procedures being used. Chapters 5 and 6 investigated the robustness of pattern detection to changes in plot extent and data representation.

The strength of pattern–process inference can be improved by using deductive rather than inductive reasoning; i.e., using ecological theory and knowledge of the focal system and species to formulate precise *a priori* hypotheses of the likely abiotic and biotic processes of importance and their expected spatial signatures (Chapter 7; McIntire and Fajardo, 2009). This approach can be extended by using point process theory to test and explore predictions. Chapter 8 used point process modelling techniques to investigate the influence of measured and unmeasured spatial covariates on the spatial structure of a *S. supranubius* population.

Based on the review of ecological spatial pattern analysis studies (Chapter 1, Appendix A) and the results of Chapters 5, 6 and 8, the following sections provide recommendations for future studies using spatial pattern analyses to infer population dynamics.

### **9.2.1 PLOT EXTENT AND REPLICATION**

In Chapter 1 (and Appendix A) it was noted that studies investigating the spatial patterns of woody species typically used small plot extents. Analyses in Chapter 5 demonstrated that the patterns detected by commonly used plot extents (< 1 ha) are spatially inconsistent, and that inconsistency was greatest at the scales at which biological interactions are presumed to occur. Furthermore, spatially inconsistent patterns were detected on both homogeneous and heterogeneous substrates. These results have implications for the utility of spatial pattern analyses as currently applied in the literature. Specifically, the patterns detected and the processes inferred from small plots may not be representative of the population as a whole. This is of particular concern when few or single plot replicates are used, as is common in the contemporary literature (Chapter 1, Appendix A). In Chapter 5, the spatial inconsistency in pattern detection was attributed to unobserved, small-scale environmental heterogeneity. Chapter 8 provides further weight to this argument by detecting a change in population spatial structure associated with local topography. This chapter concluded that small-scale environmental variability could have (spatially) non-linear effects on the local spatial structure of *S. supranubius* populations.

Therefore, it is important that the design of future studies into plant spatial patterns consider the potential importance of both long (Chapter 7) and short (Chapters 5 and 8) gradients in environmental conditions. Based on my findings, I recommend that one of two approaches is taken: either use plots with an extent larger than any anticipated environmental heterogeneity effects, or use multiple replicate plots. Both techniques allow for any effects of small-scale heterogeneity to be averaged; however, as environmental variables of importance may not always be known, the second approach is preferred. Indeed, using a single large plot may obscure the effect of important environmental gradients. If multiple plots can be used, each of a size large enough to provide a reasonable estimation of pattern (see Chapter 5), analyses may be able to start investigating which external abiotic or biotic (i.e., other species) factors are affecting the dynamics of the focal species. However, more research may be required if the results of multiple plots are to be combined into a single pattern statistic, as methods for the analysis of replicated point patterns remain relatively under-developed and untested (Diggle et al., 2000; Bell and Grunwald, 2004; Illian et al., 2008).

It is important to note that an alternative explanation may account for the increased inconsistency of  $g(r)$  at small plot extents. As noted in Chapter 5, changes in plot extent are inherently linked with the issue of sample size. Consequently, the inconsistency in  $g(r)$  at small plot extents may also be attributable to small sample sizes and associated reductions in statistical power. There is no consensus in the literature on the sample size requirements for either  $g(r)$  or  $L(r)$  (Chapter 1), and further research is needed to determine at which sample sizes the application of these indices becomes unreliable. That said, the results of Chapter 5 clearly demonstrate an increase in inconsistency with decreases in plot extent while sample sizes are high (i.e.  $n > 100$ ). Further research is needed to disassociate the effect of small-scale heterogeneity and low sample sizes at the smallest extents (i.e.  $< 1$  ha).

### **9.2.2 ASSESSING PATTERN USING MONTE CARLO SIMULATION ENVELOPES**

As sample size increases (e.g., with increases in plot extent), the width of envelopes generated from multiple Monte Carlo simulations of a hypothesised null model decreases. Changes in the width of the Monte Carlo simulation envelopes with variation in plot extent can affect the detection of 'significant' pattern (Chapter 5). This adds weight to previous criticisms that the construction of Monte Carlo simulation envelopes from the result of many simulated patterns underestimates the Type I error rate and is consequently invalid for inferring pattern significance (Loosmore and Ford, 2006). Despite these limitations, however, the majority of studies use Monte Carlo simulation envelopes to detect pattern and infer processes (Chapter 1, Appendix A), perhaps because the main authors in the field continue to recommend their usage (e.g. Diggle, 2003; Wiegand and Moloney, 2004; Illian et al., 2008). In Chapter 5 I recommend that Monte Carlo envelopes are treated as an analogue to statistical assessment via p-values; i.e., the position of the empirical function in relation to the simulation envelopes provides an indication of the importance of the pattern. Greater attention should also be paid to the magnitude of the empirical function relative to the null model expectation; an analogue of effect size.

### **9.2.3 POINT PROCESS MODELLING**

In the majority of studies, analyses of spatial point processes focus on tests of deviation from complete spatial randomness. While such analyses can provide evidence for the occurrence of a hypothesised process (Chapter 7), they are unable to explicitly test processes, and therefore may have limited applicability (Comas and Mateu, 2007). Point process modelling techniques, however, allow the effect of hypothesised abiotic and biotic processes and mechanisms on point patterns to be more rigorously tested and explored. The scope of possible models is very wide and may include spatial (Cartesian) trends, dependence upon measured covariates (Chapter 8), interpoint interactions, and dependence on marks (i.e., a categorical [e.g., species] or continuous [e.g., size] value assigned to each point; Baddeley and Turner, 2005). Despite the relatively long history of point process theory, its application in

studies of plant population dynamics remains limited (Comas and Mateu, 2007). While studies are beginning to use non-Poisson point process models such as Neyman-Scott processes (e.g., Thomas processes or Matérn cluster processes [e.g., Wiegand et al., 2007b; Yu et al., 2009]) to test and explore interactions between points, relatively few studies have used point process modelling techniques to explore the effects of abiotic gradients on either first- or second-order population spatial structure (Chapter 8; but see Shimatani and Kubota, 2004). This technique provides a promising technique to infer the influence of abiotic gradients (either measured or assumed) on population-scale structure. Such techniques could be essential for improving our understanding of how biotic and abiotic processes interact to drive population dynamics, which remains a fundamental question in ecology (Dahlgren and Ehrlén, 2009).

### **9.3 THE BIOTIC PROCESSES DRIVING ARID SHRUB POPULATION DYNAMICS**

One of the aims of plant population and community ecology is to investigate biotic interactions and predict how these translate into consequences for the whole population/community (Freckleton et al., 2009). Shrubs are important, often dominant, elements of arid and semi-arid vegetation communities. Understanding the dynamics of the dominant shrub species may be an important first step in understanding the dynamics of the ecosystem as a whole, yet little is known about their population processes and the factors underlying their dynamics (Kyncl et al., 2006; Jiménez-Lobato and Valverde, 2006). The spatial structures of *S. supranubius* populations are consistent with the operation of clonal reproduction and intra-specific competition (Chapter 7). There is also evidence that the biotic processes driving the dynamics of *S. supranubius* may be influenced by spatial environmental heterogeneity (see Section 9.4).

#### **9.3.1 CLONAL REPRODUCTION**

The spatial patterns quantified in Chapter 7 are consistent with the clonal reproduction of *S. supranubius*. It is therefore reasonable to assume that clonal reproduction is an



important process organising *S. supranubius* populations. Asexual reproduction is assumed to be favoured in marginal or harsh environments (Peck et al., 1998; Klimeš, 2008). In these locations sexual reproduction may be hampered, and so clonal reproduction enables the population to persist in the short and mid-term (Mandujano et al., 2001; Honnay and Bossyut, 2005; Wesch et al., 2005). As such, clonal growth can greatly increase the resilience of a plant population and, in the case of a keystone species, the whole community (Wesche et al., 2005). In the long-term, however, prolonged clonal reproduction can have implications for the future viability of a population (Honnay and Bossyut, 2005; Honnay et al., 2006).

Clonal reproduction can influence the spatial structure and competitive interactions in plant communities (Song et al., 2002). *S. supranubius* follows a phalanx clonal growth form: i.e., ramets form consolidated groups compared to the spreading, widely spaced characteristics of ramets typical of guerrilla clonal growth (Ye et al., 2006). The phalanx growth form increases the frequency of *S. supranubius* intra-ramet contacts and may in part explain the importance of intra-specific competition posited in Chapters 7 and 8. By ensuring close proximity and strong intra-specific competition among *S. supranubius* individuals (as observed on the homogeneous substrates, Chapter 7), clonal reproduction could have implications for the structure of the Cañadas vegetation community. Additionally, the senescence of *S. supranubius* individuals following strong intra-specific competition may leave islands of fertility which could be exploited by other species (Alvarez et al., 2009).

Clonal reproduction is believed to be a common attribute of arid shrubs (Schenk, 1999), and has been demonstrated in several species of tree and shrub in the cold deserts of central Asia (Bruehlheide et al., 2003; Qong et al., 2002; Song et al., 2002; Wesche et al., 2005). The presence of clonal species, especially when those species dominate the community, as *S. supranubius* does, could be an important mechanism driving community dynamics. However, the importance of this strategy in the dynamics of arid shrub populations, and the implications for the structure and dynamics of the wider vegetation community, has received little attention.

The use of deductive reasoning provides confidence in the interpretation of the observed spatial patterns as a signature of clonal reproduction. However, it is important to evaluate the other possible explanations. Scholes and Archer (1997; cited in Meyer et al., 2008) suggest that aggregated spatial patterns in savanna plants may be generated by topography (e.g. termite mounds), fire patchiness and soil depth. However, field observations did not indicate the presence of any topographical variation at the scales required to generate the aggregation observed (i.e. 3 - 10 m), and there is no documented evidence of fires within the caldera. Although soil depth is a possible explanation it is believed to be unlikely because of the remarkable consistency in aggregation between the substrates (Chapter 7). It is unclear how spatial variation in soil depth would be consistent between substrates that were generated at different times and have different formations (e.g. aa and pahoehoe lava flows).

### **9.3.2 INTRA-SPECIFIC COMPETITION**

Many studies have investigated the interactions between different functional groups in arid systems (e.g., shrubs, herbs and grasses; Maestre et al., 2003; Armas and Pugnaire, 2005; Anthelme and Michalet, 2009), but relatively little has been concluded about the interactions between the dominant woody components of arid systems. Bertness and Callaway's (1994) stress-gradient hypothesis proposes that facilitation should dominate species interactions in systems where extreme abiotic stress limits productivity, such as arid environments (Weedon and Facelli, 2009). Over the last 15 years there has been much debate about the relative importance of competition and facilitation in structuring arid vegetation communities, leading to the generality of Bertness and Callaway's (1994) model being questioned (Michalet, 2006, 2007; Lortie and Callaway, 2006; Maestre et al., 2005, 2009). The results presented in Chapters 7 and 8 are consistent with the operation of intra-specific competition as an important process driving the spatial structure of *S. supranubius* populations. This is in contrast to the predictions of Bertness and Callaway's (1994) stress-gradient hypothesis, perhaps suggesting that whereas facilitative interactions dominate inter-specific

interactions, intra-specific interactions are predominantly competitive. This interpretation is in contrast to a simulation study which found that neither facilitation nor competition were as important as random mortality from drought in structuring an intra-specific arid shrub population (Malkinson and Jeltsch, 2007). Notably, however, the model developed by Malkinson and Jeltsch (2007) did not incorporate clonal reproduction, which may explain the absence of strong intra-specific competition (see Section 9.3.1).

The importance of intra-specific competition in structuring *S. supranubius* populations supports previous assertions that when the limiting factor is a scarce, depletable resource, such as water, facilitation will be reduced. Under these conditions, facilitation will only occur when neighbours can increase the absolute volume of the resource (water) beyond their own requirements (Maestre and Cortina, 2004; Maestre et al., 2009). It is noted, however, that intra-specific competition alone does not fully explain the spatial structure of *S. supranubius* (Chapter 8), and that other factors (either biotic or abiotic) are important in driving the population's spatial structure.

The interpretation of intraspecific competition is given extra credence by the variety of techniques used to assess its presence: spatial structures in different size classes (Chapter 7), comparison between the hard-core distance and the maximum scale of aggregation (Chapter 7, see also 9.3.3) and size-distance regressions (Chapter 8). All three techniques produced results that were consistent with the operation of intra-specific competition. In response to criticisms that snap-shot spatial patterns cannot be used to infer processes (Mahdi and Law, 1987; Cale et al., 1989; Moravie and Robert, 2003; references in McIntire and Fajardo, 2009) it is recommended that practitioners utilise multiple approaches when investigating the presence and importance of a hypothesised process.

The following section considers the typical use of spatial pattern analysis to detect competition and makes recommendations for a more rigorous approach.

### 9.3.3 THE MYTH OF REGULARITY AND A NEW METHOD FOR DETECTING COMPETITION

Spatial pattern analysis is often used to deduce the presence, or absence, of competition in a plant population. When individuals are competing, the fitness of an individual is expected to be a function of the distance separating that individual from neighbouring individuals, and the size of those individuals. If individuals successfully outcompete all others within their zone of influence then a regular (also known as dispersed) spatial pattern will result. A regular pattern of individuals can, therefore, be considered as strong evidence for the importance of competition (Stoll and Bergius, 2005). However, there are very few documented examples of plant populations with regularly distributed individuals from any ecosystem. Indeed, regular distributions are most commonly observed in self-thinned, even-aged, single-species forest stands (Toft and Frazier, 2003). Even after removing the effects of heterogeneity (which is often cited as a cause of aggregated distributions), none of the *S. supranubius* size classes on any of the substrates displayed a regular distribution (Chapter 7).

The lack of a regular pattern does not necessarily indicate the absence of competition. Many recent articles consider a decrease in the strength of aggregation with age to be evidence of competition, even if no individual age/size class achieves regularity (e.g., Getzin et al., 2008; Metsaranta and Lieffers, 2008; Meyer et al., 2008; Gray and He, 2009). A decrease in the strength of aggregation of *S. supranubius* individuals with increases in size class was observed (Chapter 7). However, I suggest that merely investigating age/size-dependent changes in pattern is not enough to determine the existence, or otherwise, of competition. For instance, random (i.e., density-independent) mortality can have the same qualitative effect on the pattern of a population as density-dependent mortality (Toft and Frazier, 2003). Furthermore, using a spatially explicit individual-based model, Murrell et al. (2009) demonstrated that an increase in aggregation with size class may be consistent with a self-thinning process under certain conditions (i.e., slow growth, low fecundity or high juvenile–juvenile competition). Toft and Frazier (2003) suggest that in order to conclude the operation of competition, the pattern of individuals should become more dispersed

with increases in age/size than expected under the hypothesis of density-independent thinning. This approach would, however, be time-consuming; quantifying the pattern expected under density-independent thinning could only be achieved by repeatedly simulating the loss of individuals in a density-independent manner and calculating an average pattern expectation.

In Chapters 6 and 7 a more rigorous approach to detecting competition in sparse communities was introduced. The technique compares the scale of modal shrub–shrub separation within a cohort with the hard-core distance (i.e., maximum canopy diameter) below which interactions are not expected to occur. I suggest that a decrease in aggregation strength *and* an increase in the modal shrub–shrub distance (relative to the canopy diameter) with increases in age/size provides strong evidence for the operation of density-dependent thinning. Furthermore, this technique enables the relative strength of competitive forces in replicate populations to be compared, as a greater difference between the modal shrub–shrub distance and the hard-core distance indicates stronger competitive interactions. Although not tested in this thesis, this technique should be also applicable in populations with overlapping canopies where the modal shrub–shrub distance may be less than the average canopy diameter. Notably this technique is only possible if patterns are assessed using the pair-correlation function ( $g(r)$ ) which, unlike its cumulative counterpart ( $L(r)$ ), provides information on modal point–point separation. This provides further support for the application of this measure in favour of the more commonly applied cumulative measures ( $L(r)$  and  $K(r)$ ; Chapter 1, Appendix A).

#### **9.4 ABIOTIC PROCESSES DRIVING ARID SHRUB POPULATION DYNAMICS AND THEIR INTERACTION WITH BIOTIC PROCESSES**

Understanding how abiotic factors, especially water availability, influence biological processes is essential if we are to predict and manage the effects of future environmental and climatic change on arid vegetation dynamics (Snyder and Tartowski, 2006). Current efforts are being directed towards investigating (empirically and theoretically) how temporal variation in precipitation timing (Snyder et al., 2004;

West et al., 2007), magnitude (Huxman et al., 2004) and frequency (Heisler-White et al., 2009) may influence arid vegetation productivity. Comparatively little attention is being paid to the influence of spatial variation in water availability (Loik et al., 2004; Breshears et al., 2009).

Although no direct measurements of processes were made, the results presented in this thesis are consistent with the influence of spatial variation in abiotic conditions (particularly topography) on the spatial structure of *S. supranubius* populations. Analyses in Chapter 7 suggest that spatial environmental heterogeneity can influence the spatial patterns and demographics of *S. supranubius*. These analyses were extended in Chapter 8, the results of which are consistent with an influence of topography on both the density of *S. supranubius* and the importance of competitive interactions between individuals. The results of Chapter 8 provide support for the assertions of Monger and Bestelmeyer (2006) who suggested that the effect of geomorphological and topographical conditions on water and nutrient conditions can influence the dynamics of arid vegetation at a range of scales. Two possible explanations were proposed for the apparent spatial variation in the structuring force of competition: (1) the physical effect of pahoehoe ridges preventing the interaction of neighbouring root masses and reducing the effects of below-ground competition, and (2) topographically-induced spatial variation in resource (i.e. water) availability. The former explanation may also help explain why clonal reproduction appears to be less prevalent on the two pahoehoe lava flows (Substrates 3 and 4; Chapter 7).

In response to the latter explanation it is acknowledged that further research is needed to investigate how geomorphological and topographic conditions might interact with precipitation events to determine spatial variation in water availability. Currently, our understanding of these processes is limited to independent studies of the effects of geomorphic features (e.g., Pérez, 2003; Zou et al., 2010), and the effects of changes in precipitation conditions (e.g., Huxman et al., 2004; Snyder et al., 2004; West et al., 2007; Heisler-White et al., 2009). Studies of the latter are largely experimental whereas studies of both effects typically focus on the physiological response of

individual plants or the large-scale response of vegetation assemblages (i.e., biomass or productivity). Long-term datasets of water availability are needed to further our understanding of the factors that determine local water availability, the resulting natural spatio-temporal variation in water availability, and how it affects arid shrub population structure. It should also be noted that, in addition to topography, other unmeasured abiotic gradients may influence the distribution of *S. supranubius*.

Without further data the likelihood of either of the explanations suggested above cannot be assessed. Furthermore, the complexities of ecological systems mean that the spatial structure and dynamics of *S. supranubius* populations are likely to be influenced by a range of processes. As noted in Chapter 8, in addition to spatial variation in the importance of competition, the observed heterogeneity in the effect of shrub distribution on the size of those shrubs may also be influenced by the processes of clonal reproduction and interspecific competition. However, the results and speculations made raise some important questions and implications:

(1) While vertical heterogeneity in water availability has been shown to be important in determining patch dynamics (i.e., the location and dynamics of woody vegetation patches relative to herb and grass dominated areas; Sankaran et al., 2005), horizontal heterogeneity in water availability may be important in determining population spatial structure.

(2) The spatial structure and dynamics of *S. supranubius* and other arid shrub populations may be influenced by the physical effects of abiotic variables that are not temporally transient over ecologically meaningful timescales (i.e. physical geomorphological effects).

(3) Spatial variation in individual abiotic variables may have multiple effects on the dynamics of arid shrubs, and these effects may be life stage specific (Chapter 8).

(4) The effects of abiotic variables on arid shrub dynamics may be more complex than realised. It is generally accepted that the processes driving vegetation distribution vary with scale. It is a commonly held belief that whereas abiotic factors determine

vegetation heterogeneity at coarse, landscape scales, small-scale structure is determined by biotic processes (Stoyan and Penttinen, 2000; Wiegand et al., 2007b; Bisigato et al., 2009). However, the results presented in Chapter 7 and 8 suggest that spatial environmental heterogeneity (at a range of scales) may affect the fine-scale spatial structure of *S. supranubius* populations. The effect of abiotic variables on small-scale population structure can come from both long (Chapter 7) and short gradients in abiotic conditions (Chapters 5 and 8). Therefore, even when using small plots, studies of plant pattern and process should account for the potential effects of spatial environmental variation over short gradients. Research into population dynamics in all systems should place more emphasis on investigating the role of spatial environmental heterogeneity in determining plant pattern formation.

## **9.5 A CRITICAL EVALUATION OF THE THESIS**

### **9.5.1 THE MAJOR ASSUMPTIONS**

Throughout the thesis, two major assumptions were made. In both cases these assumptions were deemed necessary to enable further analysis and were encouraged by the precedent set in the literature. This section reviews each of the major assumptions, considers how the assumptions may influence the results reported in the thesis, and how these assumptions could be empirically tested.

#### **The age and size of *S. supranubius* individuals are positively correlated**

*Why and where the assumption was made:*

The relative importance of different biological processes (e.g. reproduction, competition) varies with the age of an individual. Therefore, if a researcher aims to use spatial patterns to infer the presence of multiple, age-specific processes, the spatial structure of individuals of different ages should be examined. Measuring the age of all individuals is typically unfeasible. Consequently, a common approach in the ecological literature is to use an individual's size as a proxy for its age, under the assumption that larger individuals will also be older (e.g. Meyer et al., 2008). This



assumption was utilised to define three size classes (small, medium-sized and large) which were assumed to represent cohorts of differing age. These size classes are used in the analyses found in Chapters 6 – 8.

*The potential implications of this assumption:*

Age and size are not necessarily related. For example, a young individual may grow very quickly, or an old individual that has experienced intense competition may have a low or minimal growth rate. Consequently, there may be occasions where individuals allocated to a size class (e.g. small) are not of the assumed age (e.g. young). In the context of the work presented in this thesis, deviations from this assumption would mean that the spatial signature of processes assumed to dominate at certain ages may be diluted by the presence of individuals of a differing age. However, because of the large number of individuals mapped (over 17,000) it is expected that the proportion of individuals incorrectly assigned to a size/age class would be very low and unlikely to adversely affect the detection of age-specific patterns and processes.

*How the assumption could be empirically tested:*

It is possible that this assumption could be empirically tested by taking cores of the main stem of individuals of a known size and counting growth rings to achieve an estimate of the individual's age. However, there are three issues that limit the use of this technique in the present system: (1) *S. supranubius* is a protected under regional legislation (Annex II of the Flora Order 20/02/1991). Under this protection it is illegal to remove any part of an *S. supranubius* individual, or knowingly disturb or destroy it, without governmental authorisation. (2) The dynamics of arid shrubs are event driven, and discernable growth may only occur in climatically favourable years. Consequently, *S. supranubius* individuals may not possess annual growth rings which would impede any assessment of their absolute age, although relative ages could still be assessed. (3) *S. supranubius* individuals typically have a dense canopy with little, and frequently no, gap between the canopy and the ground. This would physically

impede access being made for the purpose of coring without contravening Annex II of the Flora Order.

### **Large individuals persist in environmentally benign areas**

*Why and where the assumption was made:*

This assumption underlies the analyses found in Chapters 5 and 7. The assumption of spatial environmental homogeneity is an important limitation of the second order statistics used throughout this thesis ( $g(r)$  and  $L(r)$ ). If environmental conditions vary from one location to another (i.e. the environment is heterogeneous at a scale greater than the size of an individual shrub) the inference of biotic processes from the spatial pattern of individuals can be hindered because the distribution of plants may depend as much on the environmental template as on internal, biotic processes (Law et al., 2009). The potential effects of environmental heterogeneity must be removed if the spatial patterns driven by biotic processes are to be uncovered. To remove the effects of environmental heterogeneity, that heterogeneity must first be mapped. Without information on the types of environmental conditions that may influence the distribution of *S. supranubius*, this research utilised *S. supranubius* individuals themselves as indicators of habitat quality (following Stoyan and Penttinen, 2000; Getzin et al., 2008; Barbeito et al., 2009; Zhu et al., 2010). Specifically, it was assumed that very large *S. supranubius* individuals were either very old, or had experienced much higher growth rates than other, smaller individuals. It was consequently assumed that these individuals would be located in optimal habitat. Deviation of the distribution of very large individuals from randomness would imply that the distribution of optimal habitat was also not random and therefore that the environment was heterogeneous.

*The potential implications of the assumption:*

Individuals may attain large sizes for reasons other than being located in an environmentally benign area. For example, individuals that occupy a suboptimal habitat but experience little or no competition may reach large sizes. Conversely,

individuals in optimal habitats that undergo intense competition may not achieve large sizes. Such discrepancies may result in an inaccurate representation of the magnitude of environmental heterogeneity, although assessments of relative heterogeneity between sites should remain valid as any discrepancies are unlikely to be spatially systematic.

*How the assumption could be empirically tested:*

A more accurate assessment of the environmental heterogeneity of a site could be provided by directly mapping the abiotic factors of importance. However, such a technique would require in-depth knowledge of the abiotic variables influencing the dynamics and distribution of *S. supranubius* and an ability to map the distribution of these variables at a scale relevant to individual shrubs. This level of knowledge was not available. Identifying and dealing with heterogeneous patterns is a matter of current research (Law et al., 2009). When measurements of environmental heterogeneity are statistical (as in this thesis) it is recommended that the conclusions of the heterogeneity analyses are justified with non-statistical scientific arguments (Illian et al., 2008). In this thesis the statistical assessments of heterogeneity were supported by visual identification of a potential source of heterogeneity; the pronounced ridge-trough topography of Substrates 3 and 4.

## **9.5.2 OTHER POSSIBLE EXPLANATORY FACTORS**

### **Interspecific interactions**

*Evaluation and potential importance:*

Interspecific interactions between arid shrubs and other species are well documented (e.g. Gebauer et al., 2002; Holzapfel et al., 2006; Armas and Pugnaire, 2009). It is therefore considered likely that other species may influence the dynamics and therefore spatial structure of *S. supranubius* populations. In particular, the following species were noted for being relatively common on Substrate 1 – 5: *Adenocarpus*

*viscosus* (leguminous shrub), *Pterocephalus lasiospermus* (small shrub) and *Descurainia bourgeauana* (small shrub).

*How the process could be measured:*

Because of the slow dynamics and demographics of arid shrubs, investigating the potential influence of interspecific interactions on the dynamics of *S. supranubius* via experimental and/or observational techniques would be unfeasible. Pattern–process inference remains the most promising and accessible method of investigating interspecific interactions. The distribution of interspecific individuals could be manually mapped in the field, although the time and financial limitations of this technique would likely result in plots of small extent and low sample size. The imagery used in this thesis is not of a suitable resolution to map the distribution of small shrubs (e.g. *P. lasiospermus* and *D. bourgeauana*) and attempts to apply the imagery classification procedure utilised in Chapter 4 to the distribution of *A. viscosus* have so far been unsuccessful (data not shown). One possible technique that has so far received little attention in ecology is that of remote sensing from kite platforms. Such techniques are capable of producing high resolution imagery from which individuals of multispecies assemblages may be identifiable. One of the major advantages of this technique would be that the equipment could theoretically be deployed at optimal times (e.g. during flowering) to allow maximum species discrimination.

## **Grazing**

*Evaluation and likelihood:*

Studies suggest that livestock grazing can influence arid vegetation dynamics and the relative abundance of different functional groups (Metzger et al., 2005; Facelli and Springbett, 2009). There are two herbivores that may influence the spatial structure of *S. supranubius* populations: the Corsican mouflon and rabbits. While attempts are being made to control both species, they are still present in the caldera. Field observations made in December 2007 indicate that at some locations active grazing

was occurring, to a height of c. 70 cm. This thesis investigates the spatial structure of *S. supranubius* individuals with a canopy area of  $\geq 1 \text{ m}^2$ . These individuals are expected to have escaped the threat of grazing, although it is possible that the observed spatial patterns have been influenced by grazing-induced mortality of neighbouring individuals.

Personal observations suggest that the presence of the herbivores may not be evenly distributed across the five focal substrates. In particular, faecal evidence of rabbit inhabitation was variable between substrates, and was notably lower on those substrates with a high proportion of large surface clasts (e.g. Substrate 4). Thus it may be reasonable to assume that the any impact of grazing on the spatial structure of *S. supranubius* would not be equal between sites.

*How the process could be measured:*

Direct measurements of the impact of grazing on the spatial structure and dynamics of *S. supranubius* would likely be unfeasible because of the slow response of the species. A short-term assessment of the impact of grazing could be obtained by establishing exclosures that prevent rabbits accessing areas of the population. The growth and survival of individuals within these exclosures and nearby control areas could be compared. However, an understanding of the long term impacts of grazing may only be obtainable via mathematical models.

## **Sexual reproduction and dispersal**

*Evaluation and likelihood:*

Although field observations provided evidence of seed production, there was remarkably little evidence of regeneration via sexual reproduction (i.e. established seedlings), corresponding with previous assertions by Kyncl et al. (2006) that *S. supranubius* seedling establishment is highly episodic. These observations, however, do not preclude the possibility that sexual reproduction may influence the spatial structure of *S. supranubius*. Gravity is believed to be the major mode of *S. supranubius* seed dispersal as there is no evidence (either observational or in the

literature) to suggest that dispersal is either ballistic or is aided by animals, wind or water.

*How the process could be measured:*

It is likely that the influence of sexual reproduction and dispersal on the spatial structure and dynamics of *S. supranubius* will be strongly linked to the presence and intensity of grazing. Because of the slow demographics and dynamics of *S. supranubius* both processes would be best studied in mathematical models.

### **Disturbances**

*Evaluation and likelihood:*

The Cañadas caldera is believed to be a relatively undisturbed habitat. There is no documented evidence of fires and the low precipitation level reduces the likelihood of hydrological disturbances. Furthermore, the substrates investigated are geomorphologically stable and there is no evidence of pathogen-induced disturbance. The only potential source of disturbance is believed to be the impact of human visitors to the area. There is, however, a well-defined network of paths in area and the focal plots used in this thesis were situated away from any major roads or tracks.

### **Summarising remarks**

Pattern–process inference is typically limited to investigations of specific processes for which predicted spatial hypotheses can be developed. The technique provides little opportunity to investigate the operation and relative importance of multiple and interacting processes. With observational and experimental techniques largely unfeasible when the focal species is a slow-growing arid shrub, it seems that investigating the operation and potential interaction of alternative processes will require relatively sophisticated mathematical modelling techniques. In addition to improving the methodological application of pattern–process inference, the utilisation of mathematical simulation models is expected to be an important technique in advancing our understanding of arid shrub ecology.

## REFERENCES

---

- Ablay, G.J. & Martí, J. (2000) Stratigraphy, structure, and volcanic evolution of the Pico-Teide Pico-Viejo formation, Tenerife, Canary Islands. *Journal of Volcanology and Geothermal Research*, **103**, 175-208.
- Acevedo, M.A., Corrada-Bravo, C.J., Corrada-Bravo, H., Villanueva-Rivera, L.J., & Aide, T.M. (2009) Automated classification of bird and amphibian calls using machine learning: a comparison of methods. *Ecological Informatics*, **4**, 206-214.
- Adler, P.B., HilleRisLambers, J., & Levine, J.M. (2009) Weak effect of climate variability on coexistence in a sagebrush steppe community. *Ecology*, **90**, 3303-3312.
- Agrawal, A.A., Ackerly, D.D., Adler, F., Arnold, A.E., Cáceres, C., Doak, D.F., Post, E., Hudson, P.J., Maron, J., Mooney, K.A., Power, M., Schemske, D., Stachowicz, J., Strauss, S., Turner, M.G., & Werner, E. (2007) Filling key gaps in population and community ecology. *Frontiers in Ecology and the Environment*, **5**, 145-152.
- Akaike, H. (1974) New look at statistical-model identification. *Institute of Electrical and Electronic Engineers (IEEE) Transactions on Automatic Control*, **19**, 716-723.
- Aldrich, P.R., Parker, G.R., Ward, J.S., & Michler, C.H. (2003) Spatial dispersion of trees in an old-growth temperate hardwood forest over 60 years of succession. *Forest Ecology and Management*, **180**, 475-491.
- Ali, S. & Smith-Miles, K. (2007) An optimal degree selection for polynomial kernel with support vector machines: theoretical and empirical investigations. *International Journal of Knowledge-based and Intelligent Engineering Systems*, **11**, 1-18.
- Allen, A.P., Pockman, W.T., Restrepo, C., & Milne, B.T. (2008) Allometry, growth and population regulation of the desert shrub *Larrea tridentata*. *Functional Ecology*, **22**, 197-204.
- Alvarez, J.A., Villagra, P.E., Rossi, B.E., & Cesca, E.M. (2009) Spatial and temporal litterfall heterogeneity generated by woody species in the Central Monte desert. *Plant Ecology*, **205**, 295-303.
- Amarasekare, P. (2003) Competitive coexistence in spatially structured environments: a synthesis. *Ecology Letters*, **6**, 1109-1122.
- Ambrose, J.P. & Bratton, S.P. (1990) Trends in landscape heterogeneity along the borders of Great Smoky Mountains National-Park. *Conservation Biology*, **4**, 135-143.
- Anderson, K. & Croft, H. (2009) Remote sensing of soil surface properties. *Progress in Physical Geography*, **33**, 457-473.
- Angell, D.L. & McClaran, M.P. (2001) Long-term influences of livestock management and a non-native grass on grass dynamics in the Desert Grassland. *Journal of Arid Environments*, **49**, 507-520.
- Anthelme, F. & Michalet, R. (2009) Grass-to-tree facilitation in an arid grazed environment (Air Mountains, Sahara). *Basic and Applied Ecology*, **10**, 437-446.

- Aplin, P. (2005) Remote sensing: ecology. *Progress in Physical Geography*, **29**, 104-113.
- Arévalo, J.R. & Fernández-Palacios, J.M. (2003) Spatial patterns of trees and juveniles in a laurel forest of Tenerife, Canary Islands. *Plant Ecology*, **165**, 1-10.
- Armas, C. & Pugnaire, F.I. (2005) Plant interactions govern population dynamics in a semi-arid plant community. *Journal of Ecology*, **93**, 978-989.
- Armas, C. & Pugnaire, F.I. (2009) Ontogenetic shifts in interactions of two dominant shrub species in a semi-arid coastal sand dune system. *Journal of Vegetation Science*, **20**, 535-546.
- Baddeley, A.J., Møller, J., & Waagepetersen, R. (2000) Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica*, **54**, 329-350.
- Baddeley, A. & Turner, R. (2005) spatstat: an R package for analyzing spatial point patterns. *Journal of Statistical Software*, **12**, 1-42.
- Bailey, T.C. & Gatrell, A.C. (1995) *Interactive Spatial Data Analysis*. Longman, Harlow.
- Barbeito, I., Pardos, M., Calama, R., & Cañellas, I. (2008) Effect of stand structure on Stone Pine (*Pinus pinea* L.) regeneration dynamics. *Forestry*, **81**, 617-629.
- Barbeito, I., Fortin, M.J., Montes, F., & Cañellas, I. (2009) Response of pine natural regeneration to small-scale spatial variation in a managed Mediterranean mountain forest. *Applied Vegetation Science*, **12**, 488-503.
- Barbier, N., Couteron, P., Lefever, R., Deblauwe, V., & Lejeune, O. (2008) Spatial decoupling of facilitation and competition at the origin of gapped vegetation patterns. *Ecology*, **89**, 1521-1531.
- Barbour, M.G. (1969) Age and space distribution of desert shrub *Larrea divaricata*. *Ecology*, **50**, 679-685.
- Beckage, B. & Clark, J.S. (2003) Seedling survival and growth of three forest tree species: the role of spatial heterogeneity. *Ecology*, **84**, 1849-1861.
- Bell, M.L. & Grunwald, G.K. (2004) Mixed models for the analysis of replicated spatial point patterns. *Biostatistics*, **5**, 633-648.
- Belousov, A.I., Verzakov, S.A., & von Frese, J. (2002) A flexible classification approach with optimal generalisation performance: support vector machines. *Chemometrics and Intelligent Laboratory Systems*, **64**, 15-25.
- Bertness, M.D. & Callaway, R. (1994) Positive interactions in communities. *Trends in Ecology & Evolution*, **9**, 191-193.
- Besag, J. (1975) Statistical analysis of non-lattice data. *Statistician*, **24**, 179-195.
- Besag, J.E. (1977) Contribution to the discussion of Dr. Ripley's paper. *Journal of the Royal Statistical Society, Series B*, **39**, 193-195.
- Bestelmeyer, B.T., Ward, J.P., & Havstad, K.M. (2006) Soil-geomorphic heterogeneity governs patchy vegetation dynamics at an arid ecotone. *Ecology*, **87**, 963-973.



- Biganzoli, F., Wiegand, T., & Batista, W.B. (2009) Fire-mediated interactions between shrubs in a South American temperate savannah. *Oikos*, **118**, 1383-1395.
- Bisigato, A.J., Villagra, P.E., Ares, J.O., & Rossi, B.E. (2009) Vegetation heterogeneity in Monte Desert ecosystems: a multi-scale approach linking patterns and processes. *Journal of Arid Environments*, **73**, 182-191.
- Borcard, D., Legendre, P., Avois-Jacquet, C., & Tuomisto, H. (2004) Dissecting the spatial structure of ecological data at multiple scales. *Ecology*, **85**, 1826-1832.
- Bory de Saint-Vincent, J.B. (1802) *Essais sur les Iles Fortunées (Essays on the Fortunate Islands)*, Baudouin, Paris.
- Bowers, J.E., Turner, R.M., & Burgess, T.L. (2004) Temporal and spatial patterns in emergence and early survival of perennial plants in the Sonoran Desert. *Plant Ecology*, **172**, 107-119.
- Bowers, J.E. (2005) Effects of drought on shrub survival and longevity in the northern Sonoran Desert. *Journal of the Torrey Botanical Society*, **132**, 421-431.
- Bramwell, D. (1976). The endemic flora of the Canary Islands: distribution, relationships and phytogeography. In: Kunkel, G. (ed.) *Biogeography and Ecology in the Canary Islands*. W. Junk, The Hague, pp. 207-240.
- Bramwell, D. & Bramwell, Z. (2001) *Wild Flowers of the Canary Islands*, 2<sup>nd</sup> edition. Editorial Rueda, Madrid.
- Breshears, D.D., Myers, O.B., & Barnes, F.J. (2009) Horizontal heterogeneity in the frequency of plant-available water with woodland intercanopy-canopy vegetation patch type rivals that occurring vertically by soil depth. *Ecohydrology*, **2**, 503-519.
- Briggs, J.M., Schaafsma, H., & Trenkov, D. (2007) Woody vegetation expansion in a desert grassland: prehistoric human impact? *Journal of Arid Environments*, **69**, 458-472.
- Briones, O., Montana, C., & Ezcurra, E. (1996) Competition between three Chihuahuan desert species: evidence from plant size-distance relations and root distribution. *Journal of Vegetation Science*, **7**, 453-460.
- Briones, O., Montana, C., & Ezcurra, E. (1998) Competition intensity as a function of resource availability in a semiarid ecosystem. *Oecologia*, **116**, 365-372.
- Brooker, R.W., Maestre, F.T., Callaway, R.M., Lortie, C.L., Cavieres, L.A., Kunstler, G., Liancourt, P., Tielbörger, K., Travis, J.M.J., Anthelme, F., Armas, C., Coll, L., Corcket, E., Delzon, S., Forey, E., Kikvidze, Z., Olofsson, J., Pugnaire, F., Quiroz, C.L., Saccone, P., Schiffers, K., Seifan, M., Touzard, B., & Michalet, R. (2008) Facilitation in plant communities: the past, the present, and the future. *Journal of Ecology*, **96**, 18-34.
- Bruehlheide, H., Jandt, U., Gries, D., Thomas, F.M., Foetzki, A., Buerkert, A., Gang, W., Zhang, X.M., & Runge, M. (2003) Vegetation changes in a river oasis on the southern rim of the Taklamakan Desert in China between 1956 and 2000. *Phytocoenologia*, **33**, 801-818.
- Bruno, J.F., Stachowicz, J.J., & Bertness, M.D. (2003) Inclusion of facilitation into ecological theory. *Trends in Ecology & Evolution*, **18**, 119-125.

- Buonopane, M., Huenneke, L.F., & Remmenga, M. (2005) Community response to removals of plant functional groups and species from a Chihuahuan desert shrubland. *Oikos*, **110**, 67-80.
- Burnham, K.P. & Anderson, D.R. (2002) *Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach*, 2<sup>nd</sup> edition. Springer-Verlag, New York.
- Cade, B.S. & Noon, B.R. (2003) A gentle introduction to quantile regression for ecologists. *Frontiers in Ecology and the Environment*, **1**, 412-420.
- Caldwell, T.G., Young, M.H., Zhu, J.T., & McDonald, E.V. (2008) Spatial structure of hydraulic properties from canopy to interspace in the Mojave Desert. *Geophysical Research Letters*, **35**, L19406, doi:10.1029/2008GL035095.
- Cale, W.G., Henebry, G.M., & Yeakley, J.A. (1989) Inferring process from pattern in natural communities. *Bioscience*, **39**, 600-605.
- Call, L.J. & Nilsen, E.T. (2003) Analysis of spatial patterns and spatial association between the invasive tree-of-heaven (*Ailanthus altissima*) and the native black locust (*Robinia pseudoacacia*). *American Midland Naturalist*, **150**, 1-14.
- Camps-Valls, G., Gómez-Chova, L., Calpe-Maravilla, J., Martín-Guerrero, J.D., Soria-Olivas, E., Alonso-Chordá, L., & Moreno, J. (2004) Robust support vector method for hyperspectral data classification and knowledge discovery. *Institute of Electrical and Electronic Engineers (IEEE) Transactions on Geoscience and Remote Sensing*, **42**, 1530-1542.
- Cao, Y. (2008) *Bhattacharyya distance measure for pattern recognition*, MATLAB Central File Exchange, <http://www.mathworks.com/matlabcentral/fileexchange/18662-bhattacharyya-distance-measure-for-pattern-recognition>.
- Carine, M.A., Humphries, C.J., Guma, I.R., Reyes-Betancort, J.A., & Guerra, A.S. (2009) Areas and algorithms: evaluating numerical approaches for the delimitation of areas of endemism in the Canary Islands archipelago. *Journal of Biogeography*, **36**, 593-611.
- Carmel, Y. & Kadmon, R. (1998) Computerized classification of Mediterranean vegetation using panchromatic aerial photographs. *Journal of Vegetation Science*, **9**, 445-454.
- Carracedo, J.C. & Day, S. (2002) *Canary Islands*, Terra Publishing, Hertfordshire.
- Chapelle, O., Haffner, P., & Vapnik, V.N. (1999) Support vector machines for histogram-based image classification. *Institute of Electrical and Electronic Engineers (IEEE) Transactions on Neural Networks*, **10**, 1055-1064.
- Chaplot, V. & Le Bissonnais, Y. (2000) Field measurements of interrill erosion under different slopes and plot sizes. *Earth Surface Processes and Landforms*, **25**, 145-153.
- Chesson, P. (2000a) General theory of competitive coexistence in spatially-varying environments. *Theoretical Population Biology*, **58**, 211-237.
- Chesson, P. (2000b) Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics*, **31**, 343-366.
- Chesson, P., Gebauer, R.L.E., Schwinning, S., Huntly, N., Wiegand, K., Ernest, M.S.K., Sher, A., Novoplansky, A., & Weltzin, J.F. (2004) Resource pulses,

- species interactions, and diversity maintenance in arid and semi-arid environments. *Oecologia*, **141**, 236-253.
- Clark, P.J. & Evans, F.C. (1954) Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology*, **35**, 445-453.
- Clark-Tapia, R., Mandujano, M.C., Valverde, T., Mendoza, A., & Molina-Freaner, F. (2005) How important is clonal recruitment for population maintenance in rare plant species? The case of the narrow endemic cactus, *Stenocereus eruca*, in Baja California, Mexico. *Biological Conservation*, **124**, 123-132.
- Clos-Arceuduc, M. (1956) Etude sur photographies aériennes d'une formation végétale sahélienne: la brousse tigrée. *Bulletin de l'Institut Français d'Afrique Noire*, **7**, 678-684.
- Cody, M.L. (2000) Slow-motion population dynamics in Mojave Desert perennial plants. *Journal of Vegetation Science*, **11**, 351-358.
- Cohen, J. (1988) *Statistical power analysis for the behavioral sciences*, 2<sup>nd</sup> edition. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Comas, C. & Mateu, J. (2007) Modelling forest dynamics: a perspective from point process methods. *Biometrical Journal*, **49**, 176-196.
- Coomes, D.A., Duncan, R.P., Allen, R.B., & Truscott, J. (2003) Disturbances prevent stem size-density distributions in natural forests from following scaling relationships. *Ecology Letters*, **6**, 980-989.
- Coomes, D.A. & Allen, R.B. (2007) Mortality and tree-size distributions in natural mixed-age forests. *Journal of Ecology*, **95**, 27-40.
- Cortes, C. & Vapnik, V. (1995) Support-vector networks. *Machine Learning*, **20**, 273-297.
- Cottam, G., Curtis, J.T., & Hale, B.W. (1953) Some sampling characteristics of population of randomly dispersed individuals. *Ecology*, **34**, 741-757.
- Cousens, R., Wiegand, T., & Taghizadeh, M. (2008) Small-scale spatial structure within patterns of seed dispersal. *Oecologia*, **158**, 437-448.
- Couteron, P. (2001) Using spectral analysis to confront distributions of individual species with an overall periodic pattern in semi-arid vegetation. *Plant Ecology*, **156**, 229-243.
- Couteron, P. (2002) Quantifying change in patterned semi-arid vegetation by Fourier analysis of digitized aerial photographs. *International Journal of Remote Sensing*, **23**, 3407-3425.
- Cressie, N. & Collins, L.B. (2001) Analysis of spatial point patterns using bundles of product density LISA functions. *Journal of Agricultural Biological and Environmental Statistics*, **6**, 118-135.
- Curtis, J.T. & McIntosh, R.P. (1950) The interrelations of certain analytic and synthetic phytosociological characters. *Ecology*, **31**, 434-455.
- Dahlgren, J.P. & Ehrlén, J. (2009) Linking environmental variation to population dynamics of a forest herb. *Journal of Ecology*, **97**, 666-674.
- Dale, M.R.T. (1999) *Spatial Pattern Analysis in Plant Ecology*. Cambridge University Press, Cambridge.

- Dale, M.R.T. & Powell, R.D. (2001) A new method for characterizing point patterns in plant ecology. *Journal of Vegetation Science*, **12**, 597-608.
- 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*, **25**, 558-277.
- Davis, J.H., Howe, R.W., & Davis, G.J. (2000) A multi-scale spatial analysis method for point data. *Landscape Ecology*, **15**, 99-114.
- de la Torre, W.W. & Osorio, V.E.M. (2004). Flora vascular y vegetación. In: Canseco, V.G. (ed.) *Parque Nacional del Teide*. Canseco Editores, Talavera de la Reina, Spain, pp. 97-142.
- Díaz, F., Jimenez, C.C., & Tejedor, M. (2005) Influence of the thickness and grain size of tephra mulch on soil water evaporation. *Agricultural Water Management*, **74**, 47-55.
- Dovčiak, M., Frelich, L.E., & Reich, P.B. (2001) Discordance in spatial patterns of white pine (*Pinus strobus*) size-classes in a patchy near-boreal forest. *Journal of Ecology*, **89**, 280-291.
- Dickson, J.H., Rodriguez, J.C., & Machado, A. (1987) Invading plants at high-altitudes on Tenerife especially in the Teide National Park. *Botanical Journal of the Linnean Society*, **95**, 155-179.
- Diggle, P.J., Mateu, J., & Clough, H.E. (2000) A comparison between parametric and non-parametric approaches to the analysis of replicated spatial point patterns. *Advances in Applied Probability*, **32**, 331-343.
- Diggle, P.J. (2003) *Statistical Analysis of Spatial Point Patterns*, 2nd edition. Arnold, London.
- Dixon, B. & Candade, N. (2008) Multispectral landuse classification using neural networks and support vector machines: one or the other, or both? *International Journal of Remote Sensing*, **29**, 1185-1206.
- Drake, J.M., Randin, C., & Guisan, A. (2006) Modelling ecological niches with support vector machines. *Journal of Applied Ecology*, **43**, 424-432.
- Driscoll, R.S. & Coleman, M.D. (1974) Color for shrubs. *Photogrammetric Engineering and Remote Sensing*, **40**, 451-459.
- Eccles, N.S., Esler, K.J., & Cowling, R.M. (1999) Spatial pattern analysis in Namaqualand desert plant communities: evidence for general positive interactions. *Plant Ecology*, **142**, 71-85.
- Edgar, C.J., Wolff, J.A., Olin, P.H., Nichols, H.J., Pittari, A., Cas, R.A.F., Reiners, P.W., Spell, T.L., & Martí, J. (2007) The late quaternary Diego Hernandez Formation, Tenerife: volcanology of a complex cycle of voluminous explosive phonolitic eruptions. *Journal of Volcanology and Geothermal Research*, **160**, 59-85.
- Ehleringer, J.R. & Mooney, H.A. (1978) Leaf hairs: effects on physiological activity and adaptive value to a desert shrub. *Oecologia*, **37**, 183-200.
- Facelli, J.M. & Springbett, H. (2009) Why do some species in arid lands increase under grazing? Mechanisms that favour increased abundance of *Maireana pyramidata* in overgrazed chenopod shrublands of South Australia. *Austral Ecology*, **34**, 588-597.

- Fagerlund, S. (2007) Bird species recognition using support vector machines. *Eurasip Journal on Advances in Signal Processing*, 1-8.
- Fajardo, A. & Alaback, P. (2005) Effects of natural and human disturbances on the dynamics and spatial structure of *Nothofagus glauca* in south-central Chile. *Journal of Biogeography*, **32**, 1811-1825.
- Fajardo, A., Goodburn, J.M., & Graham, J. (2006) Spatial patterns of regeneration in managed uneven-aged ponderosa pine Douglas-fir forests of Western Montana, USA. *Forest Ecology and Management*, **223**, 255-266.
- Fajardo, A., Quiroz, C.L., & Cavieres, L.A. (2008) Distinguishing colonisation modes from spatial structures in populations of the cushion plant *Azorella madreporica* in the high-Andes of central Chile. *Austral Ecology*, **33**, 703-712.
- Fang, W. (2005) Spatial analysis of an invasion front of *Acer platanoides*: dynamic inferences from static data. *Ecography*, **28**, 283-294.
- Fehmi, J.S. & Bartolome, J.W. (2001) A grid-based method for sampling and analysing spatially ambiguous plants. *Journal of Vegetation Science*, **12**, 467-472.
- Felinks, B. & Wiegand, T. (2008) Exploring spatiotemporal patterns in early stages of primary succession on former lignite mining sites. *Journal of Vegetation Science*, **19**, 267-276.
- Fensham, R.J., Fairfax, R.J., Holman, J.E., & Whitehead, P.J. (2002) Quantitative assessment of vegetation structural attributes from aerial photography. *International Journal of Remote Sensing*, **23**, 2293-2317.
- Fernandopullé, D. (1976) Climatic characteristics of the Canary Islands. In: Kunkel, G. (ed.) *Biogeography and Ecology in the Canary Islands*. W. Junk, The Hague, pp. 185-206.
- Fonteyn, P.J. & Mahall, B.E. (1981) An experimental analysis of structure in a desert plant community. *Journal of Ecology*, **69**, 883-896.
- Foody, G.M. (2002) Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, **80**, 185-201.
- Foody, G.M. (2004) Spatial nonstationarity and scale-dependency in the relationship between species richness and environmental determinants for the sub-Saharan endemic avifauna. *Global Ecology and Biogeography*, **13**, 315-320.
- Foody, G.M. & Mathur, A. (2004) Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification. *Remote Sensing of Environment*, **93**, 107-117.
- Foody, G.M. (2005) Local characterization of thematic classification accuracy through spatially constrained confusion matrices. *International Journal of Remote Sensing*, **26**, 1217-1228.
- Foody, G.M. & Mathur, A. (2006) The use of small training sets containing mixed pixels for accurate hard image classification: training on mixed spectral responses for classification by a SVM. *Remote Sensing of Environment*, **103**, 179-189.
- Foody, G.M., Mathur, A., Sanchez-Hernandez, C., & Boyd, D.S. (2006) Training set size requirements for the classification of a specific class. *Remote Sensing of Environment*, **104**, 1-14.

- Foody, G.M. (2007) Editorial: ecological applications of remote sensing and GIS. *Ecological Informatics*, **2**, 71-72.
- Fortin, M.-J. (1999). Spatial statistics in landscape ecology. In: Klopatek, J.M. & Gardner, R.H. (eds.) *Landscape Ecological Analysis: Issues and Applications*. Springer-Verlag, New York, pp. 253-279.
- Fortin, M.-J., Dale, M.R.T., & Hoef, J.V. (2002). Spatial analysis in ecology. In: El-shaarawi, A.H. & Piegorisch, W.W. (eds.) *Encyclopedia of Environmetrics*, Vol. 4. John Wiley and Sons, Chichester, pp. 2051-2058.
- Fortin, M.-J. & Dale, M.R.T. (2005) *Spatial Analysis: A Guide for Ecologists*. Cambridge University Press, Cambridge.
- Francisco-Ortega, J., Santos-Guerra, A., & Bacallado, J.J. (2009). Canary Islands, Biology. In: Gillespie, R.G. & Clague, D.A. (eds.) *Encyclopedia of Islands*. University of California Press, London, pp. 127-133.
- Fravolini, A., Hultine, K.R., Brugnoli, E., Gazal, R., English, N.B., & Williams, D.G. (2005) Precipitation pulse use by an invasive woody legume: the role of soil texture and pulse size. *Oecologia*, **144**, 618-627.
- Freckleton, R.P., Watkinson, A.R., & Rees, M. (2009) Measuring the importance of competition in plant communities. *Journal of Ecology*, **97**, 379-384.
- Freeman, E.A. & Ford, E.D. (2002) Effects of data quality on analysis of ecological pattern using the K(d) statistical function. *Ecology*, **83**, 35-46.
- Freestone, A.L. (2006) Facilitation drives local abundance and regional distribution of a rare plant in a harsh environment. *Ecology*, **87**, 2728-2735.
- Fujita, T., Itaya, A., Miura, M., Manabe, T., & Yamamoto, S.I. (2003) Long-term canopy dynamics analysed by aerial photographs in a temperate old-growth evergreen broad-leaved forest. *Journal of Ecology*, **91**, 686-693.
- Fulé, P.Z. & Covington, W.W. (1998) Spatial patterns of Mexican pine-oak forests under different recent fire regimes. *Plant Ecology*, **134**, 197-209.
- Galindo, I., Soriano, C., Martí, J., & Pérez, N. (2005) Graben structure in the Las Cañadas edifice (Tenerife, Canary Islands): implications for active degassing and insights on the caldera formation. *Journal of Volcanology and Geothermal Research*, **144**, 73-87.
- Gebauer, R.L.E., Schwinning, S., & Ehleringer, J.R. (2002) Interspecific competition and resource pulse utilization in a cold desert community. *Ecology*, **83**, 2602-2616.
- Gehlke, C.E. & Biel, K. (1934) Certain effects of grouping upon the size of the correlation coefficient in census tract material. *Journal of the American Statistical Association*, **29**, 169-170.
- Getis, A. & Franklin, J. (1987) 2nd-order neighborhood analysis of mapped point patterns. *Ecology*, **68**, 473-477.
- Getzin, S., Dean, C., He, F.L., Trofymow, J.A., Wiegand, K., & Wiegand, T. (2006) Spatial patterns and competition of tree species in a Douglas-fir chronosequence on Vancouver Island. *Ecography*, **29**, 671-682.

- Getzin, S. & Wiegand, K. (2007) Asymmetric tree growth at the stand level: random crown patterns and the response to slope. *Forest Ecology and Management*, **242**, 165-174.
- Getzin, S., Wiegand, T., Wiegand, K., & He, F.L. (2008) Heterogeneity influences spatial patterns and demographics in forest stands. *Journal of Ecology*, **96**, 807-820.
- Gill, R., Thirlwall, M., Marriner, G., Millward, D., Norry, M., Saunders, A., & Martí, J. (1994) *Tenerife, Canary Islands*. Geological Society Publishing House, Bath.
- Goldberg, D.E. & Turner, R.M. (1986) Vegetation change and plant demography in permanent plots in the Sonoran Desert. *Ecology*, **67**, 695-712.
- Goldberg, D.E., Rajaniemi, T., Gurevitch, J., & Stewart-Oaten, A. (1999) Empirical approaches to quantifying interaction intensity: competition and facilitation along productivity gradients. *Ecology*, **80**, 1118-1131.
- Golluscio, R.A., Sala, O.E., & Lauenroth, W.K. (1998) Differential use of large summer rainfall events by shrubs and grasses: a manipulative experiment in the Patagonian steppe. *Oecologia*, **115**, 17-25.
- Golubski, A.J., Gross, K.L., & Mittelbach, G.G. (2008) Competition among plant species that interact with their environment at different spatial scales. *Proceedings of the Royal Society B-Biological Sciences*, **275**, 1897-1906.
- Goreaud, F. & Pélissier, R. (1999) On explicit formulas of edge effect correction for Ripley's K-function. *Journal of Vegetation Science*, **10**, 433-438.
- Gorgevik, D. & Cakmakov, D. (2005) Handwritten digit recognition by combining SVM classifiers. *Proceedings of the Eurocon 2005 International Conference on Computers as a Tool*, pp. 1393-1396.
- Goslee, S.C., Havstad, K.M., Peters, D.P.C., Rango, A., & Schlesinger, W.H. (2003) High-resolution images reveal rate and pattern of shrub encroachment over six decades in New Mexico, USA. *Journal of Arid Environments*, **54**, 755-767.
- Gray, L. & He, F.L. (2009) Spatial point-pattern analysis for detecting density-dependent competition in a boreal chronosequence of Alberta. *Forest Ecology and Management*, **259**, 98-106.
- Grayson, R.B., Western, A.W., Walker, J.P., Kandel, D.D., Costelloe, J.F., & Wilson, D.J. (2006). Controls on patterns of soil moisture in arid and semi-arid systems. In: D'Odorico, P. & Porporato, A. (eds.) *Dryland Ecohydrology*. Springer, Dordrecht, The Netherlands, pp. 109-128.
- Greig-Smith, P. (1983) *Quantitative Plant Ecology*, 3<sup>rd</sup> edition. Blackwell Scientific Publications, Oxford.
- Guillou, H., Carracedo, J.C., Paris, R., & Torrado, F.J.P. (2004) Implications for the early shield-stage evolution of Tenerife from K/Ar ages and magnetic stratigraphy. *Earth and Planetary Science Letters*, **222**, 599-614.
- Guisan, A., Lehmann, A., Ferrier, S., Austin, M., Overton, J.M.C., Aspinall, R., & Hastie, T. (2006) Making better biogeographical predictions of species' distributions. *Journal of Applied Ecology*, **43**, 386-392.
- Guo, Q.H., Kelly, M., & Graham, C.H. (2005) Support vector machines for predicting distribution of sudden oak death in California. *Ecological Modelling*, **182**, 75-90.

- Guo, B.F., Gunn, S.R., Damper, R.I., & Nelson, J.D.B. (2008) Customizing kernel functions for SVM-based hyperspectral image classification. *Institute of Electrical and Electronics Engineers (IEEE) Transactions on Image Processing*, **17**, 622-629.
- Haase, P. (1995) Spatial pattern analysis in ecology based on Ripley K-function: introduction and methods of edge correction. *Journal of Vegetation Science*, **6**, 575-582.
- Haase, P., Pugnaire, F.I., Clark, S.C., & Incoll, L.D. (1996) Spatial patterns in a two-tiered semi-arid shrubland in southeastern Spain. *Journal of Vegetation Science*, **7**, 527-534.
- Haase, P., Pugnaire, F.I., Clark, S.C., & Incoll, L.D. (1997) Spatial pattern in *Anthyllis cytisoides* shrubland on abandoned land in southeastern Spain. *Journal of Vegetation Science*, **8**, 627-634.
- Hamerlynck, E.P., McAuliffe, J.R., & Smith, S.D. (2000) Effects of surface and sub-surface soil horizons on the seasonal performance of *Larrea tridentata* (creosote bush). *Functional Ecology*, **14**, 596-606.
- Hamerlynck, E.P., McAuliffe, J.R., McDonald, E.V., & Smith, S.D. (2002) Ecological responses of two Mojave Desert shrubs to soil horizon development and soil water dynamics. *Ecology*, **83**, 768-779.
- Hamerlynck, E.P. & McAuliffe, J.R. (2008) Soil-dependent canopy die-back and plant mortality in two Mojave Desert shrubs. *Journal of Arid Environments*, **72**, 1793-1802.
- Hartgerink, A.P. & Bazzaz, F.A. (1984) Seedling-scale environmental heterogeneity influences individual fitness and population structure. *Ecology*, **65**, 198-206.
- Hartle, R.T., Fernandez, G.C.J., & Nowak, R.S. (2006) Horizontal and vertical zones of influence for root systems of four Mojave Desert shrubs. *Journal of Arid Environments*, **64**, 586-603.
- Heisler-White, J.L., Knapp, A.K., & Kelly, E.F. (2008) Increasing precipitation event size increases aboveground net primary productivity in a semi-arid grassland. *Oecologia*, **158**, 129-140.
- Henschel, J.R. & Seely, M.K. (2000) Long-term growth patterns of *Welwitschia mirabilis*, a long-lived plant of the Namib Desert. *Plant Ecology*, **150**, 7-26.
- Hewitt, J.E., Thrush, S.F., Dayton, P.K., & Bonsdorff, E. (2007) The effect of spatial and temporal heterogeneity on the design and analysis of empirical studies of scale-dependent systems. *American Naturalist*, **169**, 398-408.
- Hill, J.K. & Hamer, K.C. (2004) Determining impacts of habitat modification on diversity of tropical forest fauna: the importance of spatial scale. *Journal of Applied Ecology*, **41**, 744-754.
- Holzapfel, C. & Mahall, B.E. (1999) Bidirectional facilitation and interference between shrubs and annuals in the Mojave Desert. *Ecology*, **80**, 1747-1761.
- Holzapfel, C., Tielbörger, K., Parag, H.A., Kigel, J., & Sternberg, M. (2006) Annual plant-shrub interactions along an aridity gradient. *Basic and Applied Ecology*, **7**, 268-279.
- Honnay, O. & Bossuyt, B. (2005) Prolonged clonal growth: escape route or route to extinction? *Oikos*, **108**, 427-432.



- Honnay, O., Jacquemyn, H., Roldán-Ruiz, I., & Hermy, M. (2006) Consequences of prolonged clonal growth on local and regional genetic structure and fruiting success of the forest perennial *Maianthemum bifolium*. *Oikos*, **112**, 21-30.
- Huang, C., Davis, L.S., & Townshend, J.R.G. (2002) An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing*, **23**, 725-749.
- Hubert-Moy, L., Cotannec, A., Le Du, L., Chardin, A., & Perez, P. (2001) A comparison of parametric classification procedures of remotely sensed data applied on different landscape units. *Remote Sensing of Environment*, **75**, 174-187.
- Hudak, A.T. & Wessman, C.A. (1998) Textural analysis of historical aerial photography to characterize woody plant encroachment in South African savanna. *Remote Sensing of Environment*, **66**, 317-330.
- Huebner, C.D., Vankat, J.L., & Renwick, W.H. (1999) Change in the vegetation mosaic of central Arizona USA between 1940 and 1989. *Plant Ecology*, **144**, 83-91
- Hurlbert, A.H. & White, E.P. (2005) Disparity between range map- and survey-based analyses of species richness: patterns, processes and implications. *Ecology Letters*, **8**, 319-327.
- Huxman, T.E., Snyder, K.A., Tissue, D., Leffler, A.J., Ogle, K., Pockman, W.T., Sandquist, D.R., Potts, D.L., & Schwinning, S. (2004) Precipitation pulses and carbon fluxes in semiarid and arid ecosystems. *Oecologia*, **141**, 254-268.
- Ibrahim, S. & Hashim, I. (1990) Classification of mangrove forest by using 1:40000 scale aerial photographs. *Forest Ecology and Management*, **33-4**, 583-592.
- Illa, E., Carrillo, E., & Ninot, J.M. (2006) Patterns of plant traits in Pyrenean alpine vegetation. *Flora*, **201**, 528-546.
- Illian, J., Penttinen, A., Stoyan, H., & Stoyan, D. (2008) *Statistical Analysis and Modelling of Spatial Point Patterns*. John Wiley and Sons, Chichester.
- Imeson, A.C. & Prinsen, H.A.M. (2004) Vegetation patterns as biological indicators for identifying runoff and sediment source and sink areas for semi-arid landscapes in Spain. *Agriculture Ecosystems & Environment*, **104**, 333-342.
- Izquierdo, I., Martín, J.L., Zurita, N., & Arechavaleta, M.E. (2004) *Lista de especies silvestres de Canarias: hongos, plantas y animales terrestres*. Consejería de Política Territorial y Medio Ambiente, Spain.
- Jacobson, G.L., Almquist-Jacobson, H., & Winne, J.C. (1991) Conservation of rare plant habitat: insights from the recent history of vegetation and fire at Crystal Fen, Northern Maine, USA. *Biological Conservation*, **57**, 287-314.
- Jacquemyn, H., Wiegand, T., Vandepitte, K., Brys, R., Roldán-Ruiz, I., & Honnay, O. (2009) Multigenerational analysis of spatial structure in the terrestrial, food-deceptive orchid *Orchis mascula*. *Journal of Ecology*, **97**, 206-216.
- Jarabo-Lorenzo, A., Velazquez, E., Pérez-Galdona, R., Vega-Hernandez, M.C., Martínez-Molina, E., Mateos, P.F., Vinuesa, P., Martínez-Romero, E., & León-Barrios, M. (2000) Restriction fragment length polymorphism analysis of 16S rDNA and low molecular weight RNA profiling of rhizobial isolates from shrubby legumes endemic to the Canary islands. *Systematic and Applied Microbiology*, **23**, 418-425.

- Jefferson, L.V. & Pennacchio, M. (2005) The impact of shade on establishment of shrubs adapted to the high light irradiation of semi-arid environments. *Journal of Arid Environments*, **63**, 706-716.
- Jiménez-Lobato, V. & Valverde, T. (2006) Population dynamics of the shrub *Acacia bilimekii* in a semi-desert region in central Mexico. *Journal of Arid Environments*, **65**, 29-45.
- Ju, C.Y., Cai, T.J., & Yang, X.H. (2008) Topography-based modeling to estimate percent vegetation cover in semi-arid Mu Us sandy land, China. *Computers and Electronics in Agriculture*, **64**, 133-139.
- Kadmon, R. & Harari-Kremer, R. (1999) Studying long-term vegetation dynamics using digital processing of historical aerial photographs. *Remote Sensing of Environment*, **68**, 164-176.
- Kallimanis, A.S., Halley, J.M., Vokou, D., & Sgardelis, S.P. (2008) The scale of analysis determines the spatial pattern of woody species diversity in the Mediterranean environment. *Plant Ecology*, **196**, 143-151.
- Karnieli, A., Gabai, A., Ichoku, C., Zaady, E., & Shachak, M. (2002) Temporal dynamics of soil and vegetation spectral responses in a semi-arid environment. *International Journal of Remote Sensing*, **23**, 4073-4087.
- Kefi, S., Rietkerk, M., Alados, C.L., Pueyo, Y., Papanastasis, V.P., ElAich, A., & de Ruiter, P.C. (2007) Spatial vegetation patterns and imminent desertification in Mediterranean arid ecosystems. *Nature*, **449**, 213-U5.
- Kelly, M., Guo, Q., Liu, D., & Shaari, D. (2007) Modeling the risk for a new invasive forest disease in the United States: an evaluation of five environmental niche models. *Computers Environment and Urban Systems*, **31**, 689-710.
- Kenkel, N.C. (1988) Pattern of self-thinning in Jack-Pine: testing the random mortality hypothesis. *Ecology*, **69**, 1017-1024.
- Kenkel, N.C. (1993) Modeling Markovian dependence in populations of *Aralia-Nudicaulis*. *Ecology*, **74**, 1700-1706.
- Klausmeier, C.A. (1999) Regular and irregular patterns in semiarid vegetation. *Science*, **284**, 1826-1828.
- Klimeš, L., Klimešová, J., Hendriks, H., & van Groenendael, J. (1997). Clonal plant architecture: a comparative analysis of form and function. In: de Kroon, H. & van Groenendael, J. (eds.) *The Ecology and Evolution of Clonal Plants*. Backhuys Publishers, Leiden, pp. 1-29.
- Klimeš, L. (2008) Clonal splitters and integrators in harsh environments of the Trans-Himalaya. *Evolutionary Ecology*, **22**, 351-367.
- Koenker, R. (2005) *Quantile Regression*. Cambridge University Press, Cambridge.
- Koenker, R. (2009) *quantreg: Quantile Regression*. R Package Version 4.44, <http://CRAN.R-project.org/package=quantreg>.
- Koukoulas, S. & Blackburn, G.A. (2005) Spatial relationships between tree species and gap characteristics in broad-leaved deciduous woodland. *Journal of Vegetation Science*, **16**, 587-596.

- Kraaij, T. & Milton, S.J. (2006) Vegetation changes (1995-2004) in semi-arid Karoo shrubland, South Africa: effects of rainfall, wild herbivores and change in land use. *Journal of Arid Environments*, **64**, 174-192.
- Kummerow, J., Krause, D., & Jow, W. (1977) Root systems of Chaparral shrubs. *Oecologia*, **29**, 163-177.
- Kunkel, G. (1976). Introduction. In: Kunkel, G. (ed.) *Biogeography and Ecology in the Canary Islands*. W Junk, The Hague, pp. 1-14.
- Kunkel, G. (1980) *Die Kanarischen Inseln und ihre Pflanzenwelt*. Gustav Fischer, Stuttgart.
- Kyncl, T., Suda, J., Wild, J., Wildová, R., & Herben, T. (2006) Population dynamics and clonal growth of *Spartocytisus supranubius* (Fabaceae), a dominant shrub in the alpine zone of Tenerife, Canary Islands. *Plant Ecology*, **186**, 97-108.
- Lausi, D. & Nimis, P.L. (1986) Leaf and canopy adaptations in a high-elevation desert on Tenerife, Canary Islands. *Vegetatio*, **68**, 19-31.
- Law, R., Illian, J., Burslem, D.F.R.P., Gratzner, G., Gunatilleke, C.V.S., & Gunatilleke, I.A.U.N. (2009) Ecological information from spatial patterns of plants: insights from point process theory. *Journal of Ecology*, **97**, 616-628.
- Lawes, M.J., Griffiths, M.E., Midgley, J.J., Boudreau, S., Eeley, H.A.C., & Chapman, C.A. (2008) Tree spacing and area of competitive influence do not scale with tree size in an African rain forest. *Journal of Vegetation Science*, **19**, 729-738.
- Laliberte, A.S., Rango, A., Havstad, K.M., Paris, J.F., Beck, R.F., McNeely, R., & Gonzalez, A.L. (2004) Object-oriented image analysis for mapping shrub encroachment from 1937 to 2003 in southern New Mexico. *Remote Sensing of Environment*, **93**, 198-210.
- Leckie, D.G., Gougeon, F.A., Walsworth, N., & Paradine, D. (2003) Stand delineation and composition estimation using semi-automated individual tree crown analysis. *Remote Sensing of Environment*, **85**, 355-369.
- LeMay, V., Pommerening, A., & Marshall, P. (2009) Spatio-temporal structure of multi-storied, multi-aged interior Douglas fir (*Pseudotsuga menziesii* var. *glauca*) stands. *Journal of Ecology*, **97**, 1062-1074.
- Lepš, J. & Kindlmann, P. (1987) Models of the development of spatial pattern of an even-aged plant population over time. *Ecological Modelling*, **39**, 45-57.
- Levin, S.A. (1992) The problem of pattern and scale in ecology. *Ecology*, **73**, 1943-1967.
- Linares-Palomino, R. & Ponce-Alvarez, S.I. (2009) Structural patterns and floristics of a seasonally dry forest in Reserva Ecológica Chaparri, Lambayeque, Peru. *Tropical Ecology*, **50**, 305-314.
- Loik, M.E., Breshears, D.D., Lauenroth, W.K., & Belnap, J. (2004) A multi-scale perspective of water pulses in dryland ecosystems: climatology and ecohydrology of the western USA. *Oecologia*, **141**, 269-281.
- Loosmore, N.B. & Ford, E.D. (2006) Statistical inference using the G or K point pattern spatial statistics. *Ecology*, **87**, 1925-1931.
- Lortie, C.J. & Callaway, R.M. (2006) Re-analysis of meta-analysis: support for the stress-gradient hypothesis. *Journal of Ecology*, **94**, 7-16.

- Lortie, C.J. & Turkington, R. (2008) Species-specific positive effects in an annual plant community. *Oikos*, **117**, 1511-1521.
- Ludwig, J.A., Wilcox, B.P., Breshears, D.D., Tongway, D.J., & Imeson, A.C. (2005) Vegetation patches and runoff-erosion as interacting ecohydrological processes in semiarid landscapes. *Ecology*, **86**, 288-297.
- MacFadyen, W.A. (1950) Vegetation patterns in the semi-desert plains of British Somaliland. *Geographical Journal*, **116**, 199-210.
- Maestre, F.T., Bautista, S., & Cortina, J. (2003) Positive, negative, and net effects in grass-shrub interactions in mediterranean semiarid grasslands. *Ecology*, **84**, 3186-3197.
- Maestre, F.T. & Cortina, J. (2004) Do positive interactions increase with abiotic stress? A test from a semi-arid steppe. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **271**, 331-333.
- Maestre, F.T., Valladares, F., & Reynolds, J.F. (2005) Is the change of plant-plant interactions with abiotic stress predictable? A meta-analysis of field results in arid environments. *Journal of Ecology*, **93**, 748-757.
- Maestre, F.T., Callaway, R.M., Valladares, F., & Lortie, C.J. (2009) Refining the stress-gradient hypothesis for competition and facilitation in plant communities. *Journal of Ecology*, **97**, 199-205.
- Mahdi, A. & Law, R. (1987) On the spatial-organization of plant-species in a limestone grassland community. *Journal of Ecology*, **75**, 459-476.
- Maheu-Giroux, M. & de Blois, S. (2007) Landscape ecology of *Phragmites australis* invasion in networks of linear wetlands. *Landscape Ecology*, **22**, 285-301
- Malkinson, D., Kadmon, R., & Cohen, D. (2003) Pattern analysis in successional communities: an approach for studying shifts in ecological interactions. *Journal of Vegetation Science*, **14**, 213-222.
- Malkinson, D. & Jeltsch, F. (2007) Intraspecific facilitation: a missing process along increasing stress gradients - insights from simulated shrub populations. *Ecography*, **30**, 339-348.
- Malkinson, D. & Kadmon, R. (2007) Vegetation dynamics along a disturbance gradient: spatial and temporal perspectives. *Journal of Arid Environments*, **69**, 127-143.
- Mandujano, M.C., Montana, C., Franco, M., Golubov, J., & Flores-Martinez, A. (2001) Integration of demographic annual variability in a clonal desert cactus. *Ecology*, **82**, 344-359.
- Mandujano, M.C., Golubov, J., & Huenneke, L.F. (2007) Effect of reproductive modes and environmental heterogeneity in the population dynamics of a geographically widespread clonal desert cactus. *Population Ecology*, **49**, 141-153.
- Manevitz, L. & Yousef, M. (2007) One-class document classification via neural networks. *Neurocomputing*, **70**, 1466-1481.
- Manomaisupat, P., Vrusias, B., & Ahmad, K. (2006) Categorization of large text collections: Feature selection for training neural networks. *Proceedings of the 7<sup>th</sup> International Intelligent Data Engineering and Automated Learning (IDEAL) Conference*, pp. 1003-1013.

- Manor, A. & Shnerb, N.M. (2008) Facilitation, competition, and vegetation patchiness: from scale free distribution to patterns. *Journal of Theoretical Biology*, **253**, 838-842.
- Marrero-Gómez, M.V., Bafiáres-Baudet, A., & Carqué-Alamo, E. (2003) Plant resource conservation planning in protected natural areas: an example from the Canary Islands, Spain. *Biological Conservation*, **113**, 399-410.
- Martí, J. & Gudmundsson, A. (2000) The Las Cañadas caldera (Tenerife, Canary Islands): an overlapping collapse caldera generated by magma-chamber migration. *Journal of Volcanology and Geothermal Research*, **103**, 161-173.
- Mast, J.N. & Veblen, T.T. (1999) Tree spatial patterns and stand development along the pine-grassland ecotone in the Colorado Front Range. *Canadian Journal of Forest Research*, **29**, 575-584.
- Matérn, B. (1986) *Spatial Variation*. Springer, Berlin.
- Mateu, J. (2001) Parametric procedures in the analysis of replicated pairwise interaction point patterns. *Biometrical Journal*, **43**, 375-394.
- Mather, P.M. (2004) *Computer Processing of Remotely Sensed Images: An Introduction*, Wiley, Chichester.
- Mathur, A. & Foody, G.M. (2008) Crop classification by support vector machine with intelligently selected training data for an operational application. *International Journal of Remote Sensing*, **29**, 2227-2240.
- MathWorks (2007) *MATLAB: version R2007a*. The MathWorks Inc., Natick, MA.
- McAuliffe, J.R., Hamerlynck, E.P., & Eppes, M.C. (2007) Landscape dynamics fostering the development and persistence of long-lived creosotebush (*Larrea tridentata*) clones in the Mojave Desert. *Journal of Arid Environments*, **69**, 96-126.
- McClaran, M.P. & Angell, D.L. (2006) Long-term vegetation response to mesquite removal in desert grassland. *Journal of Arid Environments*, **66**, 686-697.
- McDonald, R.I., Peet, R.K., & Urban, D.L. (2003) Spatial pattern of *Quercus* regeneration limitation and *Acer rubrum* invasion in a Piedmont forest. *Journal of Vegetation Science*, **14**, 441-450.
- McDonald, A.K., Kinucan, R.J., & Loomis, L.E. (2009) Ecohydrological interactions within banded vegetation in the northeastern Chihuahuan Desert, USA. *Ecohydrology*, **2**, 66-71.
- McIntire, E.J.B. & Fajardo, A. (2009) Beyond description: the active and effective way to infer processes from spatial patterns. *Ecology*, **90**, 46-56.
- Meador, A.J.S., Moore, M.M., Bakker, J.D., & Parysow, P.F. (2009) 108 years of change in spatial pattern following selective harvest of a *Pinus ponderosa* stand in northern Arizona, USA. *Journal of Vegetation Science*, **20**, 79-90.
- Melgani, F. & Bruzzone, L. (2004) Classification of hyperspectral remote sensing images with support vector machines. *Institute of Electrical and Electronic Engineers (IEEE) Transactions on Geoscience and Remote Sensing*, **42**, 1778-1790.

- Metsaranta, J.M. & Lieffers, V.J. (2008) A fifty-year reconstruction of annual changes in the spatial distribution of *Pinus banksiana* stands: does pattern fit competition theory? *Plant Ecology*, **199**, 137-152.
- Metzger, K.L., Coughenour, M.B., Reich, R.M., & Boone, R.B. (2005) Effects of seasonal grazing on plant species diversity and vegetation structure in a semi-arid ecosystem. *Journal of Arid Environments*, **61**, 147-160.
- Meyer, P., Staenz, K., & Itten, K.I. (1996) Semi-automated procedures for tree species identification in high spatial resolution data from digitized colour infrared-aerial photography. *Journal of Photogrammetry and Remote Sensing*, **51**, 5-16.
- Meyer, K.M., Ward, D., Moustakas, A., & Wiegand, K. (2005) Big is not better: small *Acacia mellifera* shrubs are more vital after fire. *African Journal of Ecology*, **43**, 131-136.
- Meyer, K.M., Wiegand, K., Ward, D., & Moustakas, A. (2007) SATCHMO: a spatial simulation model of growth, competition, and mortality in cycling savanna patches. *Ecological Modelling*, **209**, 377-391.
- Meyer, K.M., Ward, D., Wiegand, K., & Moustakas, A. (2008) Multi-proxy evidence for competition between savanna woody species. *Perspectives in Plant Ecology Evolution and Systematics*, **10**, 63-72.
- Michalet, R. (2006) Is facilitation in arid environments the result of direct or complex interactions? Commentary. *New Phytologist*, **169**, 3-6.
- Michalet, R. (2007) Highlighting the multiple drivers of change in interactions along stress gradients. *New Phytologist*, **173**, 3-6.
- Milton, S.J. & Dean, W.R.J. (2000) Disturbance, drought and dynamics of desert dune grassland, South Africa. *Plant Ecology*, **150**, 37-51.
- Ministerio de Industria y Energia (1978a) *Llano de Uncanca*, Mapa Geologico de España, 1:25,000. Instituto Geologico y minero de España, Madrid.
- Ministerio de Industria y Energia (1978b) *Las Cañadas del Teide*, Mapa Geologico de España, 1:25,000. Instituto Geologico y Minero de Espana, Madrid.
- Ministry of the Environment. (2006) *Proposal to Inscribe Teide National Park on the World Heritage List*. Government of the Canary Islands.
- Miriti, M.N. (2006) Ontogenetic shift from facilitation to competition in a desert shrub. *Journal of Ecology*, **94**, 973-979.
- Monger, H.C. & Bestelmeyer, B.T. (2006) The soil-geomorphic template and biotic change in arid and semi-arid ecosystems. *Journal of Arid Environments*, **65**, 207-218.
- Montes, F., Rubio, A., Barbeito, I., & Cañellas, I. (2008) Characterization of the spatial structure of the canopy in *Pinus silvestris* L. stands in Central Spain from hemispherical photographs. *Forest Ecology and Management*, **255**, 580-590.
- Montesinos, D., Verdu, M., & García-Fayos, P. (2007) Moms are better nurses than dads: gender biased self-facilitation in a dioecious *Juniperus* tree. *Journal of Vegetation Science*, **18**, 271-280.
- Mooney, H.A., Ehleringer, J., & Bjorkman, O. (1977) Energy-balance of leaves of evergreen desert shrub *Atriplex hymenelytra*. *Oecologia*, **29**, 301-310.

- Moravie, M.A. & Robert, A. (2003) A model to assess relationships between forest dynamics and spatial structure. *Journal of Vegetation Science*, **14**, 823-834.
- Moustakas, A., Guenther, M., Wiegand, K., Mueller, K.H., Ward, D., Meyer, K.M., & Jeltsch, F. (2006) Long-term mortality patterns of the deep-rooted *Acacia erioloba*: the middle class shall die! *Journal of Vegetation Science*, **17**, 473-480.
- Moustakas, A., Wiegand, K., Getzin, S., Ward, D., Meyer, K.M., Guenther, M., & Mueller, K.H. (2008) Spacing patterns of an *Acacia* tree in the Kalahari over a 61-year period: how clumped becomes regular and vice versa. *Acta Oecologica*, **33**, 355-364.
- Muller-Landau, H.C., Condit, R.S., Harms, K.E., Marks, C.O., Thomas, S.C., Bunyavejchewin, S., Chuyong, G., Co, L., Davies, S., Foster, R., Gunatilleke, S., Gunatilleke, N., Hart, T., Hubbell, S.P., Itoh, A., Kassim, A.R., Kenfack, D., LaFrankie, J.V., Lagunzad, D., Lee, H.S., Losos, E., Makana, J.R., Ohkubo, T., Samper, C., Sukumar, R., Sun, I.F., Supardi, N.M.N., Tan, S., Thomas, D., Thompson, J., Valencia, R., Vallejo, M.I., Muñoz, G.V., Yamakura, T., Zimmerman, J.K., Dattaraja, H.S., Esufali, S., Hall, P., He, F.L., Hernandez, C., Kiratiprayoon, S., Suresh, H.S., Wills, C., & Ashton, P. (2006) Comparing tropical forest tree size distributions with the predictions of metabolic ecology and equilibrium models. *Ecology Letters*, **9**, 589-602.
- 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.
- Müllerová, J. (2004) Use of digital aerial photography for sub-alpine vegetation mapping: a case study from the Krkonoše Mts., Czech Republic. *Plant Ecology*, **175**, 259-272.
- Muñoz-Mari, J., Bruzzone, L., & Camps-Valls, G. (2007) A support vector domain description approach to supervised classification of remote sensing images. *Institute of Electrical and Electronics Engineers (IEEE) Transactions on Geoscience and Remote Sensing*, **45**, 2683-2692.
- Murrell, D.J., Purves, D.W., & Law, R. (2001) Uniting pattern and process in plant ecology. *Trends in Ecology & Evolution*, **16**, 529-530.
- Murrell, D.J. (2009) On the emergent spatial structure of size-structured populations: when does self-thinning lead to a reduction in clustering? *Journal of Ecology*, **97**, 256-266.
- Myers, B.J. & Benson, M.L. (1981) Rainforest species on large-scale color photos. *Photogrammetric Engineering and Remote Sensing*, **47**, 505-513.
- Nakagawa, S. & Cuthill, I.C. (2007) Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Reviews*, **82**, 591-605.
- Neatrou, M.A., Jones, R.H., & Golladay, S.W. (2007) Response of three floodplain tree species to spatial heterogeneity in soil oxygen and nutrients. *Journal of Ecology*, **95**, 1274-1283.
- Newton, A.C., Hill, R.A., Echeverría, C., Golicher, D., Benayas, J.M.R., Cayuela, L., & Hinsley, S.A. (2009) Remote sensing and the future of landscape ecology. *Progress in Physical Geography*, **33**, 528-546.
- Noy-Meir, I. (1973) Desert ecosystems: environment and producers. *Annual Review of Ecology and Systematics*, **4**, 25-51.

- Nuske, R.S., Sprauer, S., & Saborowski, J. (2009) Adapting the pair-correlation function for analysing the spatial distribution of canopy gaps. *Forest Ecology and Management*, **259**, 107-116.
- Oborny, B. & Cain, M.L. (1997). Models of spatial spread and foraging in clonal plants. In: de Kroon, H. & van Groenendael, J. (eds.) *The Ecology and Evolution of Clonal Plants*. Backhuys Publishers, Leiden, The Netherlands, pp. 155-183.
- Ogle, K. & Reynolds, J.F. (2004) Plant responses to precipitation in desert ecosystems: integrating functional types, pulses, thresholds, and delays. *Oecologia*, **141**, 282-294.
- Okeke, F. & Karnieli, A. (2006) Methods for fuzzy classification and accuracy assessment of historical aerial photographs for vegetation change analyses: Part I - algorithm development. *International Journal of Remote Sensing*, **27**, 153-176.
- Okin, G.S., Roberts, D.A., Murray, B., & Okin, W.J. (2001) Practical limits on hyperspectral vegetation discrimination in arid and semiarid environments. *Remote Sensing of Environment*, **77**, 212-225.
- Okin, G.S. & Roberts, D.A. (2004). Remote Sensing in Arid Environments: Challenges and Opportunities. In: Ustin, S. (ed.) *Manual of Remote Sensing: Remote Sensing for Natural Resource Management and Environmental Monitoring*, Vol. 4. John Wiley and Sons, New York, pp. 111-145.
- Osuna, E., Freund, R., & Girogi, F. (1997) Training support vector machines: an application to face detection. *Proceedings of the Institute of Electrical and Electronics Engineers (IEEE) Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 130-136.
- Pal, M. & Mather, P.M. (2005) Support vector machines for classification in remote sensing. *International Journal of Remote Sensing*, **26**, 1007-1011.
- Palacio, S. & Montserrat-Marti, G. (2007) Above and belowground phenology of four Mediterranean sub-shrubs: preliminary results on root-shoot competition. *Journal of Arid Environments*, **68**, 522-533.
- Park, A. (2003) Spatial segregation of pines and oaks under different fire regimes in the Sierra Madre Occidental. *Plant Ecology*, **169**, 1-20.
- Peck, J.R., Yearsley, J.M., & Waxman, D. (1998) Explaining the geographic distributions of sexual and asexual population. *Nature*, **391**, 889-892.
- Pélissier, R. (1998) Tree spatial patterns in three contrasting plots of a southern Indian tropical moist evergreen forest. *Journal of Tropical Ecology*, **14**, 1-16.
- Pélissier, R. & Goreaud, F. (2001) A practical approach to the study of spatial structure in simple cases of heterogeneous vegetation. *Journal of Vegetation Science*, **12**, 99-108.
- Pérez, F.L. (2000) The influence of surface volcanoclastic layers from Haleakala (Maui, Hawaii) on soil water conservation. *Catena*, **38**, 301-332.
- Pérez, F.L. (2003) Influence of substrate on the distribution of the Hawaiian Silversword (*Argyroxiphium sandwicense* DC.) in Haleakala (Maui, HI). *Geomorphology*, **55**, 173-202.
- Perfecto, I. & Vandermeer, J. (2008) Spatial pattern and ecological process in the coffee agroforestry system. *Ecology*, **89**, 915-920.



- Perry, J.N., Liebhold, A.M., Rosenberg, M.S., Dungan, J., Miriti, M., Jakomulska, A., & Citron-Pousty, S. (2002) Illustration and guidelines for selecting statistical methods for quantifying spatial patterns in ecological data. *Ecography*, **25**, 578-600.
- Perry, G.L.W., Miller, B.P., & Enright, N.J. (2006) A comparison of methods for the statistical analysis of spatial point patterns in plant ecology. *Plant Ecology*, **187**, 59-82.
- Perry, G.L.W., Enright, N.J., Miller, B.P., & Lamont, B.B. (2009) Nearest-neighbour interactions in species-rich shrublands: the roles of abundance, spatial patterns and resources. *Oikos*, **118**, 161-174.
- Peters, D.P.C., Bestelmeyer, B.T., Herrick, J.E., Fredrickson, E.L., Monger, H.C., & Havstad, K.M. (2006) Disentangling complex landscapes: new insights into arid and semiarid system dynamics. *Bioscience*, **56**, 491-501.
- Peters, E.M., Martorell, C., & Ezcurra, E. (2008) Nurse rocks are more important than nurse plants in determining the distribution and establishment of globose cacti (*Mammillaria*) in the Tehuacan Valley, Mexico. *Journal of Arid Environments*, **72**, 593-601.
- Peterson, C.J. & Squiers, E.R. (1995) An unexpected change in spatial pattern across 10 years in an Aspen White-Pine forest. *Journal of Ecology*, **83**, 847-855.
- Phillips, D.L. & Macmahon, J.A. (1981) Competition and spacing patterns in desert shrubs. *Journal of Ecology*, **69**, 97-115.
- Pierson, E.A. & Turner, R.M. (1998) An 85-year study of saguaro (*Carnegiea gigantea*) demography. *Ecology*, **79**, 2676-2693.
- Plotkin, J.B., Potts, M.D., Leslie, N., Manokaran, N., LaFrankie, J., & Ashton, P.S. (2000) Species-area curves, spatial aggregation, and habitat specialization in tropical forests. *Journal of Theoretical Biology*, **207**, 81-99.
- Plotkin, J.B., Chave, J.M., & Ashton, P.S. (2002) Cluster analysis of spatial patterns in Malaysian tree species. *American Naturalist*, **160**, 629-644.
- Popp, A., Blaum, N., & Jeltsch, F. (2009) Ecohydrological feedback mechanisms in arid rangelands: simulating the impacts of topography and land use. *Basic and Applied Ecology*, **10**, 319-329.
- Pottier, J., Marrs, R.H., & Bedecarrats, A. (2007) Integrating ecological features of species in spatial pattern analysis of a plant community. *Journal of Vegetation Science*, **18**, 223-230.
- Pouliot, D.A., King, D.J., Bell, F.W., & Pitt, D.G. (2002) Automated tree crown detection and delineation in high-resolution digital camera imagery of coniferous forest regeneration. *Remote Sensing of Environment*, **82**, 322-334.
- Prentice, I.C. & Werger, J.A. (1985) Clump spacing in a desert dwarf shrub community. *Vegetatio*, **63**, 133-139.
- Pueyo, Y. & Alados, C.L. (2007) Abiotic factors determining vegetation patterns in a semi-arid Mediterranean landscape: different responses on gypsum and non-gypsum substrates. *Journal of Arid Environments*, **69**, 490-505.
- Pueyo, Y., Kefi, S., Alados, C.L., & Rietkerk, M. (2008) Dispersal strategies and spatial organization of vegetation in arid ecosystems. *Oikos*, **117**, 1522-1532.

- Purves, D.W. & Law, R. (2002) Fine-scale spatial structure in a grassland community: quantifying the plant's-eye view. *Journal of Ecology*, **90**, 121-129.
- Qong, M., Takamura, H., & Hudaberdi, M. (2002) Formation and internal structure of *Tamarix* cones in the Taklimakan Desert. *Journal of Arid Environments*, **50**, 81-97.
- R Core Development Team (2009) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rahbek, C. & Graves, G.R. (2001) Multiscale assessment of patterns of avian species richness. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 4534-4539.
- Reisman-Berman, O. (2007) Age-related change in canopy traits shifts conspecific facilitation to interference in a semi-arid shrubland. *Ecography*, **30**, 459-470.
- Reyes-Betancort, J.A., Guerra, A.S., Guma, I.R., Humphries, C.J., & Carine, M.A. (2008) Diversity, rarity and the evolution and conservation of the Canary Islands endemic flora. *Anales del Jardín Botánico de Madrid*, **65**, 25-45.
- Rietkerk, M., Dekker, S.C., de Ruiter, P.C., & van de Koppel, J. (2004) Self-organized patchiness and catastrophic shifts in ecosystems. *Science*, **305**, 1926-1929.
- Rietkerk, M. & van de Koppel, J. (2008) Regular pattern formation in real ecosystems. *Trends in Ecology & Evolution*, **23**, 169-175.
- Ripley, B.D. (1976) 2nd-Order Analysis of Stationary Point Processes. *Journal of Applied Probability*, **13**, 255-266.
- Ripley, B.D. (1981) *Spatial Statistics*. Wiley, New York.
- Ripley, B.D. (2009) MASS. R Package Version 4.44. <http://cran.r-project.org/web/packages/MASS/index.html>.
- Robertson, T.R., Zak, J.C., & Tissue, D.T. (2009) Precipitation magnitude and timing differentially affect species richness and plant density in the sotol grassland of the Chihuahuan Desert. *Oecologia*, **162**, 185-197.
- Rohani, P., Lewis, T.J., Grunbaum, D., & Ruxton, G.D. (1997) Spatial self-organization in ecology: pretty patterns or robust reality? *Trends in Ecology & Evolution*, **12**, 70-74.
- Roiloa, S.R. & Retuerto, R. (2006) Small-scale heterogeneity in soil quality influences photosynthetic efficiency and habitat selection in a clonal plant. *Annals of Botany*, **98**, 1043-1052.
- Roiloa, S.R. & Retuerto, R. (2007) Responses of the clonal *Fragaria vesca* to microtopographic heterogeneity under different water and light conditions. *Environmental and Experimental Botany*, **61**, 1-9.
- Roiloa, S.R., Alpert, P., Tharayil, N., Hancock, G., & Bhowmik, P.C. (2007) Greater capacity for division of labour in clones of *Fragaria chiloensis* from patchier habitats. *Journal of Ecology*, **95**, 397-405.
- Rossi, J.P., Samalens, J.C., Guyon, D., van Halder, I., Jactel, H., Menassieu, P., & Piou, D. (2009) Multiscale spatial variation of the bark beetle *Ips sexdentatus* damage in a pine plantation forest (Landes de Gascogne, Southwestern France). *Forest Ecology and Management*, **257**, 1551-1557.

- Ryel, R.J., Leffler, A.J., Peek, M.S., Ivans, C.Y., & Caldwell, M.M. (2004) Water conservation in *Artemisia tridentata* through redistribution of precipitation. *Oecologia*, **141**, 335-345.
- Ryel, R.J., Ivans, C.Y., Peek, M.S., & Leffler, A.J. (2008). Functional differences in soil water pools: a new perspective on plant water use in water-limited ecosystems. In: Lüttge, U., Beyschlag, W. & Murata, J. (eds.) *Progress in Botany*. Springer-Verlag, Berlin, pp. 397-422.
- Saco, P.M., Willgoose, G.R., & Hancock, G.R. (2007) Eco-geomorphology of banded vegetation patterns in arid and semi-arid regions. *Hydrology and Earth System Sciences*, **11**, 1717-1730.
- Salisbury, F.B. & Ross, C.W. (1992) *Plant Physiology*, 4<sup>th</sup> edition. Wadsworth Publishing, Belmont, California.
- Sanchez-Hernandez, C. (2006) *Land cover mapping of one specific protected habitat under the requirements of the European Union Habitats Directive*. Unpublished PhD Thesis, Kingston University, Kingston upon Thames.
- Sanchez-Hernandez, C., Boyd, D.S., & Foody, G.M. (2007) One-class classification for mapping a specific land-cover class: SVDD classification of fenland. *Institute of Electrical and Electronics Engineers (IEEE) Transactions on Geoscience and Remote Sensing*, **45**, 1061-1073.
- Sandquist, D.R. & Ehleringer, J.R. (1998) Intraspecific variation of drought adaptation in brittlebush: leaf pubescence and timing of leaf loss vary with rainfall. *Oecologia*, **113**, 162-169.
- Sankaran, M., Hanan, N.P., Scholes, R.J., Ratnam, J., Augustine, D.J., Cade, B.S., Gignoux, J., Higgins, S.I., Le Roux, X., Ludwig, F., Ardo, J., Banyikwa, F., Bronn, A., Bucini, G., Caylor, K.K., Coughenour, M.B., Diouf, A., Ekaya, W., Feral, C.J., February, E.C., Frost, P.G.H., Hiernaux, P., Hrabar, H., Metzger, K.L., Prins, H.H.T., Ringrose, S., Sea, W., Tews, J., Worden, J., & Zambatis, N. (2005) Determinants of woody cover in African savannas. *Nature*, **438**, 846-849.
- Santos-Guerra, A. (2001). Flora vascular nativa. In: Fernández-Palacios, J.M. & Martín-Esquivel, J.L. (eds.) *Naturaleza de la Islas Canarias: Ecología y conservación*. Publicaciones Turquesa, Santa Cruz de Tenerife, pp. 185-192.
- Scanlon, T.M., Caylor, K.K., Levin, S.A., & Rodriguez-Iturbe, I. (2007) Positive feedbacks promote power-law clustering of Kalahari vegetation. *Nature*, **449**, 209-212.
- Schenk, H.J. (1999) Clonal splitting in desert shrubs. *Plant Ecology*, **141**, 41-52.
- Schenk, H.J. & Jackson, R.B. (2002) Rooting depths, lateral root spreads and below-ground/above-ground allometries of plants in water-limited ecosystems. *Journal of Ecology*, **90**, 480-494.
- Schenk, H.J., Holzapfel, C., Hamilton, J.G., & Mahall, B.E. (2003) Spatial ecology of a small desert shrub on adjacent geological substrates. *Journal of Ecology*, **91**, 383-395.
- Schmidt, K.S. & Skidmore, A.K. (2003) Spectral discrimination of vegetation types in a coastal wetland. *Remote Sensing of Environment*, **85**, 92-108.
- Schneider, D.C. (2001) The rise of the concept of scale in ecology. *Bioscience*, **51**, 545-553.

- Schurr, F.M., Bossdorf, O., Milton, S.J., & Schumacher, J. (2004) Spatial pattern formation in semi-arid shrubland: a priori predicted versus observed pattern characteristics. *Plant Ecology*, **173**, 271-282.
- Schwinning, S., Sala, O.E., Loik, M.E., & Ehleringer, J.R. (2004) Thresholds, memory, and seasonality: understanding pulse dynamics in arid/semi-arid ecosystems. *Oecologia*, **141**, 191-193.
- Scholes, R.J. & Archer, S.R. (1997) Tree-grass interactions in savannas. *Annual Review of Ecology and Systematics*, **28**, 517-544.
- Seguela, J.J.B. & Trujillo, F.S.D. (2004). Clima. In: Canseco, V.G. (ed.) *Parque Nacional del Teide*. Canseco Editores, Talavera de la Reina, Spain, pp. 73-96.
- Seidler, T.G. & Plotkin, J.B. (2006) Seed dispersal and spatial pattern in tropical trees. *Plos Biology*, **4**, 2132-2137.
- Shackleton, C. (2002) Nearest-neighbour analysis and the prevalence of woody plant competition in South African savannas. *Plant Ecology*, **158**, 65-76.
- Sharp, B.R. & Whittaker, R.J. (2003) The irreversible cattle-driven transformation of a seasonally flooded Australian savanna. *Journal of Biogeography*, **30**, 783-802.
- Shimatani, K. & Kubota, Y. (2004) Spatial analysis for continuously changing point patterns along a gradient and its application to an *Abies sachalinensis* population. *Ecological Modelling*, **180**, 359-369.
- Shin, H.J., Eom, D.H., & Kim, S.S. (2005) One-class support vector machines: an application in machine fault detection and classification. *Computers & Industrial Engineering*, **48**, 395-408.
- Shreve, F. & Hinckley, A.L. (1937) Thirty years of change in desert vegetation. *Ecology*, **18**, 463-478.
- Simberloff, D. (1979) Nearest neighbor assessments of spatial configurations of circles rather than points. *Ecology*, **60**, 679-685.
- Snyder, K.A., Donovan, L.A., James, J.J., Tiller, R.L., & Richards, J.H. (2004) Extensive summer water pulses do not necessarily lead to canopy growth of Great Basin and northern Mojave Desert shrubs. *Oecologia*, **141**, 325-334.
- Snyder, K.A. & Tartowski, S.L. (2006) Multi-scale temporal variation in water availability: implications for vegetation dynamics in arid and semi-arid ecosystems. *Journal of Arid Environments*, **65**, 219-234.
- Song, M.H., Dong, M., & Jiang, G.M. (2002) Importance of clonal plants and plant species diversity in the Northeast China Transect. *Ecological Research*, **17**, 705-716.
- Steinberg, E.K. & Kareiva, P. (1997). Challenges and opportunities for empirical evaluation of 'spatial theory'. In: Tilman, D. & Kareiva, P. (eds.) *Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions*. Princeton University Press, Princeton, New York, pp. 318-332.
- Sthultz, C.M., Gehring, C.A., & Whitham, T.G. (2007) Shifts from competition to facilitation between a foundation tree and a pioneer shrub across spatial and temporal scales in a semiarid woodland. *New Phytologist*, **173**, 135-145.

- Stoll, P. & Bergius, E. (2005) Pattern and process: competition causes regular spacing of individuals within plant populations. *Journal of Ecology*, **93**, 395-403.
- Stoyan, D. & Stoyan, H. (1994) *Fractals, Random Shapes and Point Fields: Methods of Geometrical Statistics*. John Wiley and Sons, Chichester.
- Stoyan, D. & Penttinen, A. (2000) Recent applications of point process methods in forestry statistics. *Statistical Science*, **15**, 61-78.
- Strand, E.K., Smith, A.M.S., Bunting, S.C., Vierling, L.A., Hann, D.B., & Gessler, P.E. (2006) Wavelet estimation of plant spatial patterns in multitemporal aerial photography. *International Journal of Remote Sensing*, **27**, 2049-2054.
- Strand, E.K., Robinson, A.P., & Bunting, S.C. (2007) Spatial patterns on the sagebrush steppe/Western juniper ecotone. *Plant Ecology*, **190**, 159-173.
- Sventenius, E.R.S. (1946) Notas sobre la flora de Las Cañadas de Tenerife. *Boletín del Instituto Nacional de Investigaciones Agronómicas*, **15**, 149-171.
- Svoray, T., Karnieli, A., & Dedieu, G. (2008) Satellite evidence of a topographic sink-source system in a small semi-arid watershed. *International Journal of Remote Sensing*, **29**, 609-616.
- Tanaka, H. & Nakashizuka, T. (1997) Fifteen years of canopy dynamics analyzed by aerial photographs in a temperate deciduous forest, Japan. *Ecology*, **78**, 612-620.
- Tax, D.M.J. (2001) *One-class classification: Concept-learning in the absence of counter-examples*. Unpublished PhD Thesis, Delft University of Technology, The Netherlands.
- Tax, D.M.J. & Juszczak, P. (2003) Kernel whitening for one-class classification. *International Journal of Pattern Recognition and Artificial Intelligence*, **17**, 333-347.
- Tax, D.M.J. & Duin, R.P.W. (2004) Support vector data description. *Machine Learning*, **54**, 45-66.
- Tax, D.M.J. (2008) *DDtools, The Data Description Toolbox for Matlab*. version 1.4.1, [http://ict.ewi.tudelft.nl/~davidt/dd\\_tools.html](http://ict.ewi.tudelft.nl/~davidt/dd_tools.html)
- Tejedor, M., Jiménez, C.C., & Díaz, F. (2002) Soil moisture regime changes in tephra-mulched soils: Implications for soil taxonomy. *Soil Science Society of America Journal*, **66**, 202-206.
- Tejedor, M., Jiménez, C., & Díaz, F. (2003) Volcanic materials as mulches for water conservation. *Geoderma*, **117**, 283-295.
- Thompson, H.R. (1956) Distribution of distance to the nth neighbour in a population of randomly distribution individuals. *Ecology*, **37**, 391-394.
- Tielbörger, K. & Kadmon, R. (1997) Relationships between shrubs and annual communities in a sandy desert ecosystem: a three-year study. *Plant Ecology*, **130**, 191-201.
- Tielbörger, K. & Kadmon, R. (2000) Temporal environmental variation tips the balance between facilitation and interference in desert plants. *Ecology*, **81**, 1544-1553.
- Tilman, D., Lehman, C.L., & Kareiva, P. (1997). Population dynamics in spatial habitats. In: Tilman, D. & Kareiva, P. (eds.) *Spatial Ecology: The Role of*

- Tiwari, A.K. & Singh, J.S. (1984) Mapping forest biomass in India through aerial photographs and nondestructive field sampling. *Applied Geography*, **4**, 151-165.
- Toft, C.A. & Fraizer, T. (2003) Spatial dispersion and density dependence in a perennial desert shrub (*Chrysothamnus nauseosus* : Asteraceae). *Ecological Monographs*, **73**, 605-624.
- Trichon, V. (2001) Crown typology and the identification of rain forest trees on large-scale aerial photographs. *Plant Ecology*, **153**, 301-312.
- Trichon, V. & Julien, M.P. (2006) Tree species identification on large-scale aerial photographs in a tropical rain forest, French Guiana: application for management and conservation. *Forest Ecology and Management*, **225**, 51-61.
- Tuda, M. (2007) Understanding mechanism of spatial ecological phenomena: a preface to the special feature on "spatial statistics". *Ecological Research*, **22**, 183-184.
- Turkington, R. & Harper, J.L. (1979) Growth, distribution and neighbor relationships of *Trifolium repens* in a permanent pasture : Part 4 - fine-scale biotic differentiation. *Journal of Ecology*, **67**, 245-254.
- Turner, R.M. (1990) Long-term vegetation change at a fully protected Sonoran Desert site. *Ecology*, **71**, 464-477.
- Turner, I.M., Wong, Y.K., Chew, P.T., & BinIbrahim, A. (1996) Rapid assessment of tropical rain forest successional status using aerial photographs. *Biological Conservation*, **77**, 177-183.
- Turner, M.G., Gardner, R.H., & O'Neill, R.V. (2001) *Landscape Ecology in Theory and Practice: Pattern and Process*. Springer-Verlag, New York.
- Tylianakis, J.M., Klein, A.M., Lozada, T., & Tscharntke, T. (2006) Spatial scale of observation affects alpha, beta and gamma diversity of cavity-nesting bees and wasps across a tropical land-use gradient. *Journal of Biogeography*, **33**, 1295-1304.
- Valiente-Banuet, A. & Verdú, M. (2008) Temporal shifts from facilitation to competition occur between closely related taxa. *Journal of Ecology*, **96**, 489-494.
- Vapnik, V.N. (1995) *The Nature of Statistical Learning Theory*. Springer-Verlag, New York.
- Vapnik, V.N. & Chervonenkis, A.Y. (1971) Theory of uniform convergence of frequencies of events to their probabilities and problems of search for an optimal solution from empirical data. *Automation and Remote Control*, **32**, 207-217.
- Vasek, F.C. (1980) Creosote bush: long-lived clones in the Mojave Desert. *American Journal of Botany*, **67**, 246-255.
- Verheyden, A., Dahdouh-Guebas, F., Thomaes, K., De Genst, W., Hettiarachchi, S., & Koedam, N. (2002) High-resolution vegetation data for mangrove research as obtained from aerial photography. *Environment Development and Sustainability*, **4**, 113-133.

- Verhulst, J., Montaña, C., Mandujano, M.C., & Franco, M. (2008) Demographic mechanisms in the coexistence of two closely related perennials in a fluctuating environment. *Oecologia*, **156**, 95-105.
- Vinuesa, P., Silva, C., Werner, D., & Martínez-Romero, E. (2005a) Population genetics and phylogenetic inference in bacterial molecular systematics: the roles of migration and recombination in *Bradyrhizobium* species cohesion and delineation. *Molecular Phylogenetics and Evolution*, **34**, 29-54.
- Vinuesa, P., León-Barrios, M., Silva, C., Willems, A., Jarabo-Lorenzo, A., Pérez-Galdona, R., Werner, D., & Martínez-Romero, E. (2005b) *Bradyrhizobium canariense* sp. nov., an acid-tolerant endosymbiont that nodulates endemic genistoid legumes (*Papilionoideae* : Genisteae) from the Canary Islands, along with *Bradyrhizobium japonicum* bv. *genistearum*, *Bradyrhizobium* *genospecies* alpha and *Bradyrhizobium* *genospecies* beta. *International Journal of Systematic and Evolutionary Microbiology*, **55**, 569-575.
- Wagner, H.H. & Fortin, M.J. (2005) Spatial analysis of landscapes: concepts and statistics. *Ecology*, **86**, 1975-1987.
- Walker, B.H. & Langridge, J.L. (1997) Predicting savanna vegetation structure on the basis of plant available moisture (PAM) and plant available nutrients (PAN): a case study from Australia. *Journal of Biogeography*, **24**, 813-825.
- Walter, H. (1971) *Ecology of tropical and subtropical vegetation*. Oliver and Boyd, Oxford.
- Wang, X.G., Hao, Z.Q., Zhang, J., Lian, J.Y., Li, B.H., Ye, J., & Yao, X.L. (2009) Tree size distributions in an old-growth temperate forest. *Oikos*, **118**, 25-36.
- Ward, D., Saltz, D., & Olsvig-Whittaker, L. (2000) Distinguishing signal from noise: long-term studies of vegetation in Makhtesh Ramon erosion cirque, Negev desert, Israel. *Plant Ecology*, **150**, 27-36.
- Ward, D. (2009) *The Biology of Deserts*. Oxford University Press, Oxford.
- Watt, A.S. (1947) Pattern and process in the plant community. *Journal of Ecology*, **35**, 1-22.
- Weedon, J.T. & Facelli, J.M. (2008) Desert shrubs have negative or neutral effects on annuals at two levels of water availability in arid lands of South Australia. *Journal of Ecology*, **96**, 1230-1237.
- Wells, M.L. & Getis, A. (1999) The spatial characteristics of stand structure in *Pinus torreyana*. *Plant Ecology*, **143**, 153-170.
- Wesche, K., Ronnenberg, K., & Hensen, I. (2005) Lack of sexual reproduction within mountain steppe populations of the clonal shrub *Juniperus sabina* L. in semi-arid southern Mongolia. *Journal of Arid Environments*, **63**, 390-405.
- West, A.G., Hultine, K.R., Burtch, K.G., & Ehleringer, J.R. (2007) Seasonal variations in moisture use in a piñon-juniper woodland. *Oecologia*, **153**, 787-798.
- Wheeler, C.T. & Dickson, J.H. (1990) Symbiotic nitrogen fixation and distribution of *Spartocytisus supranubius* on Las Cañadas, Tenerife. *Vieraea*, **19**, 309-314.
- Whitford, W.G. (2002) *Ecology of desert systems*. Academic Press, London.
- Whittaker, R.J., Fernández-Palacois, J. M. (2007) *Island Biogeography: Ecology, Evolution and Conservation*. Oxford University Press, Oxford.

- Wiegand, T. & Jeltsch, F. (2000) Long-term dynamics in arid and semiarid ecosystems: synthesis of a workshop. *Plant Ecology*, **150**, 3-6.
- Wiegand, T. (2004) *Introduction to point pattern analysis with Ripley's L and the O-ring statistic using the Programita software*. Unpublished user manual, Helmholtz Centre for Environmental Research, Leipzig, Germany.
- Wiegand, T. & Moloney, K.A. (2004) Rings, circles, and null-models for point pattern analysis in ecology. *Oikos*, **104**, 209-229.
- Wiegand, T., Kissling, W.D., Cipriotti, P.A., & Aguiar, M.R. (2006) Extending point pattern analysis for objects of finite size and irregular shape. *Journal of Ecology*, **94**, 825-837.
- Wiegand, T., Gunatilleke, S., & Gunatilleke, N. (2007a) Species associations in a heterogeneous Sri lankan dipterocarp forest. *American Naturalist*, **170**, E67-E95 (E-article).
- Wiegand, T., Gunatilleke, S., Gunatilleke, N., & Okuda, T. (2007b) Analyzing the spatial structure of a Sri Lankan tree species with multiple scales of clustering. *Ecology*, **88**, 3088-3102.
- Wiegand, T., Martinez, I., & Huth, A. (2009) Recruitment in tropical tree species: revealing complex spatial patterns. *American Naturalist*, **174**, E106-E140 (E-article).
- Wiens, J.A. (1999). The science and practice of landscape ecology. In: Klopach, J.M. & Gardner, R.H. (eds.) *Landscape Ecological Analysis*. Springer-Verlag, New York, pp. 371-383.
- Wilcox, B.P., Breshears, D.D., & Allen, C.D. (2003) Ecohydrology of a resource-conserving semiarid woodland: effects of scale and disturbance. *Ecological Monographs*, **73**, 223-239.
- Wilcox, B.P., Owens, M.K., Dugas, W.A., Ueckert, D.N., & Hart, C.R. (2006a) Shrubs, streamflow, and the paradox of scale. *Hydrological Processes*, **20**, 3245-3259.
- Wilcox, B.P., Dowhower, S.L., Teague, W.R., & Thurow, T.L. (2006b) Long-term water balance in a semiarid shrubland. *Rangeland Ecology & Management*, **59**, 600-606.
- Wilson, E. (1920) Use of aircraft in forestry and logging. *Canadian Forestry Magazine*, **16**, 439-444.
- Wilson, D.J., Western, A.W., & Grayson, R.B. (2004) Identifying and quantifying sources of variability in temporal and spatial soil moisture observations. *Water Resources Research*, **40**, W02507.0–W02507.10.
- Wu, X.B., Thurow, T.L., & Whisenant, S.G. (2000) Fragmentation and changes in hydrologic function of tiger bush landscapes, south-west Niger. *Journal of Ecology*, **88**, 790-800.
- Ye, X.H., Yu, F.H., & Dong, M. (2006) A trade-off between guerrilla and phalanx growth forms in *Leymus secalinus* under different nutrient supplies. *Annals of Botany*, **98**, 187-191.
- Yu, H., Wiegand, T., Yang, X.H., & Ci, L.J. (2009) The impact of fire and density-dependent mortality on the spatial patterns of a pine forest in the Hulun Buir sandland, Inner Mongolia, China. *Forest Ecology and Management*, **257**, 2098-2107.



- Zavala-Hurtado, J.A., Valverde, P.L., Herrera-Fuentes, M.D., & Díaz-Solís, A. (2000) Influence of leaf-cutting ants (*Atta mexicana*) on performance and dispersion patterns of perennial desert shrubs in an inter-tropical region of Central Mexico. *Journal of Arid Environments*, **46**, 93-102.
- Zeng, Z.H., Fu, Y., Roisman, G.I., Wen, Z., Hu, Y.X., & Huang, T.S. (2006). One-class classification for spontaneous facial expression analysis. *Proceedings of the 7<sup>th</sup> International Institute of Electrical and Electronics Engineers (IEEE) Computer Society Conference on Automatic Face and Gesture Recognition*, pp. 281-286.
- Zenner, E.K. & Peck, J.E. (2009) Characterizing structural conditions in mature managed red pine: spatial dependency of metrics and adequacy of plot size. *Forest Ecology and Management*, **257**, 311-320.
- Zhu, G.B. & Blumberg, D.G. (2002) Classification using ASTER data and SVM algorithms: the case study of Beer Sheva, Israel. *Remote Sensing of Environment*, **80**, 233-240.
- Zhang, Y., Zhao, Y., Zhang, C., & Zhao, X. (2009) Structure and spatial distribution of *Pinus tabulaeformis* population in the Songshan Nature Reserve, Beijing, China. *Chinese Journal of Applied and Environmental Biology*, **15**, 175-179.
- Zhu, Y., Mi, X., Ren, H., & Ma, K. (2010) Density dependence is prevalent in a heterogeneous subtropical forest. *Oikos*, **119**, 109-119.
- Zhuang, L. & Dai, H.H. (2006). Parameter estimation of one-class SVM on imbalance text classification. *Proceedings of the 19th Conference of the Canadian Society for Computational Studies of Intelligence*, pp. 538-549.
- Zou, T., Li, Y., Xu, H., & Xu, G.Q. (2010) Responses to precipitation treatment for *Haloxylon ammodendron* growing on contrasting textured soils. *Ecological Research*, **25**, 185-194.

APPENDICES

---

APPENDIX A: A REVIEW OF ARTICLES USING SECOND-ORDER SPATIAL STATISTICS TO INVESTIGATE THE SPATIAL PATTERN OF WOODY

PLANTS

General			Data collection										Pattern analysis		
Focal species / community	Life form	Arid environment	Data collection technique	Data representation	Plot extents (range) (zdp)	Plot extent (geometric mean) (zdp)	Number of spatial environmental contexts	Environmental conditions compared (i.e., when the number of environmental contexts is > 1)	Replicate plots per environmental context (range)	Replicate plots per environmental context (geometric mean)	Sample size (range)	Sample size (geometric mean)	Function	Interpretative technique	Reference
Subtropical forest	T & S		F	P	24	24	1		1	1	—	—	L & g	MC	1
<i>Euterpe edulis</i>	T		F	P	0.64	0.64	1		1	1	29	29	g	MC	2
<i>Pinus sylvestris</i>	T		F	P	0.5	0.5	4	Homogeneous and heterogeneous and two management regimes (2*2)	1	1	—	—	g	H	3
<i>Baccharis dracunculifolia</i> & <i>Eupatorium buniifolium</i>	S	✓	F	P	0.05 & 0.06	0.05	2	Fire histories	1	1	112 - 554	310	O	MC	4
<i>Eurycorymbus caraleriei</i>	T		F	P	—	—	2	Fragmented vs. continuous habitat	3	3	65 - 218	122	L	MC	5
Deciduous boreal woodland	T		F	P	1	1	4	Forest type	1	1	—	—	L	MC	6
Tropical rainforest trees	T		F	P	20	20	1		1	1	—	—	L	MC	7
Tropical trees	T		F	P	50	50	1		1	1	—	—	L	MC	8

General			Data collection										Pattern analysis		
Focal species / community	Life form	Arid environment	Data collection technique	Data representation	Plot extents (range) (2dp)	Plot extent (geometric mean) (2dp)	Number of spatial environmental contexts	Environmental conditions compared (i.e., when the number of environmental contexts is > 1)	Replicate plots per environmental context (range)	Replicate plots per environmental context (geometric mean)	Sample size (range)	Sample size (geometric mean)	Function	Interpretative technique	Reference
<i>Pseudotsuga menziesii</i> var <i>glauca</i>	T		F	P	0.05 - 0.1	0.08	1		6	6	107 - 280	175	L & g	H	9
Seasonally dry forest	T & S		F	P	1	1.00	1		1	1	36 – 249 <sup>h</sup>	68	L	MC	10
<i>Pinus ponderosa</i>	T		F	P	2.59	2.59	1		1	1	134, 219, 1487	352	L	MC	11
Tropical rainforest trees	T		F	P	6.25	6.25	1		1	1	3840	3840	L	MC	12
Pine plantation forest	T		C	P	—	—	1		23, 25	24	—	—	L	MC	13
Deciduous forest	T		F	P	0.24 - 3.79	1.06	5	Human influence	1	1	41 - 186	106	L	MC	14
Mongolian pine forest	T		F	P	1	1	3	Fire histories	1	1	56 - 436	190	g	MC	15
<i>Abies forrestii</i> var <i>georgi</i> & <i>Juniperus saltuaria</i>	T		F	P	0.25 & 0.09	0.15	4	Treeline vs. timberline, north vs. south (2*2)	1	1	7 - 264	41	K	MC	16
<i>Pinus pinea</i>	T		F	P & O	0.48	0.48	2	Even aged vs. mixed ages stand	1	1	40 - 1048	267	L & g	MC	17
<i>Sequoia sempervirens</i> (Coast redwood)	T		F	P	1.5	1.5	3	Fire history and disturbance	1	1	—	—	L	MC	18
<i>Vitellaria paradoxa</i>	T		F	P	1	1	2	Farmed vs. protected land	7	7	—	—	g	MC	19

General			Data collection										Pattern analysis		
Focal species / community	Life form	Arid environment	Data collection technique	Data representation	Plot extents (range) (zdp)	Plot extent (geometric mean) (zdp)	Number of spatial environmental contexts	Environmental conditions compared (i.e., when the number of environmental contexts is > 1)	Replicate plots per environmental context (range)	Replicate plots per environmental context (geometric mean)	Sample size (range)	Sample size (geometric mean)	Function	Interpretative technique	Reference
Western hemlock	T		F	P	0.7 - 0.9	0.8	2	Habitat heterogeneity	1	1	101 - 403	209	L & g	H	20
<i>Picea abies</i> & <i>Abies alba</i>	T		F	P	0.25	0.25	1		2	2	322 - 932	549	K	MC	21
African rain forest	T		F	P	4	4	1		1	1	—	—	O	MC	22
<i>Castanopsis chinensis</i> , <i>Schinus molle</i> & <i>Engelhardtia roxburghiana</i>	T		F	P	20	20	1		1	1	—	—	L	MC	23
<i>Acer Saccharum</i>	T		F	P	5.04	5.04	1		1	1	c. 150 – 600°	—	L	MC	24
<i>Quercus petraea</i>	T		F	P	0.21	0.21	2	Thinning conditions	1,2	1	29 - 138	62	L	MC	25
<i>Pinus Banksiana</i>	T		F	P	0.09	0.09	4	Shield vs. plant, mesic vs. xeric (2*2)	1	1	—	—	L	MC	26
<i>Acacia mellifera</i>	S	✓	F	P	0.02	0.02	1		20	20	<sup>h</sup> —	—	g	MC	27
<i>Pinus silvestris</i>	T		C <sup>a</sup>	P	0.5	0.05	5	Silvicultural practices	1	1	—	—	L	MC	28
<i>Acacia erioloba</i>	T	✓	C	P	149 & 197	171.33	1		2	2	230 - 1078	528	O	MC	29

General			Data collection										Pattern analysis		
Focal species / community	Life form	Arid environment	Data collection technique	Data representation	Plot extents (range) (2dp)	Plot extent (geometric mean) (2dp)	Number of spatial environmental contexts	Environmental conditions compared (i.e., when the number of environmental contexts is > 1)	Replicate plots per environmental context (range)	Replicate plots per environmental context (geometric mean)	Sample size (range)	Sample size (geometric mean)	Function	Interpretative technique	Reference
Abies forest	T		F	P	0.01	0.01	1		1	1	—	—	g	MC	30
Temperate forest	T		F	P	1.3 & 1.96	1.6	2	Upland hardwood vs. pine forest	1	1	—	—	L	MC	31
Sycamore & ash	T		C	P	14	14	1		1	1	—	—	g	MC	32
Juniperus virginiana & Lonicera maackii	T & S		F	P	1	1	1		1	1	—	—	L & g	MC	33
Tropical forest	T		F	P	0.45	0.45	2	Late successional vs. secondary forest	2, 1	1	10 - 201	51	L	MC	34
Deciduous & coniferous woodland	T		F	P	0.04 - 0.14	0.09 <sup>d</sup>	4	Conifer vs. deciduous, different slopes (2*2)	1	1	41 - 103	75	g	MC	35
Temperate forest	T		F	P	25	25	1		1	1	—	—	O	MC	36
Carapa guianensis	T		F	P	16	16	2	Forest types	2	2	93 - 264	167	L	MC	37
Conifer & broadleaved forest	T		F	P	2.25	2.25	1		1	1	—	—	L	MC	38
Arid shrubs	S	✓	C	P	0.25	0.25	3	Disturbance types	1	1	10 - 118	21	L	MC	39
Pinus sylvestris	T		F	P	0.8 - 1	0.93	3	Plantation vs. semi-natural, and young vs. old	1	1	—	—	L	MC	40

General			Data collection										Pattern analysis		Reference
Focal species / community	Life form	Arid environment	Data collection technique	Data representation	Plot extents (range) (2dp)	Plot extent (geometric mean) (2dp)	Number of spatial environmental contexts	Environmental conditions compared (i.e., when the number of environmental contexts is > 1)	Replicate plots per environmental context (range)	Replicate plots per environmental context (geometric mean)	Sample size (range)	Sample size (geometric mean)	Function	Interpretative technique	
<i>Juniperus</i>	T		F	P	1	1	1		1	1	269	269	L	MC	41
<i>Juniperus occidentalis</i>	T & S	✓	C	P	10	10	1		1	1	77 - 600	226	L & g	MC	42
<i>Shorea congestiflora</i>	T		F	P	25	25	1		1	1	—	—	g	MC	43
Dipterocarp forest	T		F	P	25	25	1		1	1	— <sup>j</sup>	—	g	MC	44
<i>Picea–Betula</i> forest	T		F	P	0.3 – 0.4	0.37	4	Stages in forest cycle	1	1	—	—	K	MC	45
<i>Pinus ponderosa</i> & douglas fir	T		F <sup>b</sup>	P	0.11	0.11	1		10	10	6 - 351	59	L	MC	46
Douglas fir forest	T		F	P	0.5 – 1.2	0.74	1		1	1	98 - 1061	397	g	MC	47
<i>Abies–Tsuga</i> forest	T		F	P	1	1	1		1	1	—	—	L	MC	48
Deciduous lowland forest	T		F	P	0.25 - 0.55	0.32	2	Parkland vs. forest	2	2	19 - 307	—	L	MC	49
<i>Nothofagus obliqua</i> forest	T		F	P	1	1	1		1	1	17 - 661	102	L	MC	50
Tropical trees	T		F	P	50	50	1		1	1	—	—	K	H	51

General			Data collection										Pattern analysis		Reference
Focal species / community	Life form	Arid environment	Data collection technique	Data representation	Plot extents (range) (zdp)	Plot extent (geometric mean) (zdp)	Number of spatial environmental contexts	Environmental conditions compared (i.e., when the number of environmental contexts is > 1)	Replicate plots per environmental context (range)	Replicate plots per environmental context (geometric mean)	Sample size (range)	Sample size (geometric mean)	Function	Interpretative technique	
<i>Pinus ponderosa</i>	T		F	P	9.3	9.3	1		1	1	—	—	L	MC	52
<i>Pinus sylvestris</i>	T		F	P	0.02 - 0.09	0.04	1		7	7	—	—	K	MC	53
<i>Pinus uncinata</i>	T		F	P	0.2	0.2	1		1	1	26 - 313	89	L	MC	54
Forest	T		F <sup>b</sup>	P	0.25 - 1	0.79	3	Forest types (mesic to xeric)	3,1,1	1	35 - 1816	184	L	MC	55
<i>Nothofagus glauca</i>	T		F	P	0.12 - 0.14	0.13	3	Disturbance types	2	2	23 - 463	101	L	MC	56
<i>Acer platanoides</i>	T		F	P	0.5	0.5	1		1	1	—	—	L	H	57
Deciduous woodland	T		C	P	8	8	1		1	1	—	—	L	MC	58
<i>Pifon-Juniper woodlands</i>	T		F <sup>b</sup>	P	0.78 - 1.56	1.1	3	Soil types	2	2	—	—	L	MC	59
European larch and Swiss stone pine	T		F	P	1	1	1		1	1	—	—	L	MC	60
<i>Pinus nigra</i>	T		F	P	0.1	0.1	3	Growth rates	1	1	—	—	L	MC	61
Mixed deciduous forest	T		F	P	4.3 & 4.4	4.35	1		2	2	—	—	L	MC	62



General			Data collection										Pattern analysis		Reference
Focal species / community	Life form	Arid environment	Data collection technique	Data representation	Plot extents (range) (zdp)	Plot extent (geometric mean) (zdp)	Number of spatial environmental contexts	Environmental conditions compared (i.e., when the number of environmental contexts is > 1)	Replicate plots per environmental context (range)	Replicate plots per environmental context (geometric mean)	Sample size (range)	Sample size (geometric mean)	Function	Interpretative technique	
Shrubs & trees	S & T		F	P	0.49	0.49	1		1	1	—	—	L	MC	63
<i>Quercus liaotungensis</i> , <i>Betula dahurica</i> , <i>Acer mono</i>	T		F	P	5	5	2	Homogeneous subplots (i.e., different tree densities)	1	1	32 - 930	224	L	MC	64
<i>Shorea leprosula</i> & <i>Shorea ovalis</i> spp. <i>Sericea</i>	T		F	P	50	50	1		1	1	54 - 64	58	L	MC	65
Western hemlock–douglas fir and mixed-conifer forest	T		F	P	8.4 & 12	10	2	Forest types	1	1	—	—	K	MC	66
Semi-arid Karoo shrubland	S	✓	F	P	0.02	0.02	1		1	1	158 - 683	371	L & g	MC	67
<i>Pinus ponderosa</i>	T		F <sup>b</sup>	P	1	1	3	Landscapes	5, 10, 12	8	—	—	L	MC	68
Mixed conifer stands	T		F	P	1	1	1		1	1	—	—	L	MC	69
Temperate hardwood forest	T		F	P	3.24	3.24	1		2	2	10 – 306 <sup>f</sup>	35	L	MC	70
Laurel forest	T		F	P	0.06	0.06	2	Altitude and aspect	4	4	— <sup>f</sup>	—	K	MC	71
<i>Ailanthus altissima</i> & <i>Robinia pseudoacacia</i>	T		F	P	0.25	0.25	3	Disturbance regimes	2	2	—	—	L	MC	72
Trees	T	✓	F	P	0.25 - 1	0.5	1		10	10	—	—	L	MC	73

General			Data collection										Pattern analysis		
Focal species / community	Life form	Arid environment	Data collection technique	Data representation	Plot extents (range) (zdp)	Plot extent (geometric mean) (zdp)	Number of spatial environmental contexts	Environmental conditions compared (i.e., when the number of environmental contexts is > 1)	Replicate plots per environmental context (range)	Replicate plots per environmental context (geometric mean)	Sample size (range)	Sample size (geometric mean)	Function	Interpretative technique	
			Reference												
Arid shrubs	S	✓	C	P	0.2	0.2	3	Successional stages	1	1	42 – 222 <sup>h</sup>	62	L	MC	74
Quercus & Acer rubrum	T		F	P	1.96	1.96	1		1	1	—	—	L	H	75
Wet evergreen forest	T		F	P	0.49	0.49	1		1	1	—	—	L	MC	76
Pine & oak	T		F	P	0.25 - 0.36	0.33	3	Fire histories	2, 2, 1	2	20 - 137	46	L	MC	77
Ambrosia dumosa	S	✓	F <sup>C</sup>	P	0.04	0.04	2	Geological substrates	4, 6	5	—	—	J	MC	78
Asparagus albus & Ziziphus lotus	S	✓	F	P	2.25	2.25	1		1	1	129 & 79	101	K	MC	79
Cryptocarya chinensis	T		F	P	0.64	0.64	1		1	1	103 - 213	162	L	MC	80
Aporosa species	T		F	P	50 - 52	51	1		2	2	5 - 6357	307	L	MC	81
Pinus uncinata	T		F	P	0.42	0.42	2	Contrasting ecotones	1	1	17 - 414	54	K	MC	82
Cedrela lilloi (Meliaceae)	T		F	P	1.5	1.5	1		3	3	—	—	L	MC	83
Douglas fir forest	T		F	P	0.89	0.89	1		1	1	—	—	L	MC	84

General			Data collection										Pattern analysis		
Focal species / community	Life form	Arid environment	Data collection technique	Data representation	Plot extents (range) (zdp)	Plot extent (geometric mean) (zdp)	Number of spatial environmental contexts	Environmental conditions compared (i.e., when the number of environmental contexts is > 1)	Replicate plots per environmental context (range)	Replicate plots per environmental context (geometric mean)	Sample size (range)	Sample size (geometric mean)	Function	Interpretative technique	Reference
<i>Quercus ilex</i> and <i>Quercus pubescens</i>	T		F	P	0.01	0.01	7	Pine densities	2,3,3,2 1,2,1	2	—	—	L	MC	85
Old growth conifer forest	T		F	P	1.57-4	2.3	1		6	6	1 - 1834	50	L	MC	86
<i>Acacia raddiana</i>	T	✓	F	P	—	—	1		2	2	c. 200	—	O	MC	87
Xerophytic shrubs	S	✓	F	P	0.03	0.03	1		1	1	6 - 80	18	K	MC	88
<i>Borassus aethiopum</i>	T	✓	F	P	2.25 - 5	3.63	3	Savanna types	2,2,1	2	—	—	L	MC	89
Spruce & fir forest	T		F	P	2	2	1		1	1	—	—	L	MC	90
Desert plant community	S (+H)	✓	F	P	0.03 & 0.15	0.07	2	Short vs. medium strandveld	1	1	— <sup>i</sup>	—	L	MC	91
<i>Carapa-procera</i> & <i>voucacapoua americana</i>	T		F	P	1.44 - 20	6.28	1		1	1	23 - 112	68	L	MC	92
<i>Pinus ponderosa</i>	T		F	P	0.05 - 1.5	0.32	5	Aspect	3,3,3,3 4	3	39 – 71 <sup>i</sup>	49	K	MC	93
<i>Pinus torreyana</i>	T		F	P	1	1	3	Disturbance types	1	1	44 - 197	111	L	MC	94
Tropical forest	T		F	P	0.5	0.5	5	Experimental treatments	1	1	—	—	L	MC	95

General			Data collection										Pattern analysis		Reference
Focal species / community	Life form	Arid environment	Data collection technique	Data representation	Plot extents (range) (zdp)	Plot extent (geometric mean) (zdp)	Number of spatial environmental contexts	Environmental conditions compared (i.e., when the number of environmental contexts is > 1)	Replicate plots per environmental context (range)	Replicate plots per environmental context (geometric mean)	Sample size (range)	Sample size (geometric mean)	Function	Interpretative technique	
Pine & oak forest	T		F	P	0.02 - 0.06	0.05	3	Fire histories	2, 4	3	20 – 989 <sup>g</sup>	181	K	MC	96
<i>Rubus hirtus</i>	S		F	P	< 0.01	< 0.00	1		2	2	—	—	K	MC	97
Tropical moist evergreen forest	T		F	P	0.4	0.4	3	Disturbance and topographical heterogeneity	1	1	71 - 291	173	L & g	MC, H	98
Woody plants	T & S	✓	F	P	10.24	10.24	1		1	1	—	—	L & g	MC	99
<i>Anthyllis cytisoides</i> & <i>Artemisia barrelieri</i>	S	✓	F	P	0.01 – 0.02	0.01	1		3	3	24 - 209	87	K	MC	100
Semi-arid woodland	T	✓	F	P	0.79	0.79	1		1	1	—	—	L	MC	101
Western hemlock & redcedar	T		F	P	0.13 – 0.36	0.25 <sup>d</sup>	1		6	6	—	—	K	MC	102
<i>Retama sphaerocarpa</i> & <i>Artemisia barrelieri</i>	S	✓	F	P	0.2	0.2	1		1	1	17 - 182	47	K	MC	103
Deciduous forest	T		F	P	4	4	1		1	1	197 - 1359	629	L	MC	104
Aspen & white pine forest	T		F	P	0.16	0.16	1		1	1	59 - 371	153	L	MC, H	105
Canopy trees	T		F	P	0.13 - 0.36	0.2	1		6	6	—	—	L	MC	106

General			Data collection										Pattern analysis		Reference
Focal species / community	Life form	Arid environment	Data collection technique	Data representation	Plot extents (range) (zdp)	Plot extent (geometric mean) (zdp)	Number of spatial environmental contexts	Environmental conditions compared (i.e., when the number of environmental contexts is > 1)	Replicate plots per environmental context (range)	Replicate plots per environmental context (geometric mean)	Sample size (range)	Sample size (geometric mean)	Function	Interpretative technique	
<i>Quercus laevis</i>	T		F	P	0.17 - 0.43	0.29	6	Experimental fire regimes	1	1	25 - 470	115	L	MC	107
Tropical trees	T		F	P	0.64	0.64	1		1	1	20 - 43	34	L	MC	108
<i>Eucalyptus obliqua</i>	T		F	P	0.03 - 0.3	0.17	1		12	12	8 - 313	58	L	MC	109

#### Notes

Life form – tree (T), shrub (S); Data collection technique – field collection (F), computational (C); Data representation – point (P), object (O); Function – Ripley's  $K(r)$  (K),  $L(r)$ -function (L), pair-correlation function (g), o-ring statistic (O); Interpretation – comparison to Monte Carlo simulation envelopes (MC), empirical curve height (H), other quantitative technique (Q)

<sup>a</sup> Hemispherical photos, <sup>b</sup> Laser range finder, <sup>c</sup> Total station, <sup>d</sup> only range of plot extents available (plot extent calculated as arithmetic mean), <sup>e</sup> sample size estimated from bar graphs, <sup>f</sup> minimum sample size of 10 individuals, <sup>g</sup> minimum sample size of 20 individuals, <sup>h</sup> minimum sample size of 30 individuals, <sup>i</sup> minimum sample size of 40 individuals, <sup>j</sup> minimum sample size of 70 individuals.

References given below.

- [1] Zhu, Y., Mi, X., Ren, H., & Ma, K. (2010) Density dependence is prevalent in a heterogeneous subtropical forest. *Oikos*, **119**, 109-119.
- [2] Vieira, F.d.A., de Carvalho, D., Higuchi, P., Machado, E.L.M., & dos Santos, R.M. (2010) Spatial pattern and fine-scale genetic structure indicating recent colonization of the palm *Euterpe edulis* in a Brazilian Atlantic forest fragment. *Biochemical Genetics*, **48**, 96-103.
- [3] Barbeito, I., Fortin, M.J., Montes, F., & Cañellas, I. (2009) Response of pine natural regeneration to small-scale spatial variation in a managed Mediterranean mountain forest. *Applied Vegetation Science*, **12**, 488-503.
- [4] Biganzoli, F., Wiegand, T., & Batista, W.B. (2009) Fire-mediated interactions between shrubs in a South American temperate savannah. *Oikos*, **118**, 1383-1395.
- [5] Gao, P.X., Kang, M., Wang, J., Ye, Q.G., & Huang, H.W. (2009) Neither biased sex ratio nor spatial segregation of the sexes in the subtropical dioecious tree *Eurycorymbus cavaleriei* (Sapindaceae). *Journal of Integrative Plant Biology*, **51**, 604-613.
- [6] Gray, L. & He, F.L. (2009) Spatial point-pattern analysis for detecting density-dependent competition in a boreal chronosequence of Alberta. *Forest Ecology and Management*, **259**, 98-106.
- [7] Lan, G.Y., Zhu, H., Cao, M., Hu, Y.H., Wang, H., Deng, X.B., Zhou, S.S., Cui, J.Y., Huang, J.G., He, Y.C., Liu, L.Y., Xu, H.L., & Song, J.P. (2009) Spatial dispersion patterns of trees in a tropical rainforest in Xishuangbanna, southwest China. *Ecological Research*, **24**, 1117-1124.
- [8] Leithead, M.D., Anand, M., & Deeth, L. (2009) A synthetic approach for analyzing tropical tree spatial patterns through time. *Community Ecology*, **10**, 45-52.
- [9] LeMay, V., Pommerening, A., & Marshall, P. (2009) Spatio-temporal structure of multi-storied, multi-aged interior Douglas fir (*Pseudotsuga menziesii* var. *glauca*) stands. *Journal of Ecology*, **97**, 1062-1074.
- [10] Linares-Palomino, R. & Ponce-Alvarez, S.I. (2009) Structural patterns and floristics of a seasonally dry forest in Reserva Ecologica Chaparri, Lambayeque, Peru. *Tropical Ecology*, **50**, 305-314.
- [11] Meador, A.J.S., Moore, M.M., Bakker, J.D., & Parysow, P.F. (2009) 108 years of change in spatial pattern following selective harvest of a *Pinus ponderosa* stand in northern Arizona, USA. *Journal of Vegetation Science*, **20**, 79-90.
- [12] Picard, N., Bar-Hen, A., Mortier, F., & Chadoeuf, J. (2009) Understanding the dynamics of an undisturbed tropical rain forest from the spatial pattern of trees. *Journal of Ecology*, **97**, 97-108.
- [13] Rossi, J.P., Samalens, J.C., Guyon, D., van Halder, I., Jactel, H., Menassieu, P., & Piou, D. (2009) Multiscale spatial variation of the bark beetle *Ips sexdentatus* damage in a pine plantation forest (Landes de Gascogne, Southwestern France). *Forest Ecology and Management*, **257**, 1551-1557.
- [14] Rozas, V., Zas, R., & Solla, A. (2009) Spatial structure of deciduous forest stands with contrasting human influence in northwest Spain. *European Journal of Forest Research*, **128**, 273-285.
- [15] Yu, H., Wiegand, T., Yang, X.H., & Ci, L.J. (2009) The impact of fire and density-dependent mortality on the spatial patterns of a pine forest in the Hulun Buir

- sandland, Inner Mongolia, China. *Forest Ecology and Management*, **257**, 2098-2107.
- [16] Zhang, Y., Zhao, Y., Zhang, C., & Zhao, X. (2009) Structure and spatial distribution of *Pinus tabulaeformis* population in the Songshan Nature Reserve, Beijing, China. *Chinese Journal of Applied and Environmental Biology*, **15**, 175-179.
- [17] Barbeito, I., Pardos, M., Calama, R., & Cañellas, I. (2008) Effect of stand structure on Stone pine (*Pinus pinea* L.) regeneration dynamics. *Forestry*, **81**, 617-629.
- [18] Dagley, C.M. (2008) Spatial pattern of coast redwood in three alluvial flat old-growth forests in northern California. *Forest Science*, **54**, 294-302.
- [19] Djossa, B.A., Fahr, J., Wiegand, T., Ayihouenou, B.E., Kalko, E.K., & Sinsin, B.A. (2008) Land use impact on *Vitellaria paradoxa* C.F. Gaerten. stand structure and distribution patterns: a comparison of Biosphere Reserve of Pendjari in Atacora district in Benin. *Agroforestry Systems*, **72**, 205-220.
- [20] Getzin, S., Wiegand, T., Wiegand, K., & He, F.L. (2008) Heterogeneity influences spatial patterns and demographics in forest stands. *Journal of Ecology*, **96**, 807-820.
- [21] Hofmeister, S., Svoboda, M., Soucek, J., & Vacek, S. (2008) Spatial pattern of Norway spruce and silver fir natural regeneration in uneven-aged mixed forests of northeastern Bohemia. *Journal of Forest Science*, **54**, 92-101.
- [22] Lawes, M.J., Griffiths, M.E., Midgley, J.J., Boudreau, S., Eeley, H.A.C., & Chapman, C.A. (2008) Tree spacing and area of competitive influence do not scale with tree size in an African rain forest. *Journal of Vegetation Science*, **19**, 729-738.
- [23] Li, L., Wei, S.G., Huang, Z.L., Ye, W.H., & Cao, H.L. (2008) Spatial patterns and interspecific associations of three canopy species at different life stages in a subtropical forest, China. *Journal of Integrative Plant Biology*, **50**, 1140-1150.
- [24] Lin, Y.C. & Augspurger, C.K. (2008) Long-term spatial dynamics of *Acer saccharum* during a population explosion in an old-growth remnant forest in Illinois. *Forest Ecology and Management*, **256**, 922-928.
- [25] Longuetaud, F., Seifert, T., Leban, J.M., & Pretzsch, H. (2008) Analysis of long-term dynamics of crowns of sessile oaks at the stand level by means of spatial statistics. *Forest Ecology and Management*, **255**, 2007-2019.
- [26] Metsaranta, J.M. & Leiffers, V.J. (2008) A fifty-year reconstruction of annual changes in the spatial distribution of *Pinus banksiana* stands: does pattern fit competition theory? *Plant Ecology*, **199**, 137-152.
- [27] Meyer, K.M., Ward, D., Wiegand, K., & Moustakas, A. (2008) Multi-proxy evidence for competition between savanna woody species. *Perspectives in Plant Ecology Evolution and Systematics*, **10**, 63-72.
- [28] Montes, F., Rubio, A., Barbeito, I., & Cañellas, I. (2008) Characterization of the spatial structure of the canopy in *Pinus silvestris* L. stands in Central Spain from hemispherical photographs. *Forest Ecology and Management*, **255**, 580-590.
- [29] Moustakas, A., Wiegand, K., Getzin, S., Ward, D., Meyer, K.M., Guenther, M., & Mueller, K.H. (2008) Spacing patterns of an Acacia tree in the Kalahari over a

61-year period: how clumped becomes regular and vice versa. *Acta Oecologica*, **33**, 355-364.

- [30] Suzuki, S.N., Kachi, N., & Suzuki, J.I. (2008) Development of a local size hierarchy causes regular spacing of trees in an even-aged *Abies* forest: Analyses using spatial autocorrelation and the mark correlation function. *Annals of Botany*, **102**, 435-441.
- [31] Xi, W.M., Peet, R.K., & Urban, D.L. (2008) Changes in forest structure, species diversity and spatial pattern following hurricane disturbance in a Piedmont North Carolina forest, USA. *Journal of Plant Ecology*, **1**, 43-57.
- [32] Atkinson, P.M., Foody, G.M., Gething, P.W., Mathur, A., & Kelly, C.K. (2007) Investigating spatial structure in specific tree species in ancient semi-natural woodland using remote sensing and marked point pattern analysis. *Ecography*, **30**, 88-104.
- [33] Castellano, S.M. & Boyce, R.L. (2007) Spatial patterns of *Juniperus virginiana* and *Lonicera maackii* on a road cut in Kentucky, USA. *Journal of the Torrey Botanical Society*, **134**, 188-198.
- [34] Franklin, J. & Rey, S.J. (2007) Spatial patterns of tropical forest trees in Western Polynesia suggest recruitment limitations during secondary. *Journal of Tropical Ecology*, **23**, 1-12.
- [35] Getzin, S. & Wiegand, K. (2007) Asymmetric tree growth at the stand level: random crown patterns and the response to slope. *Forest Ecology and Management*, **242**, 165-174.
- [36] Hao, Z.Q., Zhang, J., Song, B., Ye, J., & Li, B.H. (2007) Vertical structure and spatial associations of dominant tree species in an old-growth temperate forest. *Forest Ecology and Management*, **252**, 1-11.
- [37] Klimas, C.A., Kainer, K.A., & Wadt, L.H.O. (2007) Population structure of *Carapa guianensis* in two forest types in the southwestern Brazilian Amazon. *Forest Ecology and Management*, **250**, 256-265.
- [38] Kubota, Y., Kubo, H., & Shimatani, K. (2007) Spatial pattern dynamics over 10 years in a conifer/broadleaved forest, northern Japan. *Plant Ecology*, **190**, 143-157.
- [39] Malkinson, D. & Kadmon, R. (2007) Vegetation dynamics along a disturbance gradient: spatial and temporal perspectives. *Journal of Arid Environments*, **69**, 127-143.
- [40] Mason, W.L., Connolly, T., Pommerening, A., & Edwards, C. (2007) Spatial structure of semi-natural and plantation stands of Scots pine (*Pinus sylvestris* L.) in northern Scotland. *Forestry*, **80**, 564-583.
- [41] Montesinos, D., Verdú, M., & García-Fayos, P. (2007) Moms are better nurses than dads: gender biased self-facilitation in a dioecious *Juniperus* tree. *Journal of Vegetation Science*, **18**, 271-280.
- [42] Strand, E.K., Robinson, A.P., & Bunting, S.C. (2007) Spatial patterns on the sagebrush steppe/Western juniper ecotone. *Plant Ecology*, **190**, 159-173.
- [43] Wiegand, T., Gunatilleke, S., & Gunatilleke, N. (2007a) Species associations in a heterogeneous Sri lankan dipterocarp forest. *American Naturalist*, **170**, E67-E95 (E-article).



- [44] Wiegand, T., Gunatilleke, S., Gunatilleke, N., & Okuda, T. (2007b) Analyzing the spatial structure of a Sri Lankan tree species with multiple scales of clustering. *Ecology*, **88**, 3088-3102.
- [45] Dolezal, J., Srutek, M., Hara, T., Sumida, A., & Penttila, T. (2006) Neighborhood interactions influencing tree population dynamics in nonpyrogenous boreal forest in northern Finland. *Plant Ecology*, **185**, 135-150.
- [46] Fajardo, A., Goodburn, J.M., & Graham, J. (2006) Spatial patterns of regeneration in managed uneven-aged ponderosa pine Douglas-fir forests of Western Montana, USA. *Forest Ecology and Management*, **223**, 255-266.
- [47] Getzin, S., Dean, C., He, F.L., Trofymow, J.A., Wiegand, K., & Wiegand, T. (2006) Spatial patterns and competition of tree species in a Douglas-fir chronosequence on Vancouver Island. *Ecography*, **29**, 671-682.
- [48] Kubota, Y. (2006) Spatial pattern and regeneration dynamics in a temperate *Abies-Tsuga* forest in southwestern Japan. *Journal of Forest Research*, **11**, 191-201.
- [49] Rozas, V. (2006) Structural heterogeneity and tree spatial patterns in an old-growth deciduous lowland forest in Cantabria, northern Spain. *Plant Ecology*, **185**, 57-72.
- [50] Salas, C., Lemay, V., Nuñez, P., Pacheco, P., & Espinosa, A. (2006) Spatial patterns in an old-growth *Nothofagus obliqua* forest in south-central Chile. *Forest Ecology and Management*, **231**, 38-46.
- [51] Seidler, T.G. & Plotkin, J.B. (2006) Seed dispersal and spatial pattern in tropical trees. *PLoS Biology*, **4**, 2132-2137.
- [52] Boyden, S., Binkley, D., & Shepperd, W. (2005) Spatial and temporal patterns in structure, regeneration, and mortality of an old-growth ponderosa pine forest in the Colorado Front Range. *Forest Ecology and Management*, **219**, 43-55.
- [53] Brumelis, G., Elferts, D., Liepina, L., Luce, I., Tabors, G., & Tjarve, D. (2005) Age and spatial structure of natural *Pinus sylvestris* stands in Latvia. *Scandinavian Journal of Forest Research*, **20**, 471-480.
- [54] Camarero, J.J., Gutierrez, E., Fortin, M.J., & Ribbens, E. (2005) Spatial patterns of tree recruitment in a relict population of *Pinus uncinata*: forest expansion through stratified diffusion. *Journal of Biogeography*, **32**, 1979-1992.
- [55] Druckenbrod, D.L., Shugart, H.H., & Davies, I. (2005) Spatial pattern and process in forest stands within the Virginia piedmont. *Journal of Vegetation Science*, **16**, 37-48.
- [56] Fajardo, A. & Alaback, P. (2005) Effects of natural and human disturbances on the dynamics and spatial structure of *Nothofagus glauca* in south-central Chile. *Journal of Biogeography*, **32**, 1811-1825.
- [57] Fang, W. (2005) Spatial analysis of an invasion front of *Acer platanoides*: dynamic inferences from static data. *Ecography*, **28**, 283-294.
- [58] Koukoulas, S. & Blackburn, G.A. (2005) Spatial relationships between tree species and gap characteristics in broad-leaved deciduous woodland. *Journal of Vegetation Science*, **16**, 587-596.

- [59] Landis, A.G. & Bailey, J.D. (2005) Reconstruction of age structure and spatial arrangement of *piñon-juniper* woodlands and savannas of Anderson Mesa, northern Arizona. *Forest Ecology and Management*, **204**, 221-236.
- [60] Motta, R. & Lingua, E. (2005) Human impact on size, age, and spatial structure in a mixed European larch and Swiss stone pine forest in the Western Italian Alps. *Canadian Journal of Forest Research*, **35**, 1809-1820.
- [61] Tonon, G., Panzacchi, P., Grassi, G., Gianfranco, M., Cantoni, L., & Bagnaresi, U. (2005) Spatial dynamics of late successional species under *Pinus nigra* stands in the northern Apennines (Italy). *Annals of Forest Science*, **62**, 669-679.
- [62] Wolf, A. (2005) Fifty year record of change in tree spatial patterns within a mixed deciduous forest. *Forest Ecology and Management*, **215**, 212-223.
- [63] Dolezal, J., St'astna, P., Hara, T., & Srutek, M. (2004) Neighbourhood interactions and environmental factors influencing old-pasture succession in the Central Pyrenees. *Journal of Vegetation Science*, **15**, 101-108.
- [64] Hou, J.H., Mi, X.C., Liu, C.R., & Ma, K.P. (2004) Spatial patterns and associations in a *Quercus-Betula* forest in northern China. *Journal of Vegetation Science*, **15**, 407-414.
- [65] Ng, K.K.S., Lee, S.L., & Koh, C.L. (2004) Spatial structure and genetic diversity of two tropical tree species with contrasting breeding systems and different ploidy levels. *Molecular Ecology*, **13**, 657-669.
- [66] North, M., Chen, J.Q., Oakley, B., Song, B., Rudnicki, M., Gray, A., & Innes, J. (2004) Forest stand structure and pattern of old-growth western hemlock/Douglas-fir and mixed-conifer forests. *Forest Science*, **50**, 299-311.
- [67] Schurr, F.M., Bossdorf, O., Milton, S.J., & Schumacher, J. (2004) Spatial pattern formation in semi-arid shrubland: a priori predicted versus observed pattern characteristics. *Plant Ecology*, **173**, 271-282.
- [68] Youngblood, A., Max, T., & Coe, K. (2004) Stand structure in eastside old-growth ponderosa pine forests of Oregon and northern California. *Forest Ecology and Management*, **199**, 191-217.
- [69] Zenner, E.K. (2004) Does old-growth condition imply high live-tree structural complexity? *Forest Ecology and Management*, **195**, 243-258.
- [70] Aldrich, P.R., Parker, G.R., Ward, J.S., & Michler, C.H. (2003) Spatial dispersion of trees in an old-growth temperate hardwood forest over 60 years of succession. *Forest Ecology and Management*, **180**, 475-491.
- [71] Arévalo, J.R. & Fernández-Palacios, J.M. (2003) Spatial patterns of trees and juveniles in a laurel forest of Tenerife, Canary Islands. *Plant Ecology*, **165**, 1-10.
- [72] Call, L.J. & Nilsen, E.T. (2003) Analysis of spatial patterns and spatial association between the invasive tree-of-heaven (*Ailanthus altissima*) and the native black locust (*Robinia pseudoacacia*). *American Midland Naturalist*, **150**, 1-14.
- [73] Caylor, K.K., Shugart, H.H., Dowty, P.R., & Smith, T.M. (2003) Tree spacing along the Kalahari transect in southern Africa. *Journal of Arid Environments*, **54**, 281-296.

- [74] Malkinson, D., Kadmon, R., & Cohen, D. (2003) Pattern analysis in successional communities: an approach for studying shifts in ecological interactions. *Journal of Vegetation Science*, **14**, 213-222.
- [75] McDonald, R.I., Peet, R.K., & Urban, D.L. (2003) Spatial pattern of *Quercus* regeneration limitation and *Acer rubrum* invasion in a Piedmont forest. *Journal of Vegetation Science*, **14**, 441-450.
- [76] Moravie, M.A. & Robert, A. (2003) A model to assess relationships between forest dynamics and spatial structure. *Journal of Vegetation Science*, **14**, 823-834.
- [77] Park, A. (2003) Spatial segregation of pines and oaks under different fire regimes in the Sierra Madre Occidental. *Plant Ecology*, **169**, 1-20.
- [78] Schenk, H.J., Holzapfel, C., Hamilton, J.G., & Mahall, B.E. (2003) Spatial ecology of a small desert shrub on adjacent geological substrates. *Journal of Ecology*, **91**, 383-395.
- [79] Tirado, R. & Pugnaire, F.I. (2003) Shrub spatial aggregation and consequences for reproductive success. *Oecologia*, **136**, 296-301.
- [80] Wang, Z.F., Peng, S.L., Liu, S.Z., & Li, Z. (2003) Spatial pattern of *Cryptocarya chinensis* life stages in lower subtropical forest, China. *Botanical Bulletin of Academia Sinica*, **44**, 159-166.
- [81] Debski, I., Burslem, D., Palmiotto, P.A., Lafrankie, J.V., Lee, H.S., & Manokaran, N. (2002) Habitat preferences of *Aporosa* in two Malaysian forests: Implications for abundance and coexistence. *Ecology*, **83**, 2005-2018.
- [82] Camarero, J.J., Gutierrez, E., & Fortin, M.J. (2000) Spatial pattern of subalpine forest-alpine grassland ecotones in the Spanish Central Pyrenees. *Forest Ecology and Management*, **134**, 1-16.
- [83] Grau, H.R. (2000) Regeneration patterns of *Cedrela lilloi* (Meliaceae) in northwestern Argentina subtropical montane forests. *Journal of Tropical Ecology*, **16**, 227-242.
- [84] He, F.L. & Duncan, R.P. (2000) Density-dependent effects on tree survival in an old-growth Douglas fir forest. *Journal of Ecology*, **88**, 676-688.
- [85] Lookingbill, T.R. & Zavala, M.A. (2000) Spatial pattern of *Quercus ilex* and *Quercus pubescens* recruitment in *Pinus halepensis* dominated woodlands. *Journal of Vegetation Science*, **11**, 607-612.
- [86] Van Pelt, R. & Franklin, J.F. (2000) Influence of canopy structure on the understory environment in tall, old-growth, conifer forests. *Canadian Journal of Forest Research*, **30**, 1231-1245.
- [87] Wiegand, T., Milton, S.J., Esler, K.J., & Midgley, G.F. (2000) Live fast, die young: estimating size-age relations and mortality pattern of shrubs species in the semi-arid Karoo, South Africa. *Plant Ecology*, **150**, 115-131.
- [88] Zavala-Hurtado, J.A., Valverde, P.L., Herrera-Fuentes, M.D., & Díaz-Solís, A. (2000) Influence of leaf-cutting ants (*Atta mexicana*) on performance and dispersion patterns of perennial desert shrubs in an inter-tropical region of Central Mexico. *Journal of Arid Environments*, **46**, 93-102.

- [89] Barot, S., Gignoux, J., & Menaut, J.C. (1999) Demography of a savanna palm tree: predictions from comprehensive spatial pattern analyses. *Ecology*, **80**, 1987-2005.
- [90] Chen, J.Q. & Bradshaw, G.A. (1999) Forest structure in space: a case study of an old growth spruce-fir forest in Changbaishan Natural Reserve, PR China. *Forest Ecology and Management*, **120**, 219-233.
- [91] Eccles, N.S., Esler, K.J., & Cowling, R.M. (1999) Spatial pattern analysis in Namaqualand desert plant communities: evidence for general positive interactions. *Plant Ecology*, **142**, 71-85.
- [92] Forget, P.M., Mercier, F., & Collinet, F. (1999) Spatial patterns of two rodent-dispersed rain forest trees *Carapa procera* (Meliaceae) and *Vouacapoua americana* (Caesalpiniaceae) at Paracou, French Guiana. *Journal of Tropical Ecology*, **15**, 301-313.
- [93] Mast, J.N. & Veblen, T.T. (1999) Tree spatial patterns and stand development along the pine-grassland ecotone in the Colorado Front Range. *Canadian Journal of Forest Research*, **29**, 575-584.
- [94] Wells, M.L. & Getis, A. (1999) The spatial characteristics of stand structure in *Pinus torreyana*. *Plant Ecology*, **143**, 153-170.
- [95] Batista, J.L.F. & Maguire, D.A. (1998) Modeling the spatial structure of tropical forests. *Forest Ecology and Management*, **110**, 293-314.
- [96] Fulé, P.Z. & Covington, W.W. (1998) Spatial patterns of Mexican pine-oak forests under different recent fire regimes. *Plant Ecology*, **134**, 197-209.
- [97] Pancer-Koteja, E., Szwagrzyk, J., & Bodziarczyk, J. (1998) Small-scale spatial pattern and size structure of *Rubus hirtus* in a canopy gap. *Journal of Vegetation Science*, **9**, 755-762.
- [98] Péliissier, R. (1998) Tree spatial patterns in three contrasting plots of a southern Indian tropical moist evergreen forest. *Journal of Tropical Ecology*, **14**, 1-16.
- [99] Couteron, P. & Kokou, K. (1997) Woody vegetation spatial patterns in a semi-arid savanna of Burkina Faso, West Africa. *Plant Ecology*, **132**, 221-227.
- [100] Haase, P., Pugnaire, F.I., Clark, S.C., & Incoll, L.D. (1997) Spatial pattern in *Anthyllis cytisoides* shrubland on abandoned land in southeastern Spain. *Journal of Vegetation Science*, **8**, 627-634.
- [101] Martens, S.N., Breshears, D.D., Meyer, C.W., & Barnes, F.J. (1997) Scales of above-ground and below-ground competition in a semi-arid woodland detected from spatial pattern. *Journal of Vegetation Science*, **8**, 655-664.
- [102] Moeur, M. (1997) Spatial models of competition and gap dynamics in old-growth *Tsuga heterophylla* - *Thuja plicata* forests. *Forest Ecology and Management*, **94**, 175-186.
- [103] Haase, P., Pugnaire, F.I., Clark, S.C., & Incoll, L.D. (1996) Spatial patterns in a two-tiered semi-arid shrubland in southeastern Spain. *Journal of Vegetation Science*, **7**, 527-534.
- [104] Ward, J.S., Parker, G.R., & Ferrandino, F.J. (1996) Long-term spatial dynamics in an old-growth deciduous forest. *Forest Ecology and Management*, **83**, 189-202.

- [105] Peterson, C.J. & Squiers, E.R. (1995) An unexpected change in spatial pattern across 10 years in an Aspen White-Pine forest. *Journal of Ecology*, **83**, 847-855.
- [106] Moeur, M. (1993) Characterizing spatial patterns of trees using stem-mapped data. *Forest Science*, **39**, 756-775.
- [107] Rebertus, A.J., Williamson, G.B., & Moser, E.B. (1989) Fire-induced changes in *Quercus-Laevis* spatial pattern in Florida sandhills. *Journal of Ecology*, **77**, 638-650.
- [108] Sterner, R.W., Ribic, C.A., & Schatz, G.E. (1986) Testing for life historical changes in spatial patterns of four tropical tree species. *Journal of Ecology*, **74**, 621-633.
- [109] West, P.W. (1984) Inter-tree competition and small-scale pattern in monoculture of *Eucalyptus obliqua* Lherit. *Australian Journal of Ecology*, **9**, 405-411.

APPENDIX B: SPATIAL VARIATION IN WITHIN-SUBSTRATE *S. SUPRANUBIUS*  
SPECTRAL RESPONSE

The *S. supranubius* spectral response on the red, green and blue wavebands was compared between transects to assess the level of within-substrate spatial variation. All tables show the mean differences of spectral intensity values ranging from 0 to 255 (column minus row). Statistically significant differences (assessed by pairwise Mann Whitney U-tests) at the 0.05 and 0.001 level are shown by '\*' and '\*\*' respectively. Cohen's *d* values are given in parentheses. Cohen's *d* values describing a medium or large effect (i.e., > 0.5) are highlighted in bold. No results are shown for Substrate 5 as data were only collected from one transect.

**Substrate 1**

Transect	1-A	1-B	1-C
1-A	–		
1-B	0.81 (0.06)	–	
1-C	-3.76 (0.28)	-4.58* (0.44)	–

Comparing the transect level spectral response of *S. supranubius* on the red waveband on Substrate 1.

Transect	1-A	1-B	1-C
1-A	–		
1-B	0.24 (0.02)	–	
1-C	-6.82* (0.52)	-7.06* (0.66)	–

Comparing the transect level spectral response of *S. supranubius* on the green waveband on Substrate 1.

Transect	1-A	1-B	1-C
1-A	–		
1-B	0.14 (0.01)	–	
1-C	-5.17 (0.42)	-5.31* (0.51)	–

Comparing the transect level spectral response of *S. supranubius* on the blue waveband on Substrate 1.

**Substrate 2**

Transect	2-A	2-B	2-C	2-D
2-A	–			
2-B	7.11* (0.52)	–		
2-C	-13.40** (1.09)	-20.51** (1.50)	–	
2-D	2.63 (0.21)	-4.48 (0.32)	16.03** (1.26)	–

Comparing the transect level spectral response of *S. supranubius* on the red waveband on Substrate 2.

Transect	2-A	2-B	2-C	2-D
2-A	–			
2-B	1.38 (0.11)	–		
2-C	-7.27* (0.65)	-8.65* (0.67)	–	
2-D	1.81 (0.16)	0.43 (0.03)	9.08** (0.79)	–

Comparing the transect level spectral response of *S. supranubius* on the green waveband on Substrate 2.

Transect	2-A	2-B	2-C	2-D
2-A	–			
2-B	0.04 (0.00)	–		
2-C	-5.44* (0.49)	-5.48* (0.21)	–	
2-D	1.43 (0.13)	1.40 (0.05)	6.88* (0.69)	–

Comparing the transect level spectral response of *S. supranubius* on the blue waveband on Substrate 2.

**Substrate 3**

<b>Transect</b>	<b>3-A</b>	<b>3-B</b>	<b>3-C</b>
<b>3-A</b>	–		
<b>3-B</b>	-2.27 (0.18)	–	
<b>3-C</b>	-11.96** (0.98)	-9.69** (0.68)	–

Comparing the transect level spectral response of *S. supranubius* on the red waveband on Substrate 3.

<b>Transect</b>	<b>3-A</b>	<b>3-B</b>	<b>3-C</b>
<b>3-A</b>	–		
<b>3-B</b>	-1.43 (0.12)	–	
<b>3-C</b>	-13.11** (1.09)	-11.69** (0.87)	–

Comparing the transect level spectral response of *S. supranubius* on the green waveband on Substrate 3.

<b>Transect</b>	<b>3-A</b>	<b>3-B</b>	<b>3-C</b>
<b>3-A</b>	–		
<b>3-B</b>	1.53 (0.14)	–	
<b>3-C</b>	-7.84* (0.63)	-9.37** (0.73)	–

Comparing the transect level spectral response of *S. supranubius* on the blue waveband on Substrate 3.



**Substrate 4**

Transect	4-A	4-B
4-A	–	
4-B	1.62 (0.10)	–

Comparing the transect level spectral response of *S. supranubius* on the red waveband on Substrate 4.

Transect	4-A	4-B
4-A	–	
4-B	1.60 (0.12)	–

Comparing the transect level spectral response of *S. supranubius* on the green waveband on Substrate 4.

Transect	4-A	4-B
4-A	–	
4-B	1.87 (0.14)	–

Comparing the transect level spectral response of *S. supranubius* on the blue waveband on Substrate 4.

# APPENDIX C: EXAMPLE OF CLASSIFIER ACCURACY OUTPUT

Classification accuracies produced by the polynomial kernel using the INCSVDD classifier model. Classifiers trained on target training data only (dataset A). Accuracies are calculated from a total of 250. 'p' provides the order of the polynomial kernel, and 'C' details the fraction of target cases that can be rejected by the data description (the rejection error). The classifier highlighted in bold is the final model used in the image classification.

p	C	Target accuracy (#)	Outlier accuracy (#)	Overall accuracy (%)
1	0.1	230	240	94.0
	0.01	250	228	95.6
	0.001	250	224	94.8
	0.0001	250	224	94.8
2	0.1	235	240	95.0
	0.01	250	228	95.6
	<b>0.001</b>	<b>250</b>	<b>225</b>	<b>95.0</b>
	0.0001	250	225	95.0
3	0.1	245	237	96.4
	0.01	250	227	95.4
	0.001	250	225	95.0
	0.0001	250	225	95.0
4	0.1	245	235	96.0
	0.01	250	226	95.2
	0.001	250	225	95.0
	0.0001	250	225	95.0
5	0.1	245	235	96.0
	0.01	250	228	95.6
	0.001	250	225	95.0
	0.0001	250	225	95.0
6	0.1	245	235	96.0
	0.01	250	227	95.4
	0.001	250	225	95.0
	0.0001	250	225	95.0
7	0.1	245	231	95.2
	0.01	250	225	95.0
	0.001	250	225	95.0
	0.0001	250	225	95.0
8	0.1	246	229	95.0
	0.01	250	226	95.2
	0.001	250	225	95.0
	0.0001	250	225	95.0
9	0.1	245	232	95.4
	0.01	250	225	95.0
	0.001	250	225	95.0
	0.0001	250	225	95.0
10	0.1	246	229	95.0
	0.01	250	224	94.8
	0.001	250	225	95.0
	0.0001	250	225	95.0

APPENDIX D: INFORMATION ON QUADRATS USED IN CHAPTER 5

Extent (ha)	Quadrat number	Substrate 2			Substrate 4		
		Number of individuals	ULX	ULY	Number of individuals	ULX	ULY
0.0625	1	6	340990	3124460	10	342994	3124162
	2	8	340421	3124566	5	343185	3124102
	3	4	341053	3124892	5	343196	3124176
	4	6	340528	3124073	5	342640	3124067
	5	2	340456	3124805	6	343248	3123911
	6	6	341006	3124534	4	342892	3123889
	7	14	340501	3124930	10	342631	3123937
	8	14	340806	3124491	4	342878	3123825
	9	4	340860	3124825	7	342647	3123767
	10	5	340510	3124752	2	343021	3123652
0.25	1	5	340687	3124946	33	342674	3123682
	2	30	340555	3124936	27	342863	3123779
	3	17	340446	3124932	27	343210	3123851
	4	35	340431	3124641	42	343102	3123972
	5	22	340589	3124982	42	342980	3124063
	6	22	340649	3124485	28	342606	3123803
	7	47	340791	3124488	30	342730	3123907
	8	43	340708	3124408	30	342779	3123642
	9	11	340763	3124953	22	342788	3124249
	10	24	340451	3125012	31	342718	3124085
1	1	31	340652	3124995	169	343044	3123654
	2	65	340777	3124622	155	342720	3123731
	3	119	340764	3124689	140	342673	3123947
	4	130	340910	3124426	130	343129	3123991
	5	115	340435	3124436	87	342908	3123718
	6	95	340649	3124613	114	342851	3123662
	7	75	340984	3124979	131	342851	3123662
	8	79	340944	3124553	119	342858	3123854
	9	146	340442	3124523	143	343079	3123746
	10	124	340648	3124733	97	342918	3123935
2.25	1	260	340806	3124637	285	342757	3123874
	2	246	340455	3124835	305	342614	3123647
	3	186	340476	3124656	274	342925	3123947
	4	290	340543	3124480	288	342817	3123604
	5	235	340612	3124688	247	342639	3123840
	6	61	340694	3124941	277	342616	3124089
	7	252	340824	3124468	278	343103	3123853
	8	137	340672	3124833	337	343025	3123684
	9	272	340735	3124750	436	342998	3124131
	10	250	340780	3124746	280	342828	3124047
4	1	395	340431	3124596	399	342817	3123666
	2	410	340879	3124406	651	343011	3123061
	3	181	340701	3124891	464	342911	3123926
	4	393	340645	3124568	500	342668	3123880

	5	408	340752	3124555	531	342992	3123725
	6	409	340859	3124761	482	342765	3124034
	7	320	340515	3124889	528	343038	3123852
	8	349	340882	3124892	493	342600	3123672
	9	524	340692	3124439	409	342818	3123826
	10	383	340412	3124786	551	342967	3123608
<b>6.25</b>	1	388	340619	3124815	791	342993	3123744
	2	622	340552	3124563	744	342870	3123913
	3	634	340627	3124678	731	342759	3123606
	4	665	340786	3124611	814	343045	3123939
	5	644	340751	3124722	763	342641	3123803
	6	799	340421	3124431	750	342644	3124045
	7	594	340836	3124475	726	342605	3123695
	8	648	340657	3124481	858	342985	3123622
	9	587	340411	3124702	956	342942	3124037
	10	613	340420	3124579	891	342843	3124036

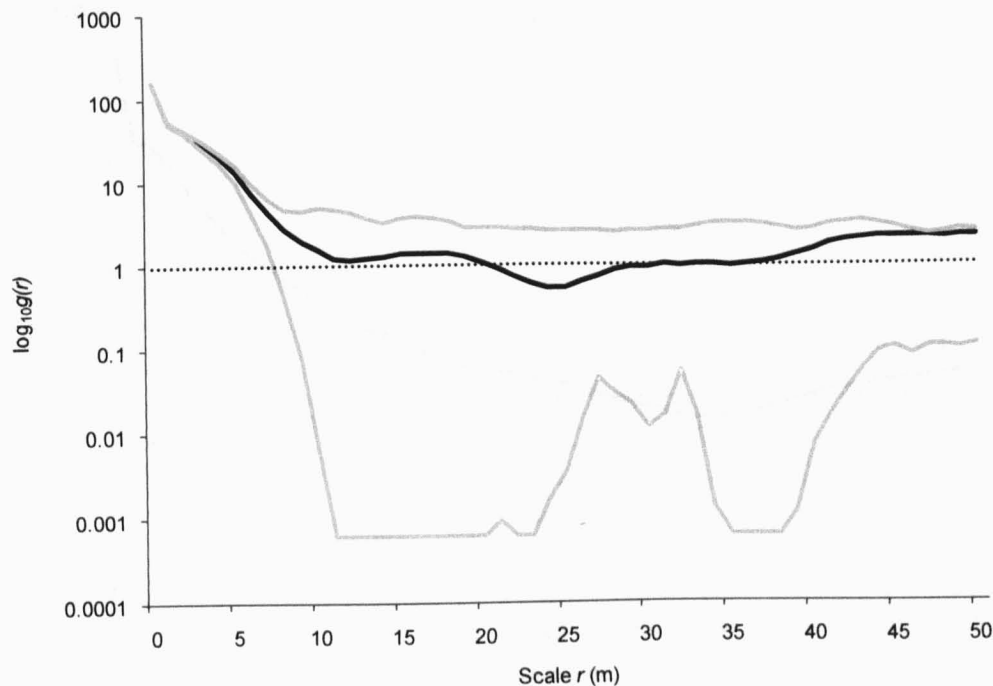
Notes:

ULX = X-coordinate (UTM) of the lower left corner of the quadrat.

ULY = Y-coordinate (UTM) of the lower left corner of the quadrat

All quadrats are square.

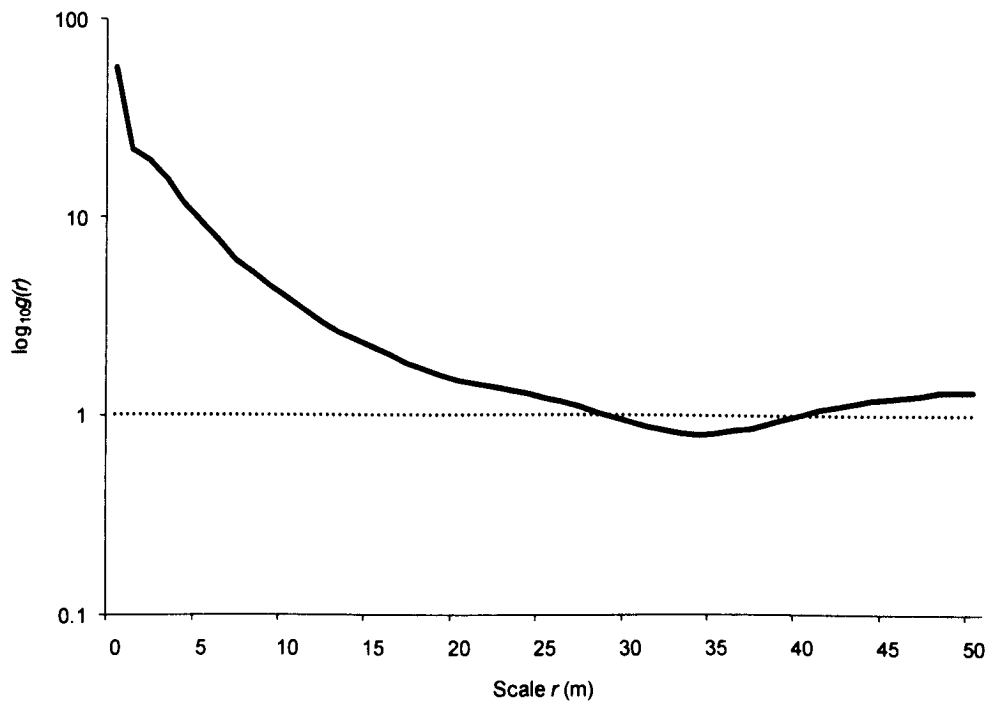
APPENDIX E: ASSESSING THE HETEROGENEITY OF SUBSTRATE 5 USING  
REAL-SHAPE ANALYSIS (CHAPTER 7)



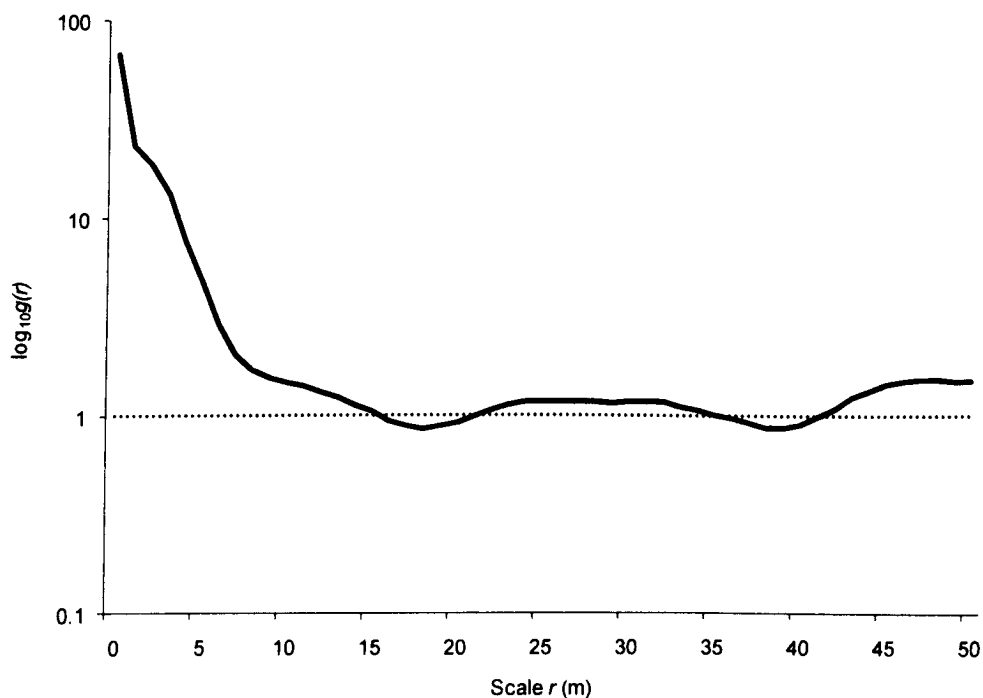
The pattern of the largest individuals ( $\geq 30 \text{ m}^2$ ) on Substrate 5 was compared to the null model of CSR using the homogeneous  $g(r)$  to assess the presence of environmental heterogeneity. In contrast to the original analyses (Section 7.4.1), the real size and shape of the shrubs was preserved during analysis to ameliorate the effect of the small sample size ( $n = 15$ ). While the empirical function (black line) falls below the CSR expectation of  $g(r) = 1$  between 31 and 30 m, the empirical function remains within the 99% CSR simulation envelope (constructed from the highest and lowest value of 99 simulations of the null model).

Note: the lower simulation envelope returned values of 0 at several scales. To enable plotting on log-axes these values were replaced with the second lowest value for the lower simulation envelope (0.00062).

APPENDIX F: ANALYSING THE SPATIAL PATTERN OF LARGE INDIVIDUALS ON  
SUBSTRATES 3 AND 5 USING REAL-SHAPE ANALYSIS (CHAPTER 7)



The pattern of the large individuals on Substrate 3 was compared to the null model of CSR using the homogeneous  $g(r)$ . In contrast to the original analyses (Section 7.4.3) the size and shape of the shrubs was preserved during analysis to ameliorate the effect of the small sample size ( $n = 83$ ). The empirical function (black line) falls below the CSR expectation of  $g(r) = 1$  between 29 and 39 m.



The pattern of the large individuals on Substrate 5 was compared to the null model of CSR using the homogeneous  $g(r)$ . In contrast to the original analyses (Section 7.4.3) the size and shape of the shrubs was preserved during analysis to ameliorate the effect of the small sample size ( $n = 43$ ). The empirical function (black line) falls below the CSR expectation of  $g(r) = 1$  between 16 and 21 m, and between 36 and 41 m.