

Bangor University

DOCTOR OF PHILOSOPHY

Perspectives in vegetation monitoring: an evaluation of approaches currently used in the UK

Hearn, Suzanne M.

Award date: 2009

Awarding institution: Bangor University

Link to publication

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- · Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
 You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Perspectives in vegetation monitoring: an evaluation of approaches currently used in the UK

A thesis submitted for the degree of Doctor of Philosophy to Bangor University

by

Suzanne M. Hearn BA (Hons), M.Sc.

School of the Environment and Natural Resources, Bangor University Bangor

August 2009



Abstract

Environmental managers are interested in detecting change in order to manage natural resources. In an applied context, managers want to know i) how much of a resource there is and ii) the quality or value of the resource. I assayed UK conservation practitioners' views of the effectiveness of current vegetation monitoring methods. Concern was expressed that vegetation mapping and condition assessment (used to assess quantity and quality of a resource respectively) techniques involve too much subjectivity with consequent uncertainty and devaluation of information on which to base management decisions.

These perceptions were explored by quantifying the amount of between-observer variation in maps prepared using the National Vegetation Classification (NVC) and in assessments produced using approaches to Common Standards Monitoring (CSM), both commonly used systems in the UK. NVC mapping produced unacceptably high levels of between-observer variation, with average spatial agreement between seven surveyors of only 34.2%. CSM approaches were shown to produce inconsistent results, with major implications for reporting against national condition targets.

Quantitative methods utilising systematically located plots were also investigated and demonstrated that variation is unpredictable through spatial scales and that sampling intensity and choice of metric both influence the effort required to detect a specified change. Furthermore, quantitative and qualitative approaches require unfeasible levels of effort to detect meaningful sizes of change with sufficient power.

This leads to the following recommendations:

 a) Vegetation mapping is unsuitable for monitoring change and should only be used to assess state with full acknowledgement of its subjectivity and uncertainty. Condition assessment could provide widespread snapshots of condition, but should not be used for monitoring temporal change without some quantifiable controls.

b) Quantitative methods must include i) *a priori* definition of the minimum (or threshold) change that should be detected, ii) relative importance of type I and II error rates and iii) definition of sample population and sampling frame which specifically includes consideration of spatial heterogeneity.

Finally, this study suggests that research should focus on the development of a multiscale and nested quantitative sampling design which could address monitoring and

ii

surveillance questions in the UK. Since resources do not permit universal application of this approach, quantitative data from a small number of sites could be used via diagnostic test methodology (as used in medicine to assess the accuracy of screening tests) to validate widespread condition assessment.

Acknowledgements

I would like to thank my supervisors in the School of the Environment and Natural Resources (SENR) at Bangor University, Dr Morag McDonald and Dr John Healey for their support, advice and pragmatism throughout the duration of this project. Thanks are also due to Dr Jenny Wong of Wild Resources Ltd who proposed the original idea, provided a shared office space and offered plenty of enthusiasm. Dr Gavin Stewart arrived at Bangor University with the Centre for Evidence Based Conservation about one year into my PhD and was invaluable in providing an understanding of issues of quantitative sampling and data analysis. My external examiner Dr Robin Pakeman also offered useful suggestions and insights.

During the course of the PhD, I have encountered a great number of people who have helped in many ways. Thanks to everyone who took part in the interviews across the UK, especially Keith Kirby from Natural England, Simon Smart at the Centre for Ecology and Hydrology and James Williams from the Joint Nature Conservation Committee. Thanks also to all of the surveyors who took part in the field trials of the various approaches; experts, consultants and volunteers. The Countryside Council for Wales (CCW) has a great deal of expertise in vegetation monitoring and I am particularly grateful to David Allen, Alan Brown, Julie Creer, Dan Guest, Clive Hurford, Dylan Lloyd and Alex Turner for their time and support. Dr Terry Rowell, retired head of monitoring at CCW was enormously helpful in providing a background to monitoring in the UK, explaining the development of Common Standards Monitoring and commenting on drafts. Helen Buckingham and Sabine Nouvet from the National Trust discussed general issues and provided specifics of current practice at Hafod y Llan in Snowdonia.

Within SENR, I would particularly like to thank Ian Harris for his tireless help with GIS and James Gibbons for statistical advice. I would also like to thank my fellow PhD students in SENR for their support especially Paul Cross, Steve Murphy, Chris Stiefvater-Thomas and James Walmsley. Also in Bangor, Andrew Packwood at CASZ (Centre for Arid Zone Studies) provided help on spatial analysis of data and Zoe Frogbrook at CEH demystified geostatistics and helped with kriging of the data. Special thanks go to Alejandra Hernandez Guzman for help with field sampling and for her company in the hills of North Wales.

Finally thanks go to my family, Miles and Barry for their endless patience and unquestioning acceptance.

Contents	5
----------	---

Content		Page
Declaration. Abstract Acknowledg Table of cor List of tables List of figure	ements itents s	i iv v viii xii
List of appe		XVII
Chapter 1	General introduction	1
1.1 1.1.1 1.1.2 1.1.3 1.2 1.2.1 1.2.2 1.3 1.4	Background to environmental monitoring Monitoring worldwide Monitoring in Europe Monitoring in the UK Vegetation monitoring in the UK Species and communities Monitoring plant communities: problems and methods Definition of monitoring and surveillance Rationale for the project	2 3 3 6 9 10 14 16
1.5	Specific context: the problem of monitoring vegetation	
1.6 1.7 Chapter 2	change across large, complex sites Thesis objectives Thesis structure Vegetation monitoring methods in the UK: the	17 21 21
	practitioners' perspective	24
2.0 2.1 2.1.2 2.2 2.2.1 2.2.2 2.2.3 2.3	Abstract Introduction Objectives Methods People interviewed Interview questions and procedure Analysis Results	25 26 27 29 29 30 32 33
2.3.1	Vegetation monitoring methods in use in the UK and how	-
2.3.2	Perceptions of the effectiveness of Common Standards Monitoring	33 37
2.3.3 2.3.4 2.4 2.4.1 2.5	Criteria for an effective monitoring system Stakeholder category Discussion Limitations to this study Conclusions.	39 39 41 44 45

Chapter 3	The repeatability of the National Vegetation 4 Classification approach to vegetation mapping	6
3.0	Abstract 4	7
3.1	Introduction 4	8
3.1.1	Vegetation mapping and classification 4	8
3.1.2	This study and the National Vegetation Classification 5	0
3.1.3	Issues to be addressed in this study 5	2
3.1.4	Objectives	6
3.2	Methods 5	7
3.2.1	Selection of surveyors for the study	7
3.2.2	Organisation of the field trial 5	7
3.2.3	Analysis of variation amongst maps	0
3.3	Results 6	4
3.3.1	Overall differences amongst 2008 maps 6	4
3.3.2	Detail of variation amongst 2008 maps 6	6
3.3.3	Variation between 2008 maps and map produced in 1999 8	0
3.4	Discussion 8	4
3.4.1	Limitations to this study 8	9
3.5	Conclusions	0
Chapter 4	Consistency in habitat condition assessments	1
4.0	Abstract	12
4.1	Introduction	13
4.1.1	Environmental management and condition assessment 9	13
4.1.2	Common Standards Monitoring 9	14
4.1.3	Condition assessment methods used in this study	8
4.1.4	Accuracy and consistency 10	11
4.1.5	Objectives	15
4.2	Methods	16
4.2.1	Field trials	16
4.3	Results	1
4.3.1	Comparison of condition assessment methods	1
4.3.2	Discussion	9
4.4	Discussion	:0
4.4.1	consistency of different approaches to condition)E
442	Effect of experience 12	20
4.4.2	Limitations to this study 12	21
4.4.5	Conclusions to this study	22
4.5		2
Chapter 5	Quantitative vegetation monitoring; sampling for the detection of temporal change 13	33
5.0	Abstract	34
5.1	Introduction 13	35
5.1.1	The need for vegetation monitoring methods which work 13	35
5.1.2	Detection of change: theory 13	38
5.1.3	Importance of spatial pattern 14	10
5.1.4	Spatial pattern analysis 14	11
5.1.5	Spatial pattern, sampling design and monitoring 14	15

5.1.6	Objectives	147
5.2	Methods	148
5.2.1	Data collection and collation	148
5.2.2	Data analysis	154
5.3	Results	161
5.3.1	Spatial pattern of vegetation and relationship to	
	environmental variables	161
5.3.2	Detection of change	169
5.4	Discussion	175
5.4.1	Limitations to this study and recommendations for the	
	future	181
5.5	Conclusions	182
Chapter 6	General discussion: implications of this study for	
enapter e	ocheral discussion. Implications of this study for	
onaptor o	monitoring change across complex sites and	
	monitoring change across complex sites and landscapes	183
6.1	monitoring change across complex sites and landscapes	183
6.1	wonitoring change across complex sites and landscapes	183
6.1 6.2	wonitoring change across complex sites and landscapes	183 184
6.1 6.2	with the state of the stat	183 184 185
6.1 6.2 6.3	wonitoring change across complex sites and landscapes Why is so much vegetation monitoring in the UK a 'waste of time'? Implications of using qualitative methods for vegetation monitoring Implications of using qualitative methods for vegetation	183 184 185
6.1 6.2 6.3	wonitoring change across complex sites and landscapes Why is so much vegetation monitoring in the UK a 'waste of time'?	183 184 185 190
6.1 6.2 6.3 6.4	wonitoring change across complex sites and landscapes	183 184 185 190
6.1 6.2 6.3 6.4	Why is so much vegetation monitoring in the UK a 'wasteof time'?Implications of using qualitative methods for vegetationmonitoringImplications of using quantitative methods for vegetationmonitoringConclusions and recommendations for practice and futureresearch.	183 184 185 190 194
6.1 6.2 6.3 6.4	wonitoring change across complex sites and landscapes	183 184 185 190 194
 6.1 6.2 6.3 6.4 References 	wonitoring change across complex sites and landscapes	183 184 185 190 194 197
 6.1 6.2 6.3 6.4 References 	wonitoring change across complex sites and landscapes	183 184 185 190 194 197

List of tables

Table

-	S		
	\sim	2	0
-	0	u	e

		-
	Chapter 1	
1.1	Typology of vegetation monitoring methods, including information about scale, measurement of species and community change, main purposes, strengths and weaknesses, major sources of error, cost and examples currently in use in the UK	13
1.2	Objectives, opinions and obstacles for vegetation monitoring on Hafod y Llan, Snowdonia National Park in 2006. Information obtained from National Trust reports and various meetings and conversations with H. Buckingham and K. Jones from the National Trust in 2006 (NT 2000; 2004)	17
	Chapter 2	
2.1	Location of sixty conservation practitioners interviewed about vegetation monitoring in 2006-8	29
2.2	Stakeholder status of sixty conservation practitioners interviewed about vegetation monitoring in 2006-8	29
2.3	Organisations employing sixty conservation practitioners interviewed about vegetation monitoring in 2006-8	30
2.4	Description of methods listed by conservation practitioners interviewed in 2006-8 about vegetation monitoring in the UK. Four of the monitoring methods are defined by specific publications: Countryside Survey (Haines-Young <i>et al.</i> , 2002); Common Standards Monitoring (JNCC, 2004c); Environmental Change Network (Sykes and Lane, 1996); Phase 1 (JNCC, 2004d) and National Vegetation Classification (Rodwell, 2006)	33
2.5	Strengths and weaknesses of vegetation monitoring methods, according to sixty conservation practitioners interviewed in 2006-8 about vegetation monitoring in the UK	36
2.6	Criteria for effective vegetation monitoring methods, according to sixty conservation practitioners interviewed in 2006-8	39
	Chapter 3	
3.1	Status, experience with the NVC and upland habitat surveying, and cost charged per survey for seven surveyors involved in a field trial carried out in Snowdonia in 2008	57
3.2	Date of survey, weather and visibility conditions during survey, time taken, length of route and way in which patches were assigned to NVC type for seven surveyors involved in a field trial of NVC mapping in Snowdonia in 2008. BPC= British Plant	

	Communities books	59
3.3	Example of a matrix of correspondence in area (ha) between two maps, A and B, with five NVC community types i, ii, iii, iv and v	61
3.4	Information about the media that seven surveyors drew their maps onto (AP= aerial photograph, OS= Ordnance Survey map), number and size of polygons and proportion mapped to sub-community in their maps. Data were collected in a field trial carried out in Snowdonia, 2008	64
3.5	The areas of National Vegetation Classification (a) habitats (and non-NVC categories), (b) communities and (c) sub- communities in each surveyor's map, as a percentage of total study area. Number of vegetation types recorded in each map are also shown. Data were collected in a field trial carried out in Snowdonia, 2008. * indicates area < 0.05%	67
3.6	Spatial concurrence between surveys assessed using percentage agreement between each pair of maps. Concurrence was assessed using original and buffered maps. Sub-community figures only include the four maps with >90% of the study area mapped to sub-community. Data were collected in a field trial carried out in Snowdonia in 2008	69
3.7	The percentage of study area agreed on in (a) whole study area and (b) Part A and Part B of study area in none of the maps to all seven maps in terms of National Vegetation Classification habitats, communities and sub-communities (note that sub-community figures only include the four maps with >90% of the study area mapped to sub-community). Data collected in a field trial carried out in Snowdonia in 2008	73
3.8	Matrices of confusion and similarity between pairs of National Vegetation Classification communities. (a) Amount of confusion amongst surveyors based on the mean proportion of total study area in agreement in 21 pair-wise combinations of seven NVC maps produced during a field trial in Snowdonia in 2008 (0 = no confusion, 100 = total confusion). Bold indicates pairs with levels of confusion \ge 30%. Figures in italics in the central diagonal indicate agreement between NVC types, with 0 = no agreement and 100 = complete agreement). (b) Amount of similarity based on co-efficients derived from MATCH (Malloch 1990) according to species composition and frequency (Rodwell, 1991b, 1992), (0 = no similarity, 100 = total similarity). Bold indicates pairs with levels of similarity \ge 60%. (c) Proportion of constant species in common based on floristic tables in MATCH (Malloch 1990) (0 = no constant species in common, 100 = all constant species in common). Bold indicates pairs with \ge 60% constant species in common. M16 was only present in one map	76
	Mitb was only present in one map	76

Spearman's rank correlation co-efficients between mean area 3.9

of agreement per surveyor and number of types listed per survey for NVC major, community and sub-community types and number of years of surveyor NVC mapping experience, number of years surveying in upland habitats and cost charged per survey. Data were collected in field trials carried out in 79 Snowdonia. 2008..... 3.10 The areas of National Vegetation Classification (a) habitats (and non-NVC categories), (b) communities and (c) subcommunities in the map produced in 1999 as part of statutory habitat mapping of the Eryri SSSI, as a percentage of area mapped. Number of vegetation types recorded are also shown. 81 3.11 Spatial concurrence between the 1999 map and each of the 2008 maps, assessed using percentage agreement. Subcommunity figures only include the 1999 map and the four 2008 maps with >90% of the study area mapped to subcommunity. Data were collected in Snowdonia during statutory mapping in 1999 and in field trials in 2008..... 82 3.12 Spearman's rank correlation co-efficients between proportions of different NVC, communities in pairs of maps. Data were collected during statutory habitat mapping in 1999 and in field trials carried out in Snowdonia. 2008. r = correlation coefficient..... 83 **Chapter 4** 4.1 Sampling strategy and target thresholds used in field trials of three condition-monitoring methods (Grid System, standard CSM and a rapid assessment) at Aberffraw to Abermenai Dunes SAC in 2008. Further information is provided in appendix 4.1.... 109 4.2 Sensitivity and specificity of site condition monitoring methods (standard CSM and rapid assessment) based on betweenobserver agreement with another method, the Grid System in field trials with 19 observers at Aberffraw to Abermenai Dunes SAC in 2008. 112 4.3 Index of conspicuousness (specific to the site and time of year) for species recorded in field trials at Aberffraw to Abermenai Dunes SAC in 2008. Information collected from the site the Ecological Flora Database and the Wild Flower Key (Rose, 2006; Fitter and Peat, 2008)..... 114 4.4 Difference in frequency of seven positive indicator species (actual and mean) estimated using a rapid assessment method and recorded using standard CSM guidance to assess condition of fixed dune grassland during field trials at Aberffraw to Abermenai Dunes SAC in 2008..... 115 4.5 Difference in mean frequency of positive indicator species recorded by volunteer and professional observers using the

Х

	Grid system to assess condition of fixed dune grassland during field trials at Aberffraw to Abermenai Dunes SAC in 2008	120
4.6	Meta-analysis of measures of sensitivity and specificity within groups of volunteer and professional observers using the Grid System approach to condition assessment during field trials at Aberffraw to Abermenai Dunes SAC in 2008	122
4.7	Type I error rate between and within assessments by pairs of volunteer and professional observers using the Grid System approach to condition assessment during field trials at Aberffraw to Abermenai Dunes SAC in 2008	124
4.7	Condition assessment data for selected habitats and all terrestrial habitats from England, Northern Ireland, Scotland and Wales from the first cycle of habitat condition assessments of designated conservation sites: Site of Special Scientific Interest (SSSI), Special Area of Conservation (SAC), Special Protection Area (SPA) and Area of Special Scientific Interest (ASSI) (TEN 2009; NILEA 2009; SNIL 2009; MAC 2009)	100
	(ASSI) (EN, 2005, NIEA, 2006, SNH, 2006, WAG, 2006)	120
	Chapter 5	
5.1	Classification of habitats across Hafod y Llan, Snowdonia National Park in terms of National Vegetation Classification (NVC) and prioritisation at UK and European levels	148
5.2	Taxa (species or groups of species) measured at sample plots across a grid at Hafod y Llan, Snowdonia National Park in 2006/7. Taxa in bold are those used in analysis	151
5.3	Environmental variables collected at sample points and in plots in Hafod y Llan, Snowdonia National Park in 2006/7, for details of measurements see Appendix 5.1	152
5.4	Number of sample points recorded per day at various sampling intensities; based on field methodology used in Hafod y Llan, Snowdonia National Park in 2006/7	159
5.5	Information about taxa: percentage of sample points found in, skewness, lag distance used to compute variogram and model parameters. Data collected from Hafod y Llan, Snowdonia National Park in 2006/7	161
5.6	Percentage of variation explained and intraset correlations (correlation between the ordination axes and environmental variables) from the Canonical Correspondence Analysis of taxa height data and environmental variables collected at	
	Hafod y Llan, Snowdonia National Park in 2006-7	166

List of figures

Figure		Page
	Chapter 1	
1.1	Alternative definitions of monitoring and surveillance in use in conservation literature	15
1.2	Photo monitoring of a fixed point quadrat at Hafod y Llan, Snowdonia National Park in 1999 and 2006 (no record of photo taken in 2003)	19
1.3	Mean cover of (a) dwarf shrub species and (b) competitive species recorded in seven Fixed Point Quadrats in Hafod y Llan, Snowdonia National Park in 1999, 2003 and 2006. Bars show +/- 1 SE	20
1.4	The structure and layout of the thesis	23
	Chapter 2	
2.1	Proportion of sixty conservation practitioners interviewed about vegetation monitoring in 2006-8 who use different vegetation monitoring methods	34
2.2	Mean rating for vegetation monitoring methods according to sixty conservation practitioners interviewed about vegetation monitoring in 2006-8, $0 =$ not effective, $10 =$ very effective. Some interviewees listed the same method category more than once, including different versions of a single method; all of their scores are used to calculate the mean score. Bars show +/- 1 SE. Figures in brackets show number of participants' scores in each group	35
2.3	Mean rating of versions of Common Standards Monitoring (CSM) according to 34 conservation practitioners, $0 = not$ effective, $10 = very$ effective. Figures in brackets show number of people involved; some interviewees listed more than one version of CSM. Bars show +/- 1 SE.	37
2.4	Mean rating for habitat sets of Common Standards Monitoring guidance, $0 =$ not effective, $10 =$ very effective. Figures in brackets show number of participants' scores in each group, bars show +/- 1 SE.	38
2.5	Mean rating for all vegetation monitoring methods according to groups of stakeholders interviewed, $0 = \text{not effective}$, $10 = \text{very effective}$. Bars show +/- 1 SE	40

Chapter 3

3.1	Study area used for field trial of NVC mapping. The area is part of The National Trust's Hafod y Llan estate in Snowdonia National Park in North Wales (SH6252). Crown copyright/database right 2008. An Ordnance Survey/EDINA supplied service.	58
3.2	Maps A-G of National Vegetation Classification community. Data collected in a field trial carried out in Snowdonia, 2008. See appendix 3.2 for larger scale versions	65
3.3	The number of surveyors' maps which agree for each 10 m x 10 m grid square in terms of classification to (a) habitat, (b) community and (c) sub-community National Vegetation Classification type. Data were collected in a field trial carried out in Snowdonia, 2008.	70
3.4	The proportion of area of National Vegetation Classification (a) habitats (M= mires and wet heaths, U= calcifugous grasslands and montane communities, H= dry heaths and CG= calcicolous grasslands), (b) communities and (c) sub-communities agreed on in 7 maps (all seven maps in agreement) to 1 map (no maps in agreement). Bars are ranked according to NVC types with highest amount of agreement at the left and least at the right. Data were collected in a field trial carried out in Snowdonia in 2008.	75
3.5	Relationship between proportion of confusion between pairs of NVC communities in the maps of seven surveyors and (a) similarity coefficients derived from match ($y = 0.32x + 41.96$, $r^2 = 0.108$), (b) proportion of constants in common based on floristic tables in MATCH ($y = 0.35x + 1.56$, $r^2 = 0.33$). Data were collected in field trials carried out in Snowdonia, 2008	78
3.6	Map of National Vegetation Classification community produced during statutory mapping in Snowdonia in 1999	80
	Chapter 4	
4.1	Potential sources of inconsistency in the application of Common Standards Monitoring guidance	96
4.2	Schematic illustration of terms precision and accuracy taken from Wong et al. (2005)	102
4.3	Relationships among terms in sensitivity/specificity matrix, taken from Hopley <i>et al.</i> (2001)	103
4.4	Location of study site in Wales used for field trials of condition assessments in 2008.	106
4.5	Location of 50 x 50 m area of fixed dune grassland in Aberffraw to Abermenai Dunes SAC and detail of design of 'W' walk used in standard CSM approach and grid used in the Grid System in	

Number of observers (out of 19) who assessed fixed dune 4.6 grassland to be in favourable condition using three condition monitoring methods; Grid System, Standard CSM and rapid assessment. Data collected in a field trial at Aberffraw to Abermenai Dunes SAC in 2008..... 111 4.7 Mean frequency of positive indicator species estimated using a rapid assessment against that recorded using standard CSM guidance to assess condition of fixed dune grassland during field trials at Aberffraw to Abermenai Dunes SAC in 2008..... 113 4.8 Photos of species recorded in field trials of methods to assess condition of fixed dune grassland during field trials at Aberffraw to Abermenai Dunes SAC in 2008. Photos taken by S. Hearn and C. Farmer (Farmer, 2008)..... 114 4.9 Number of observers who assess the area as being in favourable condition under a range of thresholds from 0 to 100% (in 5% increments) applied to data from field trials of three condition monitoring methods, the Grid System, standard CSM method and a rapid method carried out at Aberffraw to Abermenai Dunes SAC in 2008..... 116 4.10 Asymmetrical Summary Receiver Operating Characteristics plots of sensitivity against specificity calculated using thresholds from 0 to 100% (in 5% increments) for (a) the standard CSM method and (b) a rapid assessment method against the Grid System as the 'gold standard'. The circles are the pairs of specificity and sensitivity calculated for each 5% increment between 0 and 100%, the middle curve is the line of best fit and the two outer curves indicate the 95% confidence intervals and AUC=area under the curve (summarises the accuracy of the method as a single number, a perfect test will have an AUC close to 1 and poor tests have AUCs close to 0.5 (Hopley and Schalkwyk. 2001). Data were collected during field trials at Aberffraw to Abermenai Dunes SAC in 2008..... 118 4.11 Mean frequency of seven positive indicator species recorded by pairs of volunteer and professional observers using the Grid System method during a field trial at Aberffraw to Abermenai Dunes SAC in 2008. 119 4.12 Mean percentage of sample points assessed as being in favourable condition by pairs of volunteer (n=7) and professional observers (n=4) using the Grid System methodology during a field trial at Aberffraw to Abermenai Dunes SAC in 2008. Bars 121 are ± 1SE.... Pooled sensitivity and specificity according to agreement 4.13 between (a) volunteer observer and (b) professional observers in

condition assessment of plots using the Grid System in fixed

dune grassland during field trials at Aberffraw to Abermenai Dunes SAC in 2008. Circles show sensitivity/specificity values for each unique combination of pairs of observers in volunteer (n=21) and professional groups (n=6), diamond shows pooled sensitivity/specificity respectively and dashed line shows 95% confidence interval.

Chapter 5

5.1	Forms of variogram (a) unbounded (b) bounded (c) pure nugget (Oliver and Frogbrook, 1998)	144
5.2	Layout of sample plots in Hafod y Llan, Snowdonia National Park in 2006/7	150
5.3	Experimental variograms and fitted models for height of six taxa sampled in Hafod y Llan, Snowdonia National Park in 2006/7. Symbols are the experimental semivariances and the line is the fitted model	162
5.4	Maps of kriged predictions for height in cm for six taxa sampled in Hafod y Llan, Snowdonia National Park in 2006/7. Black outline shows site boundary	163
5.5	Map of kriging variance for <i>C. vulgaris</i> ; data collected in Hafod y Llan, Snowdonia National Park in 2006/7. Black outline shows site boundary	164
5.6	Map of interpolated axis 1 and 2 sample point scores from a DCA ordination of all height data. Crosses show sample points at 400 m, 200 m and 100 m spacing (50 m and 25 m not shown),black outline is the site boundary and black lines within site show compartment boundaries. Data collected in Hafod y Llan, Snowdonia National Park in 2006/7	164
5.7	Variance in combined axes 1 and 2 DCA ordination scores and sampling intensities (400 m, 200 m, 100 m distance between points) using presence/absence data from a spatially homogeneous area and a spatially heterogeneous areas in Hafod y Llan, Snowdonia National Park in 2006/7. There are insufficient sampling locations to include sampling intensities of 1600 m, 800 m, 50 m and 25 m	165
5.8	Canonical Correspondence Analysis of taxa height data and six environmental variables collected at Hafod y Llan, Snowdonia National Park in 2006-7. Ordination diagrams of: (a) Axis 1 and 2; (b) Axis 1 and 3 and (c) Axis 2 and 3. Environmental variables are represented by lines and taxa by crosses (sample points are omitted for ease of interpretation)	167
5.9	Sample size required to detect change over a range of 10-90% across different sample intensities (1600 m, 800 m, 400 m, 200 m, 100 m, 50 m, and 25 m distance between points) using combined axes 1 and 2 scores from DCA ordination of (a)	

- 5.10 Sample size required to detect change over a range of 10-90% across different sample intensities (400 m, 200 m and 100 m distance between points) using combined axes 1 and 2 scores from DCA ordination of presence/absence data for (a) homogenous and (b) heterogeneous area. Sample size is converted into number of days of sampling for (c) homogenous and (d) heterogeneous area. There are insufficient sampling points to include sampling intensities of 1600 m, 800 m, 50 m and 25 m.
- 5.11 Sample size required to detect change over a range of 10-90% at a sampling intensity of (a) 50 m and (b) 400 m distance between points (chosen to be illustrative of the relationship across sampling intensities) using combined axes 1 and 2 scores from DCA ordination of presence/absence. Sample size converted into number of days of sampling is also shown for (c) 50 m and (d) 400 m. Data collected at Hafod y Llan in Snowdonia National Park in 2006/7.
- 5.12 (a) Sample size and (b) sample size converted into number of days required to achieve a given power to detect a 20% change at different sample intensities (1600 m, 800 m, 400 m, 200 m, 100 m, 50 m, and 25 m distance between points) using combined axes 1 and 2 scores from DCA ordination of presence/absence data recorded at sample points in Hafod y Llan in Snowdonia National Park in 2006/7.....
- 5.13 Mean number of Phase I land cover types recorded per day according to re-sampling points from different sample intensities (1600 m, 800 m, 400 m, 200 m, 100 m, 50 m, and 25 m distance between points). Phase I data recorded during Habitats of Wales Phase I survey 1979-1997, data across various sampling intensities recorded in Hafod y Llan in Snowdonia National Park in 2006/7. Bars show +/- 1 SE.
- 5.14 (a) Sample size and (b) number of days required to detect a change from unfavourable condition wet heath in 2008 to favourable condition wet heath using a range of power levels (given a Type I error rate of 0.05). This is based on data collected in Hafod y Llan, Snowdonia National Park in 2008. 174

170

171

List of appendices

Appendix		Page
2.1	Questionnaire used as the basis for semi-structured interviews with conservation practitioners in the UK between 2006 and 2008.	221
3.1	Aerial photographs and Ordnance Survey maps as provided to all surveyors involved in the field trial of the National Vegetation Classification approach to vegetation mapping in Hafod y Llan in summer 2008. Aerial photographs and maps not to scale	223
3.2	Large scale maps A-G of National Vegetation Classification community. Data collected in a field trial carried out in Snowdonia, 2008.	226
4.1	Targets, field methods and criteria for JNCC CSM guidance from CSM Guidance for sand dune habitats and the Grid System taken from the CCW Aberffraw to Abermenai Dunes SAC UK0020021 SAC monitoring report. Text in bold indicates aspects used in field trials at Aberffraw to Abermenai Dunes SAC in 2008	233
5.1	Protocol used in grid sampling at Hafod y Llan, Snowdonia National Park 2006-7	234
5.2	Targets for sample points using Common Standards Monitoring guidance for wet heath habitats, amended by Countryside Council for Wales for use across the Eryri Special Area of Conservation (JNCC 2005; NT 2004)	236

Chapter 1

General introduction

•

1.1 Background to environmental monitoring

1.1.1 Monitoring worldwide

As concern about the environment has increased, so have the number and breadth of monitoring programmes (Spellerberg, 2005). Programmes range from the global through regional to local, and cover many different aspects across atmospheric, marine and terrestrial ecosystems. There are indices of changes in global biodiversity such as the Living Planet Index (Loh *et al.*, 2005; Collen *et al.*, 2009) and atlases of distribution for threatened ecosystems (UNEP, 2009). Much monitoring activity (or at least data analysis and presentation) has been in response to the targets agreed in the Convention on Biological Diversity, notably 'to achieve by 2010 a significant reduction of the current rate of biodiversity loss at the global, regional and national level' (UNEP, 2002). Many countries now have National Biodiversity Strategies and Action Plans, which include monitoring of status and changes in biodiversity at national scales. However, despite the recent expansion in monitoring activities and with the 2010 deadline imminent, there is insufficient information to make assessments of progress for most species, habitats and ecosystems (RS, 2003; Green *et al.*, 2005; Pereira and Cooper, 2006; Teder *et al.*, 2007).

For a few taxa in some regions, good knowledge of current status and certain recent trends at the species level exist. For example, time series of population estimates are available for many species of British birds (Gregory *et al.*, 2002). Some countries also have useful measures of abundance for other taxa, such as commercially exploited fish species. However, this level of information is frequently unavailable, particularly for groups of organisms such as invertebrates and fungi, and where there is information to make an assessment about change, it is subject to considerable uncertainty. This uncertainty is rarely reported when information is combined into national or international indicators, such as the Living Planets Index which is based on population trends for >4,000 vertebrate species but is biased towards species from temperate regions, has >25% missing values in its time series and does not differentiate between the quality (confidence) of the datasets used (Pereira and Cooper, 2006; Collen *et al.*, 2009). Marsh and Trenham (2008) stated that there is a lack of well designed sampling strategies suitable for monitoring, and that

researchers must create tools for designing effective monitoring programmes and make these widely available.

1.1.2 Monitoring in Europe

While many countries committed to reduce the rate of loss of biodiversity by 2010, European countries went a stage further by pledging to halt the loss of biodiversity in Europe by 2010 (EC, 2009a). This target requires integrated monitoring systems over large, supra-national spatial scales and over long time scales (Balmford et al., 2005; Lengyel et al., 2008b). An example of a large-scale monitoring system designed for species is the Pan-European Common Bird Monitoring scheme which attempts to quantify trends in populations of European breeding birds, and to develop an index of biodiversity to measure progress to the 2010 goals (Gregory et al., 2005). Currently, no such co-ordinated monitoring of habitats exists at the European level, although the Natura 2000 network of sites created under the Habitats Directive legislation is intended to eventually lead to integrated habitat monitoring (EC, 2009b). So far, a series of Special Areas of Conservation (SACs) have been designated for species and habitats as part of the Natura 2000 network. Other advances in co-ordinated habitat monitoring are the CORINE Biotopes and Land Cover projects which provide information on the changes in major land cover types and the BioHab project which developed and tested field-based methods using a typology based on plant life forms (Devillers et al., 1991; Bunce et al., 2008). Despite these developments, most monitoring programmes in Europe remain small in scope both spatially and temporally (Balmford et al., 2003).

1.1.3 Monitoring in the UK

The UK invests considerable effort into the development and delivery of monitoring programmes; a recent review of environmental monitoring (including monitoring and surveillance) listed over 400 environmental monitoring activities (Slater *et al.*, 2006). Furthermore, they emphasize that this is a partial review, with marine data excluded, along with short term research, base-line data surveys and surveys related to management practices. The list includes at least 100 separate biodiversity monitoring schemes implemented by 30 institutions with a combined annual budget of approx £7

million, not counting the volunteer inputs which is estimated at nearly three times this figure (Slater *et al.*, 2006; JNCC, 2009b). Monitoring therefore represents around 2% of the total UK biodiversity conservation budget; whether this is a worthwhile investment depends on the extent to which monitoring is meeting its objectives.

There are several notable programmes monitoring biodiversity in the UK. The Breeding Bird Survey carried out by the British Trust for Ornithology (BTO) is considered to be an extremely robust system, using a stratified random design and a highly organised network of 3,000 trained surveyors who sample more than 3,200 sites, providing trend data for over 100 bird species (BTO, 2009). Birds as a group are well covered by monitoring in the UK, as are fish, mammals and higher plants through schemes overseen by the Centre for Environment, Fisheries and Aquaculture Science (CEFAS, 2009), the Tracking Mammals Partnership (JNCC, 2009c) and the Botanical Society for the British Isles (BSBI) (Preston et al., 2002). The Distribution Maps Scheme run by the BSBI is one of the world's longest-running natural history distribution mapping projects and provides information about the abundance, range and changes in the distribution of vascular plants and charophytes in 10km x 10km squares across the British Isles, published as Atlases of the British Flora (BSBI, 2009). There is an issue however with data collected by volunteers such as the BSBI distribution data, in that the loss of a record can be more about loss of a recorder than the plant. The Countryside Survey programme uses a stratified systematic design to monitor vascular plant species and habitats and is a unique 'audit' of the current stock and recent change of the condition and extent of Broad Habitats, landscape features, vegetation, soils and freshwaters across the UK's countryside (CEH, 2009). Another well known scheme is the Butterfly Monitoring Scheme which uses standardised protocols and skilled observers to collect trend data for British butterflies, but recording is uneven in time and space due to the subjective choice of sampling sites and the use of volunteers (Fox et al., 2007).

Despite this monitoring activity, there are still substantial gaps, with no systematic sampling for invertebrates or non-vascular plants in the UK, with 66% of invertebrate species not included in any monitoring activity in Wales (Hockley *et al.*, In draft). Hockley et al. (In draft) further note that the quality of datasets is very varied; only 25% of biodiversity monitoring schemes in Wales are suitable for use in tracking

temporal change, due to inadequate spatial and temporal coverage, biased sampling design and inconsistent protocols.

The statutory monitoring carried out on protected sites in the UK is largely carried out in accordance with the Habitats Directive, which requires Member States to monitor and report the conservation status of species and habitats of European interest. The conservation status is illustrated in three 'traffic light' categories ('favourable' - green, 'unfavourable inadequate' - amber, 'unfavourable bad' - red, plus unknown). These habitats and species are listed in Annexes I and II of the Habitats Directive (189 habitats and 788 species) (UKCHM, 2009) and are protected by means of a network of sites (the Natura 2000 network). In the UK the Joint Nature Conservation Committee (JNCC) advises government on the application and interpretation of the Habitats Directive and developed common standards for the monitoring of nature conservation across the UK (JNCC, 1998). Country statutory agencies (Natural England, Scottish Natural Heritage, Countryside Council for Wales and the Environment and Heritage Service- Northern Ireland) are responsible for implementing this Common Standards Monitoring (CSM) on all protected sites.

In addition to statutory monitoring, a terrestrial protected site in the UK may also be monitored as part of national statutory and NGO schemes such as Countryside Survey (CEH, 2009), Local Change (Preston *et al.*, 2002), the National Inventory of Woodlands and Trees (FC, 2009), the Environmental Change Network (ECN, 2009), the Breeding Birds Survey (BTO, 2009) and the Butterfly Monitoring Scheme (UKBMS, 2009). Monitoring may also be carried out through regional projects such as agri-environmental schemes, Biodiversity Action Plan monitoring (UKBAP, 2008), vice county recording efforts and atlases of specific taxa/species for particular areas. Finally, local, one-off monitoring frequently takes place as part of research and experimental projects. A typical protected site will receive Common Standards Monitoring and several other more targeted surveys.

There are major efforts to effectively co-ordinate monitoring information and data across the UK, for instance through JNCC's Surveillance Strategy (JNCC, 2009b). This identifies gaps and overlaps in the coverage of monitoring schemes (including surveillance) in order to enable monitoring in the UK to become more useful and efficient in the future. Alongside this, the UK Environmental Observation Framework (UK-EOF) has been developed to 'facilitate the ongoing environmental evidence

required to understand the changing natural environment, thus guiding current and future environmental management, policy, science and innovation priorities for economic benefit and quality of life' (ERFF, 2009). One of the objectives of the UK-EOF is to improve data accessibility and quality. The development of data storage and transfer centres is also intended to aid information retrieval and reduce duplication; the UK has a series of Local Records Centres alongside two central facilities, the Biological Records Centre, and the National Biodiversity Network (NBN) Gateway.

The current study focuses on the monitoring of vegetation in the UK which is a large component of environmental and biodiversity monitoring; of the 100 biodiversity activities currently listed by JNCC, 41 include monitoring of vegetation (JNCC, 2009b).

1.2 Vegetation monitoring in the UK

Higher plants are a diverse group, with 4,111 species of vascular and non-vascular plants (including archaophytes, naturalised neophytes and regular casual species) in the UK (Preston *et al.*, 2002). There are several reasons why they are a focal group for monitoring purposes. Firstly, they are reliable since they stay still; this also makes them relatively easy to survey. They respond to large-scale pressures including climate change, anthropogenic disturbance and localised management practices. They are also the main primary producers in terrestrial ecosystems and, thus, are fundamental to ecosystem functioning and provide the physical structure and defining features of most terrestrial habitats (Pereira and Cooper, 2006). Furthermore, the diversity of plants is one of the best available predictors of diversity of other taxa (Sala, 2006) and thus plants represent an ideal group for monitoring the effectiveness of conservation practices and are an important feature in the selection of nature reserves (Ryti, 1992; Balmford, 1998).

Despite the importance of vegetation monitoring, it has been noted that there is a mismatch between conservation practitioners and academics due to a lack of knowledge exchange (Brown and Rowell, 1997; Legg and Nagy, 2006). There are several comprehensive texts on vegetation sampling (Grieg-Smith, 1957; Daubenmire, 1968; Bonham, 1989; Barbour *et al.*, 1999; Elzinga *et al.*, 2001) and

guidelines for monitoring and analysis of data (Clarke, 1986; Goldsmith, 1991; Spellerberg, 1992; Fowler *et al.*, 2000; Sutherland, 2000; Spellerberg, 2005). However, this information does not seem to have reached the relevant organisations in the conservation field; a recent workshop on vegetation sampling in the UK concluded that there is a lack of understanding of the suitability and effectiveness of approaches and recommended the development of vegetation sampling protocols which give good measurability and consistency (JNCC, 2008c). The workshop also noted that analyses and interpretation of datasets needs to be improved to facilitate the use of vegetation data and enhance understanding of habitat quality and change. In order to ascertain how widespread the lack of understanding of vegetation sampling out a series of interviews with conservation practitioners. It is anticipated that these interviews will also provide opinions about the relative effectiveness (with explanatory reasons) of approaches currently used for vegetation monitoring.

The lack of integration between effective published methods and their acceptance and use has attributed to a lack of attention to vegetation sampling in graduate and post-graduate courses (Legg and Nagy, 2006) and also a widespread view that *'i*) monitoring methods are simply an extension of research methods and *ii*) conservation is only a practical extension of the science of ecology, with the same methods and approaches' (Brown and Rowell, 1997). It is likely that a similar situation is present in countries outside the UK (Underwood, 1995; Stohlgren, 2007).

The lack of robust vegetation monitoring programmes is also explained by the wide variety of situations which require vegetation monitoring; there are many different research/programme objectives in many types of vegetation (Stohlgren, 2007). This means that, unlike soil or water monitoring, no particular methods have been accepted as standard and there is confusion over which method is best for a given problem (Royal Society, 2003). There are also difficulties associated with sampling plants; including (i) bias related to natural variation, (ii) variation between observers in species detection ability and (iii) taxonomic problems (Stohlgren, 2007; Bacaro *et al.*, 2009). The first group involve errors caused by natural variation such as phenology and abundance which can cause early or late flowering species to be missed in late or early surveys respectively, and errors caused by the relative abundances of species which means that small and inconspicuous species may be

missed or underestimated (Kennedy and Addison, 1987; Scott and Hallam, 2003). The second group of errors is caused by differences in observer training and ability to recognise species; with large numbers of plant species in most landscapes, and many species containing subspecies and varieties, there is the potential for large error, even between experienced surveyors (Hall and Okali, 1978; Sykes *et al.*, 1983; Prosser and Wallace, 1999; Klimes, 2003). Less experienced observers may miss species through not searching sufficiently thoroughly and observer fatigue can also contribute to errors, particularly at the end of a day or in adverse weather (Nilsson and Nilsson, 1985). Although errors are larger between observers, one study in which the same observer carried out repeat samples within a few days found 17% difference in species lists (Hope-Simpson, 1940). Studies have noted the importance of training in species identification, the use of standard protocols and regular quality assurance (Allegrini *et al.*, 2009; Berti *et al.*, 2009; Marchetto *et al.*, 2009).

The final group of errors is not often considered, and is sometimes known as the 'taxonomic inflation/deflation' problem (Isaac *et al.*, 2004), and results from variation between countries and date of the taxonomic reference system used. This means that the number of species in an area may change due to species or subspecies being lumped together or split apart, or when they are renamed as previously named taxa or new taxa (Nimis, 2001). Although it has been suggested that higher taxon measures (such as genera or families) could resolve this problem (Balmford *et al.*, 1996), this requires good evidence that species diversity is correlated at higher taxonomic levels, which is generally lacking. Many practitioners continue to believe that vegetation monitoring requires reliable identification to species level, since species are still widely regarded as the 'unit' of conservation (Brunialti *et al.*, 2002).

It is possible to quantify the accuracy and precision of plot-based measurements of vegetation (Lundstrom, 2000; Milberg *et al.*, 2008); this is the basis of forest inventory where statistical rigour has long been a preoccupation (Sheil, 1995a; Foster, 2001). In herbaceous vegetation, visual estimation of cover is one of the most common measures of abundance since it is fast and easy but it has been demonstrated that changes in cover must be greater than 20% (e.g. a larger increase/decrease than 30% to 50% or 90% to 70%) before they can be attributed to factors other than annual variation and observer error (Sykes *et al.*, 1983; Kennedy and Addison, 1987; Cheal, 2008). Efforts to reduce the amount of observer error by placing estimates

into cover classes such as the DOMIN scale, which is widely used in Europe, make the results unsuitable for detecting fine scale changes (Goldsmith, 1991). Frequency calculated from presence/absence measurements in plots are recommended as being quick, consistent and objective (Bonham, 1989; Bullock, 1996; Ejrnaes and Bruun, 2000; Ringvall *et al.*, 2005; Ramsay *et al.*, 2006) although they have limited sensitivity to change and are dependent on size and number of plots (Bonham 1989). Estimating cover through frequency of species at points or sub-plots within larger plots takes longer but is consistent between observers, robust and able to detect changes during the growing season, although measurements are still subject to the general problems outlined previously (Bonham, 1989; Ejrnaes and Bruun, 2000; Vittoz and Guisan, 2007).

1.2.1 Species and communities

Individual plant species are easier to monitor than communities as they can be recorded as present or absent at specific sites or measures of abundance such as cover (Goldsmith, 1991; Elzinga et al., 2001). Presence/absence records are used to report temporal change, examples being; the Botanical Society for the British Isles (BSBI) 'Local Change' programme which reports change in distribution for each vascular plant species in Britain in 10 km² grids (Preston et al., 2002) and the Countryside Survey which provides percentage change over time in abundance of common plant species such as stinging nettle and bramble across the UK (Haines-Young et al., 2002). When correlated with other factors, this sort of information about trends in distribution of species is very useful for assessing large scale impacts of climate and pollution. Single-species data is also useful in site-based monitoring where particular species are of interest because they are supposed to indicate climate change as in the arctic alpines in Snowdonia (CCW, 2009) or be particularly sensitive to management change (Elzinga et al., 2001). Even so, this still only provides information about individual species and it is sometimes necessary to monitor changes in whole plant communities.

Communities are defined as assemblages of species living in the same place and are described according to the presence and abundance of key species within the community. There has long been debate on the conceptualisation of what constitutes

a community of plants. Some plant ecologists, beginning with Frederick Clements viewed plant associations (or assemblages) as stable, self-sustaining communities which proceeded in a series of orderly successional stages from early 'seral stages' to 'climax' (Clements, 1936; Kent and Coker, 1992). Clements' theories were widely accepted and led to a focus on description, classification and sub-division of a landscape into a mosaic of homogenous patches each representing a discrete community. These models of vegetation cover dominate mapping of plants and are still used in climate modelling, ecosystem delineation and habitat mapping (Rowe, 1996; Mueller-Dombois and Ellenberg, 2002).

In contrast, many plant ecologists consider plant communities to be subjective entities; Gleason (1922, 1926) observed that plant species were clumped in distribution on the landscape but not as tightly associated with other species as with abiotic factors. Gleason tried to shift the emphasis from sampling in restricted homogenous patches of subjectively defined communities to sampling species distribution across larger areas. Despite this, the Braun-Blanquet relevé method which involves subjectively selecting 'representative' homogenous stands within different communities remains a dominant approach in continental Europe and is still one of the most widely used vegetation mapping approaches in the world (Stohlgren, 2007). Many ecologists assert that communities have an important role in shaping our understanding of plant interactions (Palmer and White, 1994). However, there is still concern that the level of subjective and arbitrary decision making involved in the description of plant communities outweighs their potential use (Dale, 1994).

1.2.2 Monitoring plant communities: problems and methods

Plant communities present considerable challenges for monitoring (Stohlgren, 2007). The first problem being the definition and the delineation of their occurrence on the ground; classifying and mapping heterogeneous landscapes into discrete 'homogenous' patches involves considerable error because, in reality patches are never completely homogenous (Kuchler, 1973; Cherrill and McClean, 1999a). Even placing samples in patches of homogenous vegetation relies on being able to locate a suitable patch and only provides information relevant to that patch. The second problem is the lack of understanding of community processes; the interaction of many species with each other and the environment makes it difficult to predict

whether or how the community as a whole will respond to management (Elzinga *et al.*, 2001). As a result, practitioners have applied standard methods that monitor all species (or as many as possible). However, the more species which are monitored together, the harder it is to interpret the various interactions taking place, and any single monitoring design may not work equally well for all species (Kenkel and Podani, 1991). For example, a design may over-sample common species and result in 100% frequency (a 'saturated' count) or under-sample and possibly miss altogether rare or patchily distributed species (Scott and Smart, 2006). Nested sample designs may overcome this but require careful thought and trialling during the planning stage (Critchley and Poulton, 1998).

Another problem with the monitoring of communities is the high level of observer variability (and the difficulty of measuring it satisfactorily). Vegetation studies which require that all species are identified are subject to considerable error when the number of sampled plants is high; this is particularly with plants that are in a vegetative state and inconspicuous or are similar looking (Sykes *et al.*, 1983; Scott and Hallam, 2003). Goldsmith (1991) concluded that monitoring of communities will always be imprecise. Nevertheless, it is important to attempt to quantify the errors involved and ascertain the effect that they have on the detection of temporal change. A final problem for monitoring of communities is the resource cost in terms of data collection, transfer and analysis; estimating and recording all species in a plot is very time consuming compared with just a few species.

Plant community monitoring methods can generally be classed into quantitative or qualitative techniques (Table 1.1). Quantitative techniques tend to be of higher intensity than qualitative methods, and thus are more time consuming but conversely provide more detailed information. Sampling is the most commonly used approach in quantitative monitoring; this measures a portion of a statistical population in order to estimate mean and variance of that population (Bonham, 1989; Schreuder *et al.*, 2004). Due to problems associated with monitoring plant communities, sub-sets of the community are often measured such as indicator species chosen to correspond with a community attribute, specific groups of species or a functional guild (Gillison and Carpenter, 1997; Critchley, 2000; Pywell *et al.*, 2003). Complex community information can also be reduced at the analysis stage, either by calculation of indices such as species richness or conversion to a small range of axes scores through

multivariate analysis. Although change in community composition can be measured through the trajectories of axes scores through time (Legendre and Anderson, 1999; Podani *et al.*, 2005), the results can be hard to interpret and determining the direction of change and its cause e.g. management, is difficult. This means that multivariate techniques cannot always provide answers to commonly posed questions about temporal change in whole communities. Quantitative approaches to the monitoring of change in vegetation communities are also sensitive to the nature of the measurement (e.g. density, abundance or frequency) and sampling design. When sampling units are compositionally different from each other due to spatial variability then detecting temporal change is complicated, although analyses such as repeated measures can overcome this (Grieg-Smith, 1957).

Qualitative techniques can provide an useful alternative to quantitative monitoring (Table 1.1). They tend to be low intensity and rapid, often used to compare many sites and can make use of inference based on practitioners' field experience to assess systems and processes (Hockings, 2003). Techniques include site condition assessments, photo monitoring and qualitative assessment of remotely sensed data. A widely used system of condition assessment in the UK is Common Standards Monitoring (JNCC, 2006) which attempts to measure change in the condition of plant communities using categories such as 'unfavourable' and 'favourable'. Problems with qualitative assessments are that the measurements are often biased, difficult to repeat and provide a coarse measure of change with no indication of power or error. This study recognises vegetation mapping as a further type of plant community monitoring involving detection of change in boundary position and extent of particular vegetation communities over an area (Elzinga *et al.*, 2001). Although not necessarily intended as a monitoring tool (Rodwell, 1997), plant community mapping is nevertheless used as such (Dargie, 1993; Stevens *et al.*, 2004b).

Monitoring methods for plant communities vary in spatial scale from small patches at a single site to catchments, landscapes and regions, and vary tremendously in purpose from scientific research studies to evaluation of site-based management practice (Royal Society, 2003). Choosing the right method requires a careful consideration of the objectives of the assessment as well as the trade-offs between usefulness, completeness and required effort in terms of time and other resources. No one method is best for all purposes.

Table 1.1 Typology of vegetation monitoring methods, including information about scale, measurement of species and community change, main purposes, strengths and weaknesses, major sources of error, cost and examples currently in use in the UK.

	Monitoring approach		
	Qualitative	Quantitative	Vegetation mapping
Spatial scale	Series of sites	Site or series of sites	Landscape-site
Index of change	Community condition	Species composition/diversity	Extent of patches of different vegetation communities
Main purposes	 (i) warning system (ii) assess impact of management practices (iii) report against targets of proportion of designated sites in favourable condition 	 (i) assess impact of management practices (ii) investigate causes of change (iii) understand ecosystem functioning (iv) hypothesis testing 	 (i) provide information for environmental assessment and site management planning (ii) assess large-scale pressures and impact of management
Strengths	Rapid Can utilise expert opinion	Utilises statistics to obtain estimates of mean and variance Can detect small change	Rapid Resulting information easy to communicate
Weaknesses	Generally only picks up large changes	Time consuming Sampling design, field survey and analysis may require expertise	Requires skilled observers
Major sources of error	High subjectivity in assessment	Unsuitable sampling design which misses change or finds erroneous change (Type I and II errors)	Misclassification of vegetation type Spatial error in position of boundaries
Cost	Low	High	Medium
Examples in the UK	Common Standards Monitoring (England, Wales, Northern Ireland) Site condition monitoring (Scotland)	Local Change (BSBI ¹) Environmental Change Network (ECN) Countryside Survey (CS) Long-term monitoring/research sites (often experimental)	Phase I Phase II (using National Vegetation Classification)
Relevant studies	(Shaw and Wind, 1997; Oliver, 2002; Gibbons and Freudenberger, 2006; Jackson and Gaston, 2008; JNCC, 2008a; SNH, 2008; Bussotti <i>et al.</i> , 2009)	(Grubb, 1982; Belsky, 1985; Bakker <i>et al.</i> , 1996; Thomas and Meyer, 1996; Lundstrom, 2000; Sheil <i>et al.</i> , 2000; Vaughan <i>et al.</i> , 2001; Bakker <i>et al.</i> , 2002; Marriott <i>et al.</i> , 2004; Taverna <i>et al.</i> , 2005)	(Dargie, 1993; Cherrill and McClean, 1995; Elzinga <i>et al.</i> , 2001; Stevens <i>et al.</i> , 2004b)

¹ Botanical Society for the British Isles

1.3 Definition of monitoring and surveillance

The meaning of the terms monitoring and surveillance may appear to be an academic digression (Spellerberg, 2005) but definitions can help to clarify purpose, as Allen (2006) notes; 'the clearer we are about what type of activity is being used, the more likely the activity is to succeed'. The confusion over terms both in the literature and amongst practitioners leads to miscommunication about objectives and practice and this is partly because the definitions of the terms really do overlap as the Oxford English Dictionary definitions demonstrate:

'to monitor- observe, supervise, keep under review, measure or test at intervals, especially for the purpose of regulation or control.

'surveillance- supervision for the purpose of direction or control, superintendence.'

Often monitoring is taken as any programme of repeated surveys or measurement. However, Tucker et al. (2005) point out that this is merely surveillance if there is no predetermined objective or value against which findings can be measured and note that even daily measurements of rainfall are a type of surveillance. Monitoring is more rigorously and suitably defined as 'the collection and analysis of repeated observations or measurements to evaluate changes in condition and progress toward meeting a management objective' (Elzinga *et al.*, 2001). This distinction between monitoring and surveillance tends to be used in the UK, largely originating from definitions in the Biological Monitoring Handbook by Hellawell (1978) (Fig. 1.1). To avoid repetition, this introduction uses the term monitoring to include all monitoring and surveillance activities. However, in the following data chapters and the discussion chapter, monitoring is distinguished from surveillance following the definitions in Hellawell (1978) in order to clarify the purpose of the methods investigated and to ensure consistency.

	Differentiating between 'monitoring' and 'surveillance' Monitoring is the collection and analysis of repeated observations or measurements to evaluate changes in condition and progress toward meeting a management objective. Surveillance refers to repeated systems of measurements carried out to draw general conclusions about trends and changes but with no predetermined objective or value that guides what the findings ought to be (Hellawell, 1978; Brown and Rowell, 1997; Elzinga <i>et al.</i> , 2001; Hurford and Perry, 2001; Spellerberg, 2005; Tucker <i>et al.</i> , 2005; Allen, 2006; Legg and Nagy, 2006)
•	Using 'monitoring' and 'surveillance' interchangeably The word 'surveillance' originates from the French to 'watch over', and may be used interchangeably with 'monitoring'. Both imply repeated recording over time (JNCC, 2009b)
	Using 'monitoring' to describe monitoring and surveillance Monitoring describes any activity involving taking repeated measurements over time in order to detect change in measured variables (Vaughan <i>et al.</i> , 2001; Yoccoz <i>et al.</i> , 2001; Green <i>et al.</i> , 2005)
	 Using 'monitoring' to describe monitoring and surveillance, and recognising subsets of monitoring (i) result monitoring is the measurement of change in a pressure, as an outcome of management, (ii) outcome monitoring is the measurement of change in the conservation asset of interest and (iii) surveillance monitoring is monitoring carried out in the absence of specific management actions. (Urlich and Brady, 2003) (i) status monitoring is quantitative description of the universe against a threshold as it changes with time and (ii) trend monitoring detects whether there is autocorrelation in temporal change as a basis for linking changes to specific events or interventions. (de Gruijter et al., 2006)
Figu	ire 1.1 Alternative definitions of monitoring and surveillance in use in conservation literature.

The root of the word monitoring means 'to warn' and one essential purpose of monitoring is to raise awareness about problems. For instance, early detection of an invasive species that threatens priority species or habitats may mean that it can be dealt with before it becomes well established. As well as providing important information about biodiversity change, environmental monitoring is indispensable in measuring management actions (Vaughan et al., 2001; Szaro et al., 2005). Management is designed to meet specific objectives and this defines what is measured during the monitoring; management can then be changed if the monitoring reveals a failure to meet the objectives. This framework is termed 'adaptive management' (Ringold et al., 1996; Light, 2001; Smit, 2003; Sabine et al., 2004), due to the recognition that understanding of ecosystems is incomplete and that management is essentially experimental (Brunner and Clark, 1997; Salafsky et al., 2002). The adaptive approach can be compared to the use of a thermostat to register and respond to pluses and minuses from the desired temperature, or state. The approach encompasses the use of the 'precautionary principle' which says that management has to continue in spite of incomplete knowledge (Contamin et al.,

2009). Monitoring programmes tend to focus on changes in the resource of interest (the desired state of the resource) as the most meaningful measure of management performance, but this does not necessarily provide the best information to decide how to change the management if the outcome is unsatisfactory to meet this need (Shaw and Wind, 1997; Alexander and Rowell, 1999). Other types of monitoring are recognised, for example by Ulrich and Brady (2003) who, along with *outcome monitoring*, also list *result monitoring* as the measurement of change in a pressure (anthropogenic and natural disturbances) as an outcome of management, and *surveillance monitoring* as monitoring carried out in the absence of specific management actions (Fig. 1.1).

1.4 Rationale for the project

Environmental managers wish to either maintain or improve habitats and therefore need to detect change in extent and quality of natural vegetation in order to monitor the effectiveness of current biodiversity/conservation management and to identify successful management interventions (Hockings, 2003). Change can be assessed at many levels from genes, species, populations to ecosystems and their associated processes or functions (Noss, 1990; Gaston and Spicer, 2004). Confidence in the detection of change is related to the reliability with which each measurement of state is made; only when reliability is quantified (accuracy, bias and consistency over space and time) can statements about status *and* change be made with known confidence. Generally higher levels of confidence require most intensive and therefore costly monitoring. Management practitioners therefore need to balance the risks of misjudging the effectiveness of management with cost in order to select the monitoring methodology most suited to their aims (Linke and Norris, 1993; Nichols and Williams, 2006).

However, it is claimed that there is insufficient information available about the inherent errors in commonly used vegetation monitoring approaches such as species inventory, vegetation mapping and rapid 'quality' appraisal, which means that managers are unable to make confidence/cost trade-offs (Legg and Nagy, 2006). This could have important implications for environmental reporting and management. This study aims to empirically validate current vegetation monitoring approaches in order to identify fit-for-purpose monitoring methods.

1.5 Specific context: the problem of monitoring vegetation change across large, complex sites

Successful monitoring of vegetation change across large, complex sites is a widespread problem, both in the UK and beyond. The nature reserve, Hafod y Llan (1043 ha of upland in Snowdonia National Park) typifies the problem. When the National Trust acquired the property in 2000, the site was in a degraded condition due to over grazing (NT, 2000), and management was instigated to reduce the level of sheep grazing and introduce welsh black cattle with the intention of bringing the priority habitats (Table 5.1) into favourable condition. A habitat is defined as being in favourable condition when its conservation objectives are met. Since Hafod y Llan is seen as a 'flagship' for the National Trust to demonstrate 'sustainable and conservation friendly farming' (NT, 2004), management must consider not only favourable condition but also the requirements of the livestock, and the two might not be compatible. Monitoring was required to assess change across the site, and in 2006 (the outset of this study) the final monitoring protocol was being put together and various objectives, opinions and obstacles were evident (Table. 1.2).

Table 1.2 Management objectives, opinions and obstacles for vegetation monitoring on Hafod y Llan,Snowdonia National Park in 2006. Information obtained from National Trust reports and variousmeetings with H. Buckingham and K. Jones from the National Trust in 2006 (NT 2000; 2004).

Management objectives

• Improve the condition of priority habitats currently in unfavourable condition due to overgrazing by sheep. Decrease the dominance of unpalatable, competitive species: *Molinia caerulea, Nardus stricta, Juncus effusus* and *Pteridium aquilinum.*

Opinions

- There is so much expert opinion around, we know what a habitat in favourable condition looks like therefore we don't need to monitor; we can just 'look over the wall'.
- Common Standards Monitoring does not work for upland habitats because it is too subjective.
- We need to worry about what is right for the livestock, not for 'conservation'.

Obstacles

- Vegetation sampling had been carried out on the site since 1999 but the sample size was insufficient to detect change with suitable confidence and it was incompatible with CSM monitoring.
- There was insufficient evidence to predict what effect the changes in management would have on the plant species and communities.
- There has been no validation of CSM so insufficient evidence exists about its sensitivity to the change taking place.
- No standard method has been identified that is suitable for monitoring vegetation change across such a large site of varied vegetation.

There was already some monitoring in place at the site: seven fixed point quadrats of 2 m x 2 m had been established in 1999 and re-recorded in 2003; a visual estimate of cover was made for each species in each quadrat and photographs were taken. These quadrats were re-recorded and re-photographed at the outset of this study in 2006. Whilst the example shown of one quadrat placed in an area of degraded dry heath shows a definite recovery at this point, with a large increase in cover of dwarf shrub species, *Pteridium aquilinum* is also increasing (Fig. 1.2). Analysis across the seven quadrats using a paired samples t-test (with a threshold of 0.05) of 1999-2003, 2003-2006 and 1999-2006 cover data for each species showed no significant differences between years (p>0.05 for all species). However, analysis of the vegetation height data, again using a paired samples t-test showed a significant increase between 1999 and 2006 (p = 0.01) (Fig. 1.3).

These are quite different results and it is not possible to reconcile them as data from just seven quadrats does not allow change to be quantified with sufficient precision. Besides the simple fact that seven is a very small sample size this is compounded by a lack of stratification (Elzinga *et al.*, 2001), potential bias in the placement of plots (Brown, 2001) and potential between observer error in visual estimates of cover (Kennedy and Addison, 1987). Typical of many conservation monitoring programmes, the results of this monitoring suggest that it is largely a waste of resources (Hurford and Perry, 2001; Legg and Nagy, 2006; Pereira and Cooper, 2006). In 2006, the National Trust were keen to find an effective method to monitor habitats on the site but were unsure how to proceed in Hafod y Llan.


Figure 1.2 Photo monitoring of a fixed point quadrat at Hafod y Llan, Snowdonia National Park in 1999 and 2006 (no record of photo taken in 2003).



(b)

(a)



Figure 1.3 Mean cover of (a) dwarf shrub species and (b) competitive species recorded in seven Fixed Point Quadrats in Hafod y Llan, Snowdonia National Park in 1999, 2003 and 2006. Bars show +/- 1 SE.

1.6 Thesis objectives

The principle aim of this thesis is to assess the reliability and effectiveness of a range of vegetation monitoring methodologies currently employed in the UK and to explore potential alternatives.

The specific research objectives are:

i. To ascertain the range of stakeholder perceptions of current vegetation monitoring approaches employed in the UK.

ii. To determine the reliability of vegetation mapping to measure state.

iii. To determine the consistency of different site condition assessment methods to measure state.

iv. To investigate the ability of a quantitative vegetation survey method to predict species and plant community distributions.

v. To examine the relationship between sampling intensity and spatial variation and the effect of both on the ability to detect temporal change.

vi. To evaluate the implications of using quantitative or qualitative vegetation monitoring approaches to describe state and detect temporal change.

1.7 Thesis structure

The structure of the thesis is summarised in Figure 1.4. The chapters are selfcontained and each contains an abstract, introduction, methods, results, discussion and conclusion. The thesis is divided into a total of six chapters. Chapter Two investigates the first research objective (i), by reporting responses of semi-structured interviews with a range of conservation practitioners and using these results to identify approaches to vegetation monitoring for further investigation. Addressing the second objective (ii), Chapter Three takes one of these approaches, vegetation mapping, and assesses the reliability of mapping with the most widely used sitebased system in the UK, the National Vegetation Classification. Maps of the same site produced by seven experienced surveyors are compared with each other and an earlier survey with the primary objective of determining the amount of observer error. Chapter Four tackles the third objective (iii), taking another of the vegetation monitoring approaches identified, site condition monitoring, and compares the reliability and consistency of the systems used as part of UK-wide 'Common Standards Monitoring'. Diagnostic test methodology is used to assess three different methods of Common Standards Monitoring currently in use in the UK. The effect of observer experience is also investigated.

Chapter Five addresses objectives four (iv) and five (v). It explores the feasibility of using quantitative sampling systems to describe the distribution of single species, groups of related taxa and plant communities. Semi-variograms are calculated for each variable and parameters used to produce kriged predictions across the site. The influence of sampling intensity on the sampling effort required to detect temporal change is also investigated, along with the impact of type of measurement, magnitude of change and level of power. The sampling effort required to detect a change in condition category using qualitative monitoring is also calculated.

Chapter Six is the general discussion of the thesis and considers the sixth objective (vi). Strengths and weaknesses of the various approaches to vegetation monitoring are discussed and conclusions drawn. Finally, implications for policy and practice are made along with recommendations for future research.



Chapter 2

Vegetation monitoring methods in the UK: the practitioners' perspective

2.0 Abstract

There is a concern that much conservation monitoring fails to meet its objectives because many practitioners have a limited understanding of the sample design and statistical analysis necessary to detect temporal change. This study sets out to ascertain the range of stakeholder perceptions of current vegetation monitoring approaches employed in the UK. Semi-structured interviews were carried out between 2006 and 2008 with sixty land managers, consultants, advisors, policy makers and researchers involved in conservation management across the UK. There was a tendency for the respondents to score quantitative methods as more effective than qualitative alternatives. Reasons for scores and criteria listed as necessary for effective methodology demonstrated that the practitioners value the practical aspects which make quantitative methods repeatable (standard protocols, training of observers etc) but lack understanding of theoretical aspects (statistical power, size of change to be detected etc). Only five practitioners mentioned the need to detect specified levels of change or to carry out power/sample size calculations and no one mentioned the importance of type I and II errors.

Practitioners were in agreement that condition monitoring and vegetation mapping involve great subjectivity which leads to high levels of intra and inter observer error. There was a particular concern that the different approaches used as part of Common Standards Monitoring, the condition assessment system for protected sites in the UK, may produce varying condition assessments. This study recommends that the following research is carried out: (i) empirical validation of the various condition monitoring approaches used in Common Standards Monitoring, (ii) quantification of the repeatability of habitat mapping using the National Vegetation Classification as an approach to monitoring and (iii) exploration of the ability of quantitative methods to detect specified change.

2.1 Introduction

Conservation monitoring has been criticised for being poorly organised and inefficient, with many projects failing to meet their stated objectives (Yoccoz *et al.*, 2001; Legg and Nagy, 2006). This is despite widespread agreement in published literature over the requirements for effective monitoring. These requirements can be summarised into the following characteristics: consistent protocols (Bonham, 1989; Ringold *et al.*, 1996; Treweek, 1996; Urlich and Brady, 2003; Green *et al.*, 2005; Pereira and Cooper, 2006); objective and quantitative measures (Grieg-Smith, 1957; Hurford and Schneider, 2006; Legg and Nagy, 2006); minimal between-observer error (Grieg-Smith, 1957; Bråkenhielm and Qinghong, 1995; Hurford *et al.*, 2001; Milberg *et al.*, 2008) and sufficient power to detect change of a stated level (Green, 1979, 1989; Critchley and Poulton, 1998; Brown, 2001; Foster, 2001; Stefano, 2003; Seavy and Reynolds, 2007).

This disparity between theory and application has been explained by inadequate coverage of monitoring design in degree and postgraduate courses and by a general lack of accessible, practical information on aspects of monitoring design (Legg and Nagy, 2006; Slater *et al.*, 2006). It has also been noted that the complexity of objectives of vegetation sampling and the range of types of habitat leads to numerous different methods which creates a bewildering choice (Royal Society 2003; Stohlgren 2007). A recent workshop on vegetation sampling in the UK which consisted of researchers and senior conservation practitioners concluded that there is a lack of understanding of suitability and effectiveness of methods used in vegetation monitoring and recommended the development of protocols which give good measurability and consistency (JNCC, 2008c).

Quantitative techniques for vegetation monitoring are generally chosen in situations where it is necessary to quantify the error in estimates of species' variables. Aspects of plot-based quantitative sampling such as location of sampling points, size of plot and type of measurement (cover, frequency etc) have been well studied (Sykes *et al.*, 1983; Kennedy and Addison, 1987; Bråkenhielm and Qinghong, 1995; Klimeš *et al.*, 2001; Kercher *et al.*, 2003; Ringvall *et al.*, 2005). The influence of surveyor training, experience and effort on quantitative sampling has also been explored (Sykes *et al.*, 1983; Scott and Hallam, 2003; Archaux *et al.*, 2006).

Qualitative methods for vegetation monitoring such as condition monitoring are known to be less precise than quantitative sampling (Goldsmith, 1991; Gibbons and Freudenberger, 2006) and are only appropriate for situations where it is not necessary to quantify the error. Condition monitoring is widely used in the UK since the introduction of Common Standards Monitoring (CSM) by the Joint Nature Conservation Committee (JNCC). CSM consists of series of habitat specific guidance for the condition assessment of the majority of interest features (flora, fauna, geological or physiographical elements) of designated sites (JNCC, 2008a). As in any newly introduced approach, this has led to some discussion about its suitability (Everett, 2004; Gaston *et al.*, 2006; Jackson and Gaston, 2008).

This study is intended to ascertain opinions about existing vegetation monitoring methods in the UK from a range of people involved in their design and implementation. The aim is to investigate whether there are patterns in what types of method are rated as effective or otherwise and in the reasons given. Although interviews are always value-laden, they are very useful in gaining a general overview of perception and opinion and are often used to bring disparate information together and to categorize varied opinion (Hockings, 2003). Everyone's experience is unique and this will influence the way they view things, for instance researchers and policy makers involved in designing monitoring programmes are more likely to be aware of issues of statistical robustness and detection of change whereas land managers and contractors involved in implementation are more concerned with field logistics and training requirements.

2.1.2 Objectives

This study aims to answer the following questions:

- 1. What vegetation monitoring methods are currently used in the UK?
- 2. How do practitioners rate the effectiveness of vegetation monitoring methods?
- 3. What do practitioners base their rating of methods on?

4. Do practitioner's perceptions about CSM differ with different sets of guidance or different versions of the approach?

5. What are practitioners' views about the criteria necessary for an effective monitoring system?

6. Does the stakeholder category of respondents influence their perception of the effectiveness of different methods?

This study presents the results of sixty semi-structured interviews with conservation practitioners, which were used as an exploratory survey to develop ideas about research into vegetation monitoring methods.

2.2 Methods

2.2.1 People interviewed

Sixty practitioners from across the UK were interviewed between March 2006 and March 2008. Relevant people were contacted using a snowball sampling technique (Trost, 1986; Bryman, 2001); initially people were contacted at conferences, through existing networks or by searching on the internet, after which further suggestions were followed up. Relevance was defined as 'currently involved in the implementation of vegetation monitoring methods'. People from across the UK and from a range of stakeholder groups and organisations were included to ensure a broad coverage of experience and opinions (Tables 2.1, 2.2 and 2.3).

 Table 2.1 Location of sixty conservation practitioners interviewed about vegetation monitoring in 2006-8.

Primary location	No. people interviewed	
England	6	
Scotland	14	
UK	13	
Wales	27	

 Table 2.2 Stakeholder status of sixty conservation practitioners interviewed about vegetation monitoring in 2006-8.

Stakeholder group	No. people interviewed
Scientific advisor	11
Contractor/consultant	10
Land manager	14
Policy maker	11
Researcher	14

 Table 2.3 Organisations employing sixty conservation practitioners interviewed about vegetation monitoring in 2006-8.

Organisation		No. people interviewed
Statutory organisation		
Joint Nature Conservation Committee		1
Countryside Council for Wales		17
Scottish Natural Heritage		7
Natural England		4
Environment Agency Wales		1
Forestry Commission		1
Gwynedd Council		1
Scottish Agricultural College		2
	total	34
University/research institute		
University		4
Centre for Ecology and Hydrology		6
Macaulay Institute		2
	total	12
Non Governmental Organisation		
ADAS		1
Botanical Society for the British Isles		1
Butterfly Conservation		1
National Trust		2
Plantlife		1
Snowdonia Mammal/Bat Group		1
Snowdonia National Park Authority		1
Woodland Trust		1
	total	9
Consultancy		
Private consultant		5
	total	5

2.2.2 Interview questions and procedure

The interviews began with recording basic information about each respondent, including who they worked for, where they worked and which stakeholder group they belonged to.. A standard set of questions (Appendix 2.1) was then used as the basis for the interviews; questions were mainly open-ended, starting with what type of work they were involved with, what vegetation monitoring methods they used and why they chose to use those particular methods. Next, interviewees were asked to score the methods they had listed in terms of effectiveness at meeting the specific programme objectives, from 0 to 10, and to list the most important criteria for an effective monitoring system. Finally they were asked some specific questions about CSM,

including what type of guidance or version they used and how effective it is for their work.

The questionnaire was originally intended to be self-applied, however a trial showed that participants' answers were too brief without explanations and prompts from an interviewer. To overcome this, the questions were used as the basis for semi-structured interviews, with the original form of the questionnaire retained to ensure comparability of data. Use of semi-structured interviews allowed the participants to expand their answers and discuss issues freely, since the interview questions were used as a guide rather than a schedule, with the interviewer asking the same questions each time, but altering their sequence and probing for more information as necessary (Fielding, 1993; Hall and Hall, 1996). The interviewer could also ask further questions in response to significant answers (Bryman, 2001). During the interview, answers were recorded with as few changes and as little interpretation as possible.

Before the interview, a short rationale was provided, including the broad area of the PhD research, the sort of information sought and why the participant was selected. Permission to use data from the answers in the published thesis or other publications was obtained, with the assurance that it would all be anonymous. The same interviewer (the author) was used throughout the process and potential interviewer-induced bias was avoided by reading exactly the same questions to everyone and not commenting, either positively or negatively, on participants' responses (Bryman, 2001).

Depending on people's time and interest, they either only answered the basic questions or answered and expanded on them. Some of the interviews developed into long discussions focusing on various topics including the history of conservation monitoring in the UK, the development and application of CSM and what makes a vegetation monitoring method effective. These were extremely informative in building up an overall perspective of the state of vegetation monitoring in the UK and what future research was required. Towards the end of the process, people came forward to take part in the interview as they were particularly interested in the subject; this raises a concern that these interviewees may represent extreme viewpoints about vegetation monitoring. Interviews lasted from 30 minutes to over one hour.

2.2.3 Analysis

Data were entered into an Access database for storage and manipulation. Basic calculations were performed in Excel. For guestion 1, methods were collated into six categories, with any quadrat or transect re-visited over time in the 'permanent plot' category. When participants listed more than one version of a single method category, it was only counted as one. These counts were converted into proportions. For questions 2, 4 and 6, participants' scores for each category of method/version of CSM/all methods by stakeholder group were averaged. It should be noted that sometimes participants listed two or three versions of a single method; in this case all their scores were used in order to evaluate the separate versions as fully as possible (although this may result in some participants having an unduly large effect on average scores). Where possible, scores for different groups were analysed using a one-way ANOVA (Analysis of Variance) test and checked for significant differences using Least Squares Difference at a significance threshold of 0.05 in SPSS (SPSS, 2003). In the case of the sets of guidance for different habitats, only one score was available for the coastal, grassland, vascular plants and lower plants guidance and so these were excluded from the ANOVA.

Questions 3 and 5 involved the coding of responses about reasons for scoring and criteria for effective methods. Lists of responses were explored in order to draw up a number of categories to encompass the majority of reasons or criteria. These were then ordered into frequency tables.

2.3 Results

2.3.1 Vegetation monitoring methods in use in the UK and how practitioners rate their effectiveness

The methods which people listed can be classified into six main categories, with methods which could not be categorised or which were one-offs classed as 'other' (Table 2.4).

Table 2.4 Description of methods listed by conservation practitioners interviewed in 2006-8 about vegetation monitoring in the UK. Four of the monitoring methods are defined by specific publications: Countryside Survey (Haines-Young *et al.*, 2002); Common Standards Monitoring (JNCC, 2004c); Environmental Change Network (Sykes and Lane, 1996); Phase I (JNCC, 2004d) and National Vegetation Classification (Rodwell, 2006).

Monitoring method	Type of monitoring	Description of measurements relevant to vegetation	Plots recorded?
Countryside Survey	Quantitative	Random stratified sample of 1km ² squares across the UK, repeated every 6-9 y (1978, 1984, 1990, 1997/8, 2007)	yes
Common Standards Monitoring	Qualitative	Rapid assessment of condition of habitat features on designated sites across the UK, repeated every 6 y	sometimes
Environmental Change Network	Quantitative	Species composition and abundance recorded in fixed plots at sites across the UK, repeated every 3/9 y	yes
Phase I/National Vegetation Classification	Vegetation mapping (Qualitative)	Mapping of vegetation into pre-existing classification system	sometimes
Permanent Plots	Quantitative	Any aspect of vegetation sampled in fixed plots and re-recorded over time	yes
Population counts	Quantitative	Other programmes involving repeated censuses of species	yes
Other	n/a	Methods which did not fit any category, including fixed point photography and phenology	sometimes

The most commonly listed method was Common Standards Monitoring, which was used by 34 (57%) of the interviewees (Fig. 2.1).



Figure 2.1 Proportion of sixty conservation practitioners interviewed about vegetation monitoring in 2006-8 who use different vegetation monitoring methods.

When the sixty interviewees were asked to rate each method in terms of how effective it is at meeting its objectives, the Environmental Change Network (ECN) protocols had the highest mean score of 8.4 out of 10, and the Phase I and National Vegetation Classification (NVC) methods had the lowest mean score at 5.2 (Fig. 2.2). Common Standards Monitoring (CSM) methods had the next lowest score (6.7), with the lowest standard error. The high standard errors associated with the Phase I/NVC data, the Countryside Survey (CS) and the population count methods are partly due to the small number of respondents who rated these methods. There were significant

differences between scores given for NVC/Phase I and each other method and between CSM and ECN and CSM and permanent plots (p < 0.05, F = 4.039).



Figure 2.2 Mean rating for vegetation monitoring methods according to sixty conservation practitioners interviewed about vegetation monitoring in 2006-8, 0 = not effective, 10 = very effective. Some interviewees listed the same method category more than once, including different versions of a single method; all of their scores are used to calculate the mean score. Bars show +/- 1 SE. Figures in brackets show number of participants' scores in each group.

Comparing the quantitative methods (ECN, permanent plots, population counts and CS, n = 49) with the qualitative methods (CSM and Phase I/NVC, n = 56), the former had a significantly higher mean score (7.68) than the latter (6.47) (p <0.05, F = 16.86). In a few cases, interviewees were unable to assign a score to a method and provided several explanatory reasons for this: they had insufficient results so far, the methods had not yet been validated, they felt the method was too complex to reduce to a single measure or they did not have experience of a sufficient range of methods to be able to assign relative scores.

Interviewees tended to give similar reasons for giving high scores (7-10); notably that the method is repeatable and sensitive to change, provides detailed information, meets its objectives and follows clear protocols. Reasons for low scores (6 and under) were also fairly consistent, with interviewees saying that low scoring methods were vague and unrepeatable, resulting in inaccurate, unquantifiable data with excess subjectivity. Specific comments about the monitoring methods follow this pattern (Table 2.5). It is interesting that aspects listed as strengths by some practitioners e.g. permanent plots being 'repeatable' and 'plots easy to relocate' are listed as weaknesses by other practitioners e.g. permanent plots 'time-consuming' and 'can't change sample plots'. This is due to varying experience and perspective.

Table 2.5 Strengths and weaknesses of vegetation monitoring methods, according to sixty conservation practitioners interviewed in 2006-8 about vegetation monitoring in the UK.

-	method	strengths	Weaknesses	
highest scoring	Environmental Change Network	repeatable; provides detailed data	time-consuming; frequency is a clumsy measure	
method	Permanent plots	repeatable; plots easy to relocate; detailed data, able to detect a fine scale change; uses objective measurements; errors between people less than differences over time	time-consuming; can't change sample plots; only monitors at chosen locations; cover assessments too subjective	
	Population counts	clear protocols; repeatable; quantifiable; accurate	data not very valuable	
	Countryside Survey	long term dataset; best habitat mapping method developed yet; detailed data	sample size inadequate; tension between the need for consistent methodology and constantly changing objectives; subjective	
\bigvee	Common Standards Monitoring	repeatable; robust; good at identifying problems with site management; rapid	sampling too selective; subjective; hard to interpret the guidance; criteria are too strict; too broad brush; not repeatable or robust; too static; based on little scientific evidence; hard to make it site-specific	
lowest scoring method	Phase I/National Vegetation Classification	consistent data collection; rapid; acts as a standard	too subjective; inaccurate; not enough guidance	

2.3.2 Perceptions of the effectiveness of Common Standards Monitoring

During the interviews it was clear that people had divergent opinions about the effectiveness of different versions of CSM currently used in the different statutory agencies. Of the 46 times CSM was listed, the JNCC standard guidance (used by Natural England) was listed 34 times, the Countryside Council for Wales (CCW) version seven times and the Scottish Natural Heritage (SNH) version five times. The CCW version is accepted as a distinct approach amongst the practitioners interviewed, and refers to an approach developed by CCW as part of an EU LIFE Project which uses the same attributes as the JNCC standard guidance, but specifies slightly different field methods and targets (Brown, 2001; Hurford and Schneider, 2006). The SNH version is not accepted as different from the JNCC standard guidance, but is used in this study to refer to the use of 'site condition monitoring' in Scotland, whereby the JNCC guidance is modified to take account of the needs of extensive sites with more northerly species composition. When the effectiveness rating scores were averaged, the JNCC guidance had a mean score of 6.2 compared with 9.2 and 7.5 for the modifications used by CCW and SNH (Fig. 2.3).





The scores given for the CCW version (n = 7) were significantly different from those given for the JNCC version (n = 34) (p<0.001, F = 13.86). There were no significant differences between scores for the SNH version (n = 5) and either of the other versions.

In terms of sets of guidance, the upland habitat guidance had the highest average score of 8.1, and the woodland guidance the lowest with 5.4 (Fig. 2.4). 'All habitats' is the average of scores, from participants who referred to the method used across all sets of guidance rather than specifying a particular set of habitats.



set of guidance

Figure 2.4 Mean rating for habitat sets of Common Standards Monitoring guidance, 0 = not effective, 10 = very effective. Figures in brackets show number of participants' scores in each group, bars show +/- 1 SE.

Scores for the upland habitat (n = 8) and the woodland guidance (n = 7) are significantly different (p = 0.002, F = 3.13), although neither is significantly different from the scores given for all habitats. Scores for the upland habitat (n = 8) and the fen guidance (n = 2) are also significantly different (p = 0.041, F = 3.13). Reasons given for the high scores for the upland habitat guidance included that the method is repeatable and easy to make site specific, whilst reasons given for low scores for the fen and woodland guidance included their subjectivity and the difficulty of interpreting specific attributes.

2.3.3 Criteria for an effective monitoring system

There was some consensus about the criteria necessary for an effective monitoring method; repeatability, stated objectives, sufficient resources, clear protocols and ease of implementation were each listed by over 20% of the interviewees (Table 2.6).

 Table 2.6 Criteria for effective vegetation monitoring methods, according to sixty conservation practitioners interviewed in 2006-8.

Criteria for an effective method	Number of interviewees	% of interviewees
Repeatability	28	46.7
Stated objectives	21	35.0
Sufficient resources	17	28.3
Clear protocols	14	23.3
Easy to implement	13	21.7
Suited to the experience of observers	12	20.0
Robust	11	18.3
Fit for purpose	10	16.7
Feasible	8	13.3
Consistent	8	13.3
Detect change (including change of a given magnitude)	5	8.3
Flexible	5	8.3
Answer questions	5	8.3
Sustainable	2	3.3

2.3.4 Stakeholder category

There was little difference between the groups of stakeholders in terms of their average rating of methods, with only 0.9 between land managers who gave the highest scores on average and researchers who gave the lowest scores, and no significant differences between categories (p>0.05 for all categories, F= 1.7) (Fig. 2.5). There were insufficient stakeholders to explore the breakdown of scores apportioned to different methods by each group.



Figure 2.5 Mean rating for all vegetation monitoring methods according to groups of stakeholders interviewed, 0 = not effective, 10 = very effective. Figures in brackets show number of participants' in each group. Bars show +/- 1 SE.

2.4 Discussion

Results from this study show that methods involving quantitative sampling were generally perceived to be more effective at meeting objectives than qualitative methods. The vast majority of participants emphasised the importance of repeatability, which is consistently given as a reason for the high scores for quantitative methods, and which is frequently stated as a requirement for monitoring in relevant literature (Bullock, 1996; Elzinga *et al.*, 2001; Hurford and Schneider, 2006; Bussotti *et al.*, 2009). The two highest scoring categories of methods, Environmental Change Network monitoring and permanent plots, both involve repeat visits to a fixed location. It seems that the physical re-visiting of exactly the same location provides assurance that a measured change represents a true change over time. It could also be that the association of statistical analysis with quantitative plotbased sampling is generally viewed as a positive thing since statistics such as the mean and variance can be calculated.

When asked to provide criteria for effective methods, however, the majority of participants focused on the practical aspects such as feasibility and the need for standard protocols. Although these are important to the repeatability of a monitoring programme (Chiarucci et al., 2001; Tucker et al., 2005; Pereira and Cooper, 2006), statistical criteria are also essential for detecting change over time and meeting objectives. Some of the practitioners did have a good understanding about these aspects, stating that effective techniques provide: (i) objective measurements which reduce observer error so that differences between observers are less than differences over time and (ii) detailed data which is sensitive to a particular size of change, as outlined in various studies on monitoring (Green, 1979; Vos et al., 2000; Legg and Nagy, 2006). These practitioners were in a minority, with only five people stating the 'ability to detect change' as an important criterion for monitoring (two of whom added '...at an appropriate magnitude'); two of these were policymakers and three were researchers, implying that it is people involved in the theoretical side of the industry who are aware of how to design statistically effective monitoring.. This supports concerns in the literature (Brown and Rowell, 1997; Legg and Nagy, 2006) that many monitoring schemes fail due to a lack of understanding of applied statistical requirements (and therefore how to plan and carry out appropriate sampling and analysis).

The main reason for the agreement that qualitative methods are less effective than quantitative sampling was their high subjectivity. Whilst the finding that condition monitoring involves excess personal judgement is not new (Jackson and Gaston, 2008), this does not necessarily affect a programme's ability to achieve its objectives. Studies have shown that rapid, low-cost condition assessments over large scales with explicit subjectivity can be very useful (Gibbons *et al.*, 2006). However, participants in this survey viewed the combination of measurements of variables with decision making and/or value judgements as less effective. Concerns about the Common Standards Monitoring (CSM) guidance ranged from aspects of its practical feasibility through to its lack of 'scientific evidence'.

The range of opinions associated with the different versions of the CSM guidance is most surprising. The Joint Nature Conservation Committee (JNCC) standard guidance is consistently given the lowest scores out of the versions and is most criticised for being 'unrepeatable and not robust'; respondents said that the guidance is 'hard to interpret', 'difficult to make site-specific', uses 'inflexible criteria' and that the amount of personal interpretation in assessment of condition means that different surveyors will come to different conclusions. Interviewees rated the Countryside Council for Wales version significantly higher than the JNCC guidance, stating that the CCW version is 'repeatable', 'easy to set up' and that it 'measures the right things'. The Scottish Natural Heritage (SNH) version was also rated higher (but not significantly so) than the JNCC version with reasons including that is it 'rapid' and 'well linked to management'. The positive rating of the CCW and SNH versions are likely to be because they have been devised to overcome the problems associated with the original JNCC guidance. The fact that the JNCC guidance is rated so low is of concern for the evaluation of change in condition over time, since this is the standard procedure advised by the UK government's advisory body on conservation. Problems could also arise when agencies use different versions, with the potential of resulting variation in results having important implications for the consistency of reporting categories across and between regions. Despite recommendations for validation of CSM (Gaston et al., 2006; Jackson and Gaston, 2008), there has been very little empirical published validation of assessments produced using either the JNCC standard guidance or the different versions, and no validation has been carried

out comparing results from different versions (Bealey and Cox, 2004; Ross *et al.*, 2004; Ross and Bealey, 2006).

This study furthermore highlights the perception of potential inconsistencies in condition assessments arising from the different sets of habitat-specific CSM guidance, with the respondents' scores varying widely, again with the reasons of repeatability and subjectivity given for high and low scores respectively. One problem noted was the variety of recommended field sampling methods. For instance, the woodland habitat guidance recommends 'a structured walk around the site with a series of observation stops along the way' (JNCC, 2004b), whereas the upland habitat guidance permits a choice of sampling strategies (JNCC, 2005a). Although the numbers of interviewees involved in scoring each set of guidance was small, this does suggest that further work is required to allow formal appraisal of the various sets of guidance.

The second monitoring method involving qualitative techniques, Phase I/National Vegetation Classification, was also criticised for being overly subjective, which again agrees with previous reports (Cherrill and McClean, 1995, 1999a; Stevens *et al.*, 2004b). These studies all found low levels of agreement in repeated Phase I surveys and concluded that the assignation of vegetation type is largely a matter of personal opinion. Since the NVC involves more detailed vegetation classifications and is mapped at a finer scale than Phase I (Rodwell, 2006), a higher agreement is expected (Kuchler, 1973), but again there has been no published investigation into its reliability. Neither Phase I nor the NVC was actually designed to be used for monitoring (Rodwell, 1997), but this study highlights that they are being so used and that they may not be fit for this purpose.

The fact that the stakeholder group had very little effect on responses suggests that opinions about different methods reported in this study represent a consensus between different groups involved in vegetation monitoring. There is a slight trend of scores decreasing from the most practical-minded, applied stakeholder group, land managers, through consultants to advisors to policymakers and finally to the least applied group, researchers. This suggests that people involved in designing, advising and researching about the methods tend to be more critical of them and aware of their flaws than people involved in their implementation in the field.

There are limits to the strength of findings based on interviews since all answers, including the rating of methods' effectiveness, are based on the varied knowledge base of the respondents (Hockings et al., 2000). Answers can be open to interpretation; when asked about criteria for effective monitoring, 18% of respondents included 'robust' and 17% 'fit for purpose'; both of these terms are ambiguous and could have several meanings. It is also known that interviewees may modify their answer either to correspond with what they think the interviewer wants to hear, to make an impression or because they are fearful of being critical (Ruxton and Colegrave, 2006). In this study, a neutral stance was maintained by the interviewer, no leading questions were asked and there was no particular reason to suspect that any participants changes their answers to suit the situation. One concern however is that some participants contacted the interviewer in order to take part since they had a particular interest in the subject and thus these interviewees may represent extreme viewpoints about vegetation monitoring methods and this set of respondents did tend to be strongly critical of Common Standards Monitoring. However, studies based on interviews have been shown to produce valid results which would have been very difficult to obtain otherwise (Slater et al., 2006; Jones et al., 2008).

2.4.1 Limitations to this study

Sampling a greater number of interviewees would have provided greater confidence in results, interviewing was stopped at sixty due to time restrictions. Since it was a scoping study, it was necessary to conduct it at the outset of the PhD research. There was a bias in geographic coverage, with Wales over-represented and England under-represented. For a UK wide study, Northern Ireland should also have been included.

2.5 Conclusions

This is only a partial exploration of practitioners' perceptions of vegetation monitoring methods, but does reveal some interesting patterns. This study has found that conservation practitioners are in agreement over which vegetation monitoring methods currently used in the UK are most and least effective at meeting objectives, with quantitative methods scored as more effective than qualitative alternatives. The reasons given for the scores demonstrated that practitioners value the practical aspects which make quantitative methods repeatable (standard protocols, training of observers etc), but lack understanding of theoretical aspects (statistical power, size of change to be detected etc). Practitioners were also in agreement that condition monitoring and vegetation mapping involve excess subjectivity which leads to high levels of intra- and inter-observer error. The responses showed that due to concerns over the variability produced using the standard JNCC guidance for condition monitoring, CCW and SNH have developed alternative versions which may produce differences in condition assessment. This study recommends that the following research is carried out: (i) empirical validation of the various condition monitoring approaches, (ii) quantification of the repeatability of habitat mapping using the National Vegetation Classification as an approach to monitoring and (iii) exploration of the ability of quantitative methods to detect specified change.

Chapter 3

The repeatability of the National Vegetation Classification approach to vegetation mapping

3.0 Abstract

Habitat and plant community mapping is a common component of planning and monitoring for conservation management. However, there are major concerns about its subjectivity and risk of observer bias. This study provides the first test of the consistency of habitat maps produced with the National Vegetation Classification (NVC), the most widely used system for production of habitat maps on conservation sites in Britain. Seven surveyors mapped the same upland site within five weeks in summer 2008 and the spatial correspondence of the resulting maps was assessed. The NVC is a hierarchical classification and pair-wise spatial agreement between maps decreased with lower levels of sub-classification. The average agreement between maps was 77.6% at the major (habitat) level, 34.2% at the community level and 18.5% at the sub-community level. Further comparison of each of the 2008 maps with an NVC map from 1999 produced similar levels of agreement; 76.8% at the major habitat level, 37.9% at the community level and 16.5% at the sub-community level.

Spatial disparity in the location of mapped boundaries between vegetation types only made a small contribution to overall differences; the majority of variation between maps was due to discrepancies in classification with vegetation types of similar species composition most often confused. It is recommended that NVC should not be used for monitoring or surveillance and where it is used for site description this should be with full acknowledgement of the inherently subjective and uncertain nature of the maps produced.

3.1 Introduction

3.1.1 Vegetation mapping and classification

Vegetation mapping (the representation of observable patterns of interest in the actual vegetation landscape onto a map) is an important element in environmental management, providing information for planning, monitoring and policy decisions (Kuchler, 1967; Rowe, 1996; Sutherland, 2000). Although vegetation maps are generally taken as representing the truth since they are easy to interpret and understand, as in all ecological assessments, the data contained in vegetation maps needs to be of sufficient quality for its purpose, ideally with the amount of error quantified (Treweek, 1996; Williams, 1996). There are two main measurements of error applicable to vegetation mapping; accuracy and repeatability (or precision or consistency), Accuracy is the proximity to the true value, i.e. the extent to which a map represents the observable patterns of interest in the actual landscape, and this relies on an exact measure of reality (Bråkenhielm and Qinghong, 1995; Lundstrom, 2000). Repeatability is how much variability different measures produce, i.e. the extent to which different maps by the same or different surveyors agree; this can be assessed by comparisons of maps. It has been noted that few vegetation maps contain information on either type of error, due to difficulties of quantification and presentation (Millington and Alexander, 2000).

Vegetation mapping is closely linked to vegetation classification, the identification of individual units of vegetation and subsequent arrangement in an orderly and meaningful way (Kuchler, 1973; Kent and Coker, 1992). The repeatability of maps depends on the classification system and the method of mapping, both of which contain elements of personal judgement and can never be absolutely objective (Mueller-Dombois and Ellenberg, 2002). The classification can either be devised on the ground using samples for a particular locality, or can be available before field work and any samples fitted into the pre-existing structure. The method of mapping encompasses the field method and the assignation of sample to vegetation type.

The vegetation classifications (and associated mapping methods) most widely used in conservation management in the UK, are the Phase I system (JNCC, 2004d) and the National Vegetation Classification (NVC) (Rodwell, 2006). The Phase I system consists of 155 habitat types and tends to be used at the broad, regional scale in order to identify sites of potential conservation value whereas the NVC, which comprises 681 units (JNCC, 2008b) is used in ecological survey and assessment on conservation sites to produce inventories and maps of plant communities (Kirby, 2003; Strachan and Jackson, 2003; Wilson, 2003).

There are a number of other vegetation classifications systems used in the UK. The biodiversity Broad Habitat Classification is the framework used for the UK biodiversity action planning process and for reporting the condition of protected sites (Jackson, 2000). The basis of the Broad Habitat Classification is that it should consist of a limited number of habitat types that are simple and easily understood by a range of people; there are 17 terrestrial/freshwater Broad Habitat Types (UKBG, 1998). The Biodiversity Priority Habitats are nested within the Broad Habitat Classification; these Priority Habitats (45 terrestrial/freshwater types) are those identified as in need of conservation concern, and for which Habitat Action Plans have or will be drawn up (Jackson, 2000). The Annex I habitat types drawn up at a European level are also used in the UK, mainly to inform conservation priorities at a UK level; of the 189 types, 78 are thought to occur in the UK (JNCC, 2009a).

A further classification system is that used in the Countryside Survey, the ITE Land Classification based on data on vegetation, soil, climate and topography from 1228 km² samples from the intersections of a 15km grid across the UK (Firbank *et al.*, 2003). The botanical data from over 13,000 vegetation plots recorded in the 1978 and 1990 Countryside Surveys was also used to develop the Countryside Vegetation System (CVS) which was intended as the basic building block for the subsequent development of botanical indicators and analysis of vegetation change (Bunce *et al.*, 1999). Differences in the rationale behind all these habitat classifications means that correspondence between systems is not straightforward. However, efforts have been made to bring systems into some alignment and the Broad Habitat Classification and Priority Habitats are now compatible with Phase I and Countryside Survey (Jackson, 2000).

Studies looking at the reliability of vegetation mapping using the Phase I survey classes found that the agreement between a pair of maps of the same area produced by two different surveyors at 12 months apart was only 44% of the study area (Cherrill and McClean, 1995) and that average agreement between maps of the same area produced by six different surveyors in the same month was only 25.6% of

the study area (Cherrill and McClean, 1999b, a). Cherrill and McClean concluded that the Phase I method is inherently subjective, with the decision of land cover class based on personal opinion.

Quality Assurance of the Countryside Survey data has also been carried out; in 2007, surveyors revisited some of the 1km² plots and repeated the vegetation sampling. They found that agreement between Quality Assurance and 'real' surveyors about the presence of a Broad Habitat in any plot was high (81%), as was agreement about change of most Broad Habitats; out of 19 habitat categories, 14 had agreement of >80% (Norton *et al.*, 2008). However, this analysed assignment to category rather than spatial agreement, and involved far fewer categories than in Phase I or NVC.

3.1.2 This study and the National Vegetation Classification

The present study aims to quantify the precision, or repeatability, between surveyors mapping vegetation based on a standard vegetation classification widely used in the UK, the National Vegetation Classification (NVC). The NVC is a comprehensive classification and description of the plant communities of Britain along phytosociological lines (Rodwell, 1991b, a, 1992, 1995, 2000; JNCC, 2008b). The development of the classification used about 35,000 samples, covering nearly all natural and semi-natural, and a number of highly modified vegetation communities, to characterise vegetation types using multivariate techniques such as those in TWINSPAN (Hill, 1979; JNCC, 2008b). However, the samples included a large component of records from protected sites and from Scotland in particular (Rodwell, 1991a; Turner, 2008). This may make the NVC system better suited to semi-natural vegetation, and particularly to vegetation types which occur in Scotland. The NVC is now the most widely accepted classification in the UK for use in ecological site survey and assessments to produce inventories and maps of plant communities on designated or threatened sites; it is used by UK conservation statutory agencies, non-governmental organisations and consultants (Kirby, 2003; Wilson, 2003). Maps of NVC classes provide information for management plans and policy actions, give a baseline against which to measure change and act as a framework for scientific research into relationships between plant communities and environmental factors which influence their distribution.

The NVC classifies UK vegetation into 12 habitats, such as M (mires) and H (heaths), further subdivided into 286 'communities', such as M17 (*Scirpus cespitosus-Eriophorum vaginatum* blanket mire) and M25 (*Molinia caerulea-Potentilla erecta* mire). Many of the NVC communities are broken down further into 'sub-communities', such as M17a and M17c (*Drosera rotundifolia-Sphagnum* spp. and *Juncus squarrosus-Rhytideadelphus loreus* sub-communities respectively). A very small number of especially species-rich and/or complex communities have a third level of sub-division, into 'variants'.

Associated floristic tables consisting of lists of the species have been drawn up for all NVC communities and sub-communities, these tables are based on measurements recorded in sample plots; frequency of each species (the proportion of sample plots the species is found in) from 'I' (1% - 20%) to 'V' (80% - 100%) and the range of abundance in each species across sample plots from 0 (absent) to 10 according to the 'DOMIN' scale (Rodwell, 2006). In general, communities have been named using two or more of the most frequent and abundant species, with many sub-communities named using distinctive 'preferential' species. Preferential species are those which occur at a frequency of III-V (41%-100%) in one sub-community and at a frequency of II (21% - 40%) or less in all other sub-communities within a community. The occurrence of two or more such preferential species is used to define a subcommunity, e.g. Vaccinium myrtillus and Juncus squarrosus are preferentials for M17c as they are usually absent in M17a or b. The concept of preferential species is sometimes used at the next level up in the classification hierarchy, the community, if a particular species occurs at high frequency (generally IV-V) only in that community and no others. Each community also has a letter which abbreviates the habitat vegetation type and a number which indicates the position in the sequence of communities described within that type.

The standard approach to vegetation mapping using the NVC requires surveyors to identify homogenous areas of vegetation, although this is not always possible in the field, and to assign a vegetation class to each one on the basis of the abundance of some of the plant species (vascular, bryophyte and macro-lichen species) of which it is composed (Elkington *et al.*, 2002). Samples are taken from each homogenous area, sorted and compared against the NVC floristic tables to determine which vegetation type gives the best fit. This process can either be done by hand, using the

keys in the British Plant Communities Volumes (Rodwell, 1991b, a, 1992, 1995, 2000), or entered into a computerised key such as MATCH (Malloch, 1990) or TABLEFIT (Hill, 1996), which work by making statistical comparisons using simple similarity co-efficients between survey samples and the NVC floristic tables. More experienced surveyors tend to bypass some of these stages and use specific guidelines and comparison tables between NVC tables or keys; these may or not be published formally (Elkington *et al.*, 2002; Hall *et al.*, 2002; Averis *et al.*, 2004). Furthermore, it may be possible to assign areas to a vegetation type without recording any data at all, due to knowledge acquired through repeated use of guidelines and tables or published keys (Rodwell, 2006).

3.1.3 Issues to be addressed in this study

Variation between surveyors according to level of NVC detail

It has been found that as the level of detail increases in vegetation surveys, so does the amount of between observer error; in a study of a Phase I mapping, agreement increased when Phase I land cover types were amalgamated into broader habitat groups (Stevens *et al.*, 2004b). It is likely that this will be the case with NVC mapping, thus there will be more agreement between surveyors when the sub-community types are amalgamated into community types, and when the community types are amalgamated into habitats. This is presumably because there are fewer choices available at the broader levels, and so it is easier to choose the 'correct' vegetation type, which reflects the general recommendation that in order to be practical, vegetation units in a classification should be few (Kuchler, 1973). However, there is a balance to be met; as the number of classification units available decreases, so the detail of the vegetation landscape is obscured (Grubb *et al.*, 1963). In their study, Cherrill and McClean (1999b) noted that total agreement between maps is only likely to be achieved when vegetation types have been combined to such an extent that all meaningful detail is lost.

Variation between surveyors in homogenous versus heterogeneous areas

The subjective choice of location in NVC mapping presupposes that the surveyor has a clear impression of the vegetation types that are present in the area to be mapped, that the patch of a given vegetation type within which each sample is located is sizeable and homogenous, and that the surveyor successfully locates the sample so that it is representative of the patch, e.g. through being uniform and homogenous (Kent and Coker, 1992). Where a vegetation patch is homogenous it will be relatively uniform in colour and texture (Rodwell, 2006) and surveyors should be able to recognise and avoid boundaries (at the edge of the patch) between different vegetation types, resulting in consistent assessments. However, in heterogeneous areas, vegetation will be more spatially variable, with patches of different vegetation types grading into each other in a more complex mosaic. In this case, even for an experienced surveyor, it may be difficult to decide where vegetation of one type end and another begins; this is likely to produce low levels of agreement.)

Variation between surveyors according to NVC type

Stevens et al (2004) found that agreement between surveyors was highest for Phase I land cover types which were distinctive, particularly highly modified cover types. This observation is echoed by Cherrill and McClean (1999a, b) who found that agreement between Phase I surveyors was greatest in areas of improved grassland and woodland. All of these cover types have features which are easy to recognise, and it is likely that NVC types with such clear distinguishing features will also show high levels of agreement. Linked to this is the question of which types of vegetation are most easily confused with each other; Cherrill and McClean (1999a, b) found that surveyors tended to confuse land cover types that were most similar in appearance and species composition, such as semi-improved neutral grassland.

One aspect of distinctiveness between vegetation types is the type of species characterising each; studies of surveyor variation in vegetation surveys have generally found that there is less surveyor error associated with large or distinctive species than with small species or those which have characteristics in common with many other species (Sykes *et al.*, 1983; Kennedy and Addison, 1987; Bråkenhielm and Qinghong, 1995; Prosser and Wallace, 1999; Klimeš *et al.*, 2001; Scott and Hallam, 2003). Thus NVC vegetation types characterised by large, distinctive species such as *Pteridium aquilinum* may have more agreement and be less easily confused

with each other than vegetation types characterised by smaller species which are similar in appearance to other species, such as *Tricophorum cespitosus*.

Reasons for variation amongst surveyors

a) Surveyor experience, effort and method of assignment to NVC type

There are several factors which influence the amount of variation between surveyors in vegetation surveys, including surveyor expertise and experience (Lundstrom, 2000; Scott and Smart, 2006). The more experienced a surveyor, the more they are able to charge, thus the cost of a survey may also be a corollary of its probability of agreement with other surveys. The survey effort is also likely to influence the quality of the survey; time is one measure of this (Nilsson and Nilsson, 1985; Kercher *et al.*, 2003), another is the total length of the route taken. Weather conditions are also well known to influence the quality of surveys; in cold, wet conditions not only is visibility reduced but concentration is adversely effected (Bonham, 1989). Finally, the length of time between surveys may have an effect; if there is a large gap, variation in surveys could be due to real changes in the vegetation.

The way surveyors assign patches to NVC type may also influence map agreement. As explained previously (section 1.2), there are a number of ways of assigning types, ranging from experience through to using the keys and tables in the British Plant Communities volume, to running plot data through specialised computer programs. A study of MATCH and TABLEFIT found that expert opinion and computer program sub-community agreement for given stands of vegetation was only 43% for TABLEFIT and 36% for MATCH (Palmer, 1992). Since the experts were specialist botanists their opinions were used as the standard and it was concluded that answers provided by these programs are indications rather than fact, and that the descriptive text of the British Plant Communities and other published guides should always be consulted.

b) Spatial errors

Some disagreements between maps will be due to differing perceptions and depictions of the placement of boundaries. Boundaries between vegetation patches
can be sharp, such as a land use boundary between a wheat field and a plantation, or transitional, where in the extreme there are no boundaries which can be shown logically on a map (Kuchler, 1973). In the second case, the surveyor has to make a seemingly arbitrary decision about where to place a line along a transition between two different vegetation patches, and the location of one surveyor's boundary is likely to differ from the next. Boundary errors also arise from the difficulty of depicting the complex nature of spatial transitions in vegetation onto a paper map (Rodwell, 2006) and also from the interpretation of maps during the process of digitising (Cherrill and McClean, 1999b). Errors made during the drawing and digitise in the field using a hand-held computer. This approach was used in the vegetation mapping aspect of the Countryside Survey in 2007; surveyors were also given the previous survey and asked to mark genuine change and change due to previous misallocation of boundaries or habitat type (Maskell *et al.*, 2008).

c) Resolution of mapping

Heterogeneity of vegetation is scale dependant; vegetation mapped at a very fine spatial scale will appear more heterogeneous than the same vegetation mapped at a broad scale. Rodwell (2006) recommends that the scale of a map of NVC types should determine the size of the minimum mappable unit (Rodwell, 2006). All surveyors in the present study were given maps of the same scale, so this source of error should be reduced. However, surveyors may still map at different scales onto the same map; where one surveyor may map all small patches of different vegetation located within a larger patch, another may simply map the general vegetation type of the patch.

d) Discrepancies between NVC types

Although the NVC is a rigorous classification with clearly defined units (Kuchler, 1973), the recognition of its vegetation types on the ground is intrinsically subjective (see above). This means that confusion over what NVC type to assign to a given patch of vegetation is likely to be the main source of disagreement.

3.1.4 Objectives

This study aims to answer the following questions:

1. Does between-surveyor variation increase with the level of detail in the hierarchical NVC that patches of vegetation are classified to?

2. Is between-surveyor variation greater in areas of heterogeneous vegetation than in areas of homogenous vegetation?

3. Does between-surveyor variation vary with NVC type? It is likely that NVC types with obvious features (e.g. distinctive dominant species) will show less variation than NVC types without such features and that there will be more confusion between NVC community types which are close to each other in species composition and which have a high proportion of dominant species in common.

4. Is variation amongst maps explained by:

a) Surveyor experience, effort and method of assigning vegetation to NVC type

b) Spatial errors in placements of boundaries

c) Differences in resolution of mapping

d) Discrepancies in assignment to NVC type?

5. Is between-surveyor variation from NVC maps produced at the same time less than the changes picked up between NVC maps produced at different time periods?

The study describes results from a field trial of the NVC approach applied by seven experienced surveyors to a 43 ha upland area in North Wales over a period of five weeks in June-July 2008. Spatial comparison methodology developed by Cherrill and McClean (1999a, 1999b) was used throughout to allow direct comparisons to be drawn with the previous work on Phase I maps. As NVC communities are more detailed than Phase I land cover classes, a lower agreement between NVC-based maps was expected (JNCC, 2004d; Rodwell, 2006). Finally agreement was calculated between an NVC map of the study area produced in 1999 and each of the seven 2008 maps.

3.2 Methods

3.2.1 Selection of surveyors for the study

Surveyors for the 2008 survey were chosen according to their current status, expertise in vegetation survey, experience in carrying out NVC surveys and familiarity with upland habitats (Table 3.1). The map available from 1999 had been produced by surveyor A as part of the statutory habitat mapping of the Eryri/Snowdonia Site of Special Scientific Interest.

			the second se	and the second se	
Surveyor	Current status	NVC experience (no. yrs)	Surveying in upland habitats experience (no. yrs)	Most familiar with upland habitats?	Cost of survey including travel (£)
А	Ecologist for a statutory organisation	14	14	yes	160
В	Previously an ecologist for a statutory organisation, now a self-employed consultant	10	30	yes	410
С	Ecologist for a consultancy company	4	1	no	270
D	Previously an ecologist for a statutory organisation, now a self-employed consultant	18	17	yes	410
Е	Self-employed ecological consultant	2	4	yes	190
F	Previously an ecologist for a statutory organisation, now a self-employed consultant	18	1	no	180
G	Ecologist for a consultancy company	20	14	no	350

Table 3.1 Status, experience with the NVC and upland habitat surveying, and cost charged per survey for seven surveyors involved in a field trial carried out in Snowdonia in 2008.

3.2.2 Organisation of the field trial

The area selected for the field trial was approximately 43 ha of the Hafod y Llan estate (UK grid reference SH6252, altitude range 250-350 m) which is owned and managed by the National Trust and located in the Snowdonia National Park (SNP) in North Wales. The estate is of importance for conservation and forms part of the Eryri Special Area of Conservation (SAC), Eryri Site of Special Scientific Interest (SSSI) and Yr Wyddfa/Snowdon National Nature Reserve (NNR) (Fig.3.1). The vegetation is a mosaic of dry heath, wet heath, blanket bog and acid grassland (with abundant

Pteridium aquilinum and Molinia caerulea). Two contrasting parts of the estate were selected. Part A (8.7 ha) in the north of Hafad y Llan, is a steep east-facing slope and was chosen to represent homogeneous vegetation since it is comprised of patches of *Pteridium aquilinum* at the bottom and heathland at the top. Part B (34.4 ha) in the south of Hafod y Llan is a hanging valley and was chosen to represent heterogeneous vegetation, comprising a mosaic of wet heath, blanket bog, and acid grassland all dominated by *Molinia caerulea*. The total study area is a suitable size for surveyors to map in one day and linear features were used to reduce ambiguity about the location of outer boundaries though the interior area has few artificial boundaries (e.g. fencelines).



Figure 3.1 Study area used for field trial of NVC mapping. The area is part of The National Trust's Hafod y Llan estate in Snowdonia National Park in North Wales (SH6252). Crown copyright/database right 2008. An Ordnance Survey/EDINA supplied service.

The seven surveyors were provided with digitised and orthorectified aerial photographs which were chosen for ease of use in the field; surveyors can recognise patches of vegetation and use them as landmarks to enable more accurate boundary location (Turner, 2008 *pers. comm.*). In a study looking at repeat Phase I surveys, it was found that the use of aerial photographs taken as close as possible to the date of survey produced the most accurate vegetation mapping (Dargie, 1993). The present field trial conducted in 2008 used the most recent aerial photographs available, taken in summer 2006; since this was only two years previously and

changes in upland areas are relatively slow (Marriott *et al.*, 2004), the vegetation in the photographs was expected to bear a close resemblance to the vegetation encountered in the field. Furthermore, the time of year that aerial photographs are taken makes a big difference to how much differentiation there is between patches of different vegetation type; photos are often taken in summer to increase this differentiation. The aerial photographs were 1:5000 and were overlain with opaque acetate film which could be written on. Assuming the surveyors used a fine pen or pencil, an area of 1 x 1 mm could be demarked on the acetate, translating to a patch of 5 x 5 m. The aerial photographs allowed patches of certain vegetation types to be delineated, notably *P. aquilinum* stands which show up as a dark green and dry heath which is a dark grey green (Appendix 3.1).

The surveyors were also provided with ordnance survey (OS) maps of 1:10000 to allow cross-referencing to contours and other features (Appendix 3.1). The surveyors were asked to use the standard NVC mapping approach (Rodwell, 2006). Field surveys were conducted over five weeks between 8th June and 15th July 2008. The surveyors spent one day at the study site and the surveys were arranged such that different surveyors would not be present on the site at the same time. Weather conditions, time spent on site, route taken and method of assignation to NVC type were noted for all surveyors (Table 3.2).

Table 3.2 Date of survey, weather and visibility conditions during survey, time taken, length of route
and way in which patches were assigned to NVC type for seven surveyors involved in a field trial of
NVC mapping in Snowdonia in 2008. BPC= British Plant Communities books.

Surveyor	Date of survey	Weather	Visibility	Time taken (hours)	Route length (km)	How surveyor assigned patches to NVC type
A	7 [™] June	Sun	Excellent	10	9.9	Experience, tables and alternative keys based on BPC ¹
В	9 th June	Sun	Excellent	10	4.9	BPC
С	11 th June	Cloud/wind	Good	8	3.8	Alternative keys and TABLEFIT ²
D	17 th June	Cloud/sun	Good	4.5	4.6	Experience and BPC
E	16 th June	Cloud/sun	Good	8.5	4.1	BPC
F	13 th July	Cloud/sun	Good	10	4.2	BPC
G	15 th July	Cloud/sun	Good	10	7.1	Experience and BPC

¹BPC= British Plant Communities (Rodwell, 1991b, a, 1992, 1995, 2000)

² TABLEFIT= computerised key (Hill, 1996)

The survey in 1999 was conducted in August/September, again using an aerial photograph from Spring 1992 over laid with acetate, however the aerial photographs were not digitised or orthorecitified, and were blown up to a scale of approximately 1:4000.

3.2.3 Analysis of variation amongst maps

All 2008 maps were scanned and geo-referenced in MapInfo (MapInfo, 2004) to within 2 pixels of error and digitised on screen; the 1999 map had already been digitised by the surveyor. Lists of NVC types identified and their representation (as a proportion of the study site) on each map were produced. Where surveyors listed more than one NVC vegetation type for polygons.it was decided to use the dominant NVC type for analysis. All polygons in each surveyor's map (except for the 1999 map) were buffered by adding 5 m either side of each boundary and the area lying within this 10 m wide boundary strip was excluded from certain analyses (as specified below). This was done in order to provide allowance for small inaccuracies in surveyors' placements of boundaries, and concentrate on agreement in cores areas (Cherrill and McClean, 1999a). The choice of 5 m was made to correspond with the smallest mappable unit (Rodwell, 2006). This resulted in two maps for each surveyor, one original and one buffered.

Area of agreement

Files were converted to ESRI shape files for use in ArcView (ESRI, 2002). Each pair of maps, e.g. A and B, were overlaid to create a third map containing all polygons in map A split according to the arrangement of polygons in map B. Each of the maps were joined (and polygons split) successively to produce a single layer showing all polygons from the seven maps. Polygons of less than 5 m² were deleted to reduce errors introduced from digitising and for ease of further analysis; these polygons amounted to 0.33 ha (0.77%) of the study area. For all analyses at the sub-community level only the four maps from 2008 with >90% of the area classified to sub-community were used. The area of agreement between each pair of maps was calculated using a matrix of correspondence between community types, (or habitat

type or sub-community); for example in Table 3.3, the (Ai, Bi)th cell shows that there is 1.1 ha, of community type i in Map A classed as community type i in Map B.

		3		MA	AP A		
MAP B	NVC community	Ĩ	ii	iii	iv	V	Total
	i	1.1	13.7	0.2	5.3	0.1	20.4
	li	0	0.8	0	0	0	0.8
	iii	0.2	1.5	0.2	1.5	1	4.4
	iv	0	0.9	0	1.6	0.1	2.6
	V	0.2	0.9	0	0.1	0.8	2
	Total	1.5	17.8	0.5	8.5	2	30.2

Table 3.3 Example of a matrix of correspondence in area (ha) between two maps, A and B, with five NVC community types i, ii, iii, iv and v.

The diagonal in the matrix (where each map gave the same type i.e. agreed) were totalled for all community types for each unique pair of maps, and converted to percentage of the study area. These percentages were then averaged for all the pairs of maps and used as an index of extent of overall agreement between maps. There were 21 unique pair-wise map combinations for the habitat and community types, and six for the sub-community. These data were found to be normal using the Shapiro-Wilk test in SPSS (SPSS, 2003). The mean area of agreement between pairs of maps was calculated for the original and buffered maps and the effect of buffering was tested using paired samples t-tests in SPSS (SPSS, 2003). This analysis was also carried out between the 1999 map and each of the seven 2008 maps in turn (excluding the analysis of buffered maps). In order to check whether the 2008 maps were consistently different from the 1999 map, Pearson's correlation was calculated for percentages of NVC communities in the 1999 map and each 2008 map and in pairs of 2008 maps. Percentage figures were converted to proportion and an arcsine transformation carried out before Pearson's correlation was calculated in SPSS (SPSS 2003).

Location of agreement The following analyses are carried out on the 2008 maps only. The second objective is to ascertain which locations in the trial area had least and most agreement in order to test whether surveyors tend to agree less in areas of heterogeneous than in homogenous vegetation. In MapInfo, a raster grid of 10 m spacing was imposed over the study area, and the NVC type at the centre of each grid square in each of the seven maps from 2008 was recorded. The amount of agreement in NVC types was calculated for each grid cell, resulting in cells labelled 1 (all maps different) to 7 (all maps the same). These data are displayed in Fig. 3.3 and were used to calculate the extent of agreement in NVC type as a proportion of the study area and in Parts A and B. In order to compare agreement in Part A with Part B, an index of agreement was calculated for Part A and B at the habitat, community and sub-community levels. The figures for percentages of Part A/B agreed on by 0-7 maps were each multiplied by the corresponding number of maps agreed on i.e. between 0 and 7 and then divided by the combined number of maps in agreement (2+3+4+5+6+7=27). This resulted in an index between 0 (no agreement) and 100 (total agreement).

Variation amongst surveyors according to NVC type

The third objective is to identify the frequency of agreement associated with different NVC types and to determine which types were most often confused with each other. This was only carried out for NVC communities. Firstly, the amount of 10 m x 10 m grid squares labelled as a particular NVC type in one map up to all seven maps was calculated as a proportion of the total number of grid squares labelled as that NVC type across all maps (section 3.2.4.2). Only NVC types with a cover of 5% in at least one map were included; this cut off was chosen to exclude NVC types with a marginal contribution across all maps.

Variation amongst surveyors was analysed only at the NVC community level to avoid repetition. The number of 10 m grid cells labelled as a particular community type in one map up to all seven maps was calculated as a proportion of the number of grid cells labelled as that community summed across all maps. Only communities with a cover of 5% in at least one map were included in order to exclude communities with a marginal contribution. Across all pairs of maps the proportion of community *i* in the first map labelled as community *ii* in the second map was determined and repeated reversing the order of the maps i.e. M compared to N followed by N compared to M. This resulted in 42 pair-wise comparisons. A confusion matrix showing the mean

proportion of confusion between each pair of communities was then formed and compared with a matrix of coefficients of similarity between the same pairs of communities. The measure of similarity used was based on the national species lists of these community types (Rodwell, 1991b, 1992) analysed in MATCH (Malloch, 1990), by creating exemplar constancy tables for each community based on the constancy of their species composition (Packwood, 1991). Finally a similarity matrix was formed using the proportion of constant species in the species list for each community present in the list for communities with which it has been confused. Constant species were classed as species with frequencies of between 61% and 100% (IV and V) according to the national floristic tables (Rodwell, 1991b, 1992). Spearman's rank correlation co-efficient was calculated for the confusion data and similarity coefficients, and for the confusion data and proportion of constant species.

Reasons for variation amongst surveyors

The final objective was to find reasons for the variation amongst the surveyors' maps, both through analysis and visual inspection of the maps. In addition, in order to explore the influence of surveyor experience, the number of years of experience of NVC mapping and of upland habitat survey, and the cost charged were correlated in SPSS (using Spearman's rank correlation coefficient) against the following variables used to characterise each survey: amount of agreement per surveyor (mean of the pair-wise combinations involving that surveyor, section 3.2.4.1), number of habitat, community and sub-community NVC types identified, number of polygons mapped and proportion of the study area mapped to the sub-community level. Whether upland habitats were the habitat most familiar to the surveyor was also compared with amount of agreement and number of NVC types.

Two measures of surveyor effort were also correlated against the above variables: time taken to do the survey and total length walked during the survey. The number of days between pairs of surveys was also correlated with area of agreement per pair. Weather conditions were similar for all surveys so this variable was excluded from analysis. Finally, the methods used to assign areas to vegetation type were compared with amount of agreement.

3.3 Results

3.3.1 Overall differences amongst 2008 maps

Five surveyors mapped onto the aerial photograph (AP) and two mapped onto the Ordnance Survey (OS) map. The two maps produced on the OS maps had the fewest polygons which were the largest mean size, and also contained the largest minimum and maximum sized polygons out of all of the maps (Table 3.4, Fig. 3.2, Appendix 3.2). The largest number of polygons in any one map is 445 (map A) and the smallest number is 24 (map F), (Fig. 3.2).

Table 3.4 Information about the media that seven surveyors drew their maps onto (AP= aerial photograph, OS= Ordnance Survey map), number and size of polygons and proportion mapped to sub-community in their maps. Data were collected in a field trial carried out in Snowdonia, 2008.

	map						
	А	В	С	D	E	F	G
media mapped on	AP	AP	AP	AP	OS	OS	AP
number of polygons	445	41	137	64	32	24	65
mean size of polygon (m ²)	977	10561	3152	6770	13562	18138	6671
smallest polygon (m ²)	1	29	61	59	499	982	379
largest polygon (m²) Proportion of study area	45515	238715	62764	78479	272528	243488	3244444
recorded to sub-community	91	26	17	99	99	38	97



Figure 3.2 Maps A-G of National Vegetation Classification community. Data collected in a field trial carried out in Snowdonia, 2008. See appendix 3.2 for larger scale versions.

3.3.2 Detail of variation amongst 2008 maps

Overall, three habitat types; dry heaths (H), mires and wet heaths (M) and calcifugous grasslands and montane communities (U) were common to all seven maps, with calcicolous grasslands (CG) also recorded on two maps (Table 3.5a). The surveyors agreed about the relative extent of these habitat types, with M as dominant, followed by H then U. At the community level, 20 communities were recorded in total, with between 5 and 18 recorded in individual maps (mean 10). Only four communities were recorded in all maps, and a further two were recorded in six maps (Table 3.5b). Three of these common six communities were acid grassland: U4, U5 and U20, patches of which could be distinguished on the aerial photos provided. The maps agreed fairly well about the extent of the most common communities, with M15 (wet heath), M17 (blanket mire), M25 (mire) and U20 (acid grassland) in the top three dominant communities in at least three maps.

In addition, all surveyors mapped at least some of the area to sub-community, with four surveyors recording >90% of the area as sub-communities and the other three recording between 17% and 38% to that level (Table 3.3). Altogether, 34 sub-community types were recorded, with between 3 and 32 recorded in individual maps (mean 11) and there was little correspondence about the relative extent of sub-communities (Table 3.5c).

Table 3.5 The areas of National Vegetation Classification (a) habitats (and non-NVC categories), (b) communities and (c) sub-communities in each surveyor's map, as a percentage of total study area. Number of vegetation types recorded in each map are also shown. Data were collected in a field trial carried out in Snowdonia, 2008. * indicates area < 0.05%.

1	-	1	
	ы		
٠.	~	1	

	map						
NVC habitat (code: description)	А	В	С	D	Е	F	G
CG: calcareous grassland	*	0	0	*	0	0	0
H: heath	5.8	6.6	10	7	8.3	13	62
M: mire	66	74	51	78	65	60	69
U: acid grassland	24	19	39	15	27	27	25
Number of habitats recorded	4	3	3	4	3	3	3

Non-INVO categories recorded							
Boulders/Rock/Scree	4.1	*	0.4	*	*	*	0.3
	A STATE OF A STATE OF A STATE OF A STATE		97026756				0.0

(b)

.

	Map						
NVC community (code: description)	A	В	С	D	Е	F	G
CG10: Festuca ovina-Agrostis capillaris-Thymus			10.000				
praecox grassland	*	0	0	*	0	0	0
H8: Calluna vulgaris-Ulex gallii heath	0.2	1.3	0	1.1	0	0	0
H10: Calluna vulgaris-Erica cinerea heath	5.5	0	10	5.9	8.3	13	5.1
H12: Calluna vulgaris-Vaccinium myrtillus heath	0.1	5.3	0	0	0	0	1.1
M2: Sphagnum cuspidatum/recurvum bog pool community	*	0	0	0	0	0	0
M6: Carex echinata-Sphagnum		10721	1920	10 00	-		U
recurcum/auriculatum mire	1.7	0	0	0.1	0	0	0.6
M15: Scirpus-cespitosus-Erica tetralix wet heath	11	61	13	56	1.7	44	51
M16: Erica tetralix-Sphagnum compactum wet heath	7.5	0	0	0	0	0	0
M17: Scirpus cespitosus-Eriophorum vaginatum				174	170		×
blanket mire	17	12	16	20	0	0	15
M20: Eriophorum vaginatum blanket and raised mire	0.5	0	0	0	0	0	0
M21: Narthecium ossifragum-Sphagnum papillosum vallev mire	0	0	04	0	0	0	0
M23: Juncus effusus/acutiflorus-Galium palustre	U	U	0.4	U	U	U	0
rush-pasture	0	0	01	0	0	0	0
M25: Molinia caerulea-Potentilla erecta mire	29	12	21	2	63	56	10
M32: Philontis fontana-Saxifraga stellaris spring	01	0	0	01	0	0	0
U1: Festuca ovina-Agrostis capillaris-Rumex	0.1	U	U	0.1	U	0	U
acetosella grassland	*	0	04	0	0	0	0
U4: Festuca ovina-Agrostis capillaris-Galium savatile		0	0.1	Ŭ	U	U	0
grassland	5.1	0.8	93	19	28	21	71
U5: Nardus stricta-Galium saxatile grassland	1.7	3.6	67	1.7	0	48	6
U6: Juncus squarrosus-Festuca ovina grassland	0.4	0	0	0	Ô	0	0
U20: Pteridium aquilinum-Galium saxatile		U	U	U	U	U	0
community	16	14	22	12	25	1.5	12
U21: Cryptogramma crispa-Deschampsia flexuosa							
community	0.1	0.6	0	0.3	0	0	0
Number of communities recorded	18	9	10	12	5	6	9

1		`
(0	• 1
U	L	. 1
Υ.	_	

NVC sub-community (code: description)	Мар						
	А	В	С	D	Е	F	G
H8a: Species-poor	0.1	1.3	0	1.1	0	0	0
H8b: Danthonia decumbens	*	0	0	0.1	0	0	0
H10a: Typical	4.4	0	0	5.9	8.3	0	3.3
H10c: Festuca ovina-Anthoxanthum odoratum	1.1	0	0	0	0	13	0
H12a: Calluna vulgaris	0.1	5.3	0	0	0	0	1.1
M6a: Carex echinata	0.6	0	0	0.1	0	0	0
M6b: Carex nigra- Nardus stricta	0.3	0	0	0	0	0	0
M6c: Juncus effusus	0.3	0	0	0	0	0	0
M6d: Juncus acutiflorus	0.5	0	0	0	0	0	0
M15a: Carex panicea	0.6	0	0	0	0.5	0	1.9
M15b: Typical	3.5	0	0	56	1.2	0	3.1
M15c: Cladonia spp.	0.5	0	0	0	0	0	46
M15d Vaccinium myrtillus	1.2	5.1	0	0	0	0	0
M16a: Typical	7	0	0	0	0	0	0
M16d: Juncus squarrosus-Dicranium scoparium	0.4	0	0	0	0	0	0
M17a: Drosera rotundifolia-Sphagnum spp.	15	0	0	19	0	0	15
M17c: Juncus squarrosus-Rhytideadelphus loreus	2	0	0	0	0	0	0
M20a: Species-poor	0.2	0	0	0	0	0	0
M20b: Calluna vulgaris-Cladonia spp.	0.4	0	0	0	0	0	0
M25a: Erica tetralix	4.5	0	16	0	63	0	0
M25b: Anthoxanthum odoratum	24	0	0.5	2	0	0	1.8
M32a: Sphagnum auriculatum	0.1	0	0	0.1	0	0	0
U1c: Erodium curcutarium-Teesdalia nudicaulis	0	0	0.4	0	0	0	0
U1e: Galium saxatile-Potentilla erecta	*	0	0	0	0	0	0
U4a: Typical	4.9	0	0	1.9	2.8	21	6.8
U4e: Vaccinium myrtillus-Deschampsia flexuosa	0.1	0.5	0	0	0	0	0
U5a: Species-poor	0.6	0	0	0.6	0	0	4.2
U5d: Calluna vulgaris-Danthonia decumbens	1.1	0	0	1.1	0	4.8	1.7
U5e: Racomitrium lanuginosum	0	0	0.1	0	0	0	0.1
U6c: Vaccinium myrtillus	0.1	0	0	0	0	0	0
U6d: Agrostis capillaris-Luzula multiflora	0.3	0	0	0	0	0	0
U20a: Anthoxanthum odoratum	16	0	0	12	23	0	6.9
U20b: Vaccinium myrtillus-Dicranium scoparium	0.9	0	0	0	0	0	0.8
U20c: Species-poor	0.1	14	0	0	0	0	3.9
Number of sub-communities recorded	32	5	4	13	6	3	14
Proportion of study area recorded to sub-community	91	26	17	99	99	38	97

Area of agreement between 2008 maps

On average only 34.2% of each pair of maps agreed on the community type, with a large range in values between the maximum pair-wise agreement of 69.6% and the minimum of 5.4% (Table 3.6). This compares with even lower agreement at the sub-community level of 18.5% (range 29.0%-8.6%) and much higher agreement at the habitat level with 77.6% (range 88.6%-66.6%). Introducing buffers increased mean paired agreement only slightly, showing that differences in placement of boundaries

explained only a small proportion of variation between maps. After buffering, mean agreement increased by 11.2% at the habitat level (t = -17.82, p < 0.01), 4.7% at the community level (t = -2.11, p < 0.05) and 6.9% at the sub-community level (t = -2.07, p = 0.09) (Table 3.6).

Table 3.6 Spatial concurrence between surveys assessed using percentage agreement between each pair of maps. Concurrence was assessed using original and buffered maps. Sub-community figures only include the four maps with >90% of the study area mapped to sub-community. Data were collected in a field trial carried out in Snowdonia in 2008.

	NVC type					
	ha	bitat	com	munity	sub-co	mmunity
pair of maps	original	buffered	original	buffered	original	buffered
A-B	86.2	98.4	21.0	44.7		
A-C	75.2	88.7	37.7	47.9		
A-D	85.7	97.7	38.6	52.0	29.0	42.6
A-E	81.0	96.1	41.2	45.8	18.7	20.6
A-F	72.5	83.3	27.5	27.2		
A-G	83.7	91.5	34.9	46.9	15.5	31.5
B-C	71.8	87.5	36.0	49.1		
B-D	88.6	98.3	69.6	78.9		
B-E	82.0	96.3	13.6	14.5		
B-F	71.6	83.3	5.4	2.6		
B-G	83.4	91.5	62.5	76.1		
C-D	71.1	87.4	40.4	17.7		
C-E	74.6	86.0	38.9	42.1		
C-F	66.6	75.9	27.3	29.4		
C-G	72.2	82.4	36.2	44.6		
D-E	83.0	96.2	17.8	17.4	19.4	15.6
D-F	71.4	83.5	10.7	6.2		
D-G	85.7	91.6	66.8	79.2	19.7	32.5
E-F	73.9	83.4	59.1	72.9		
E-G	80.0	88.9	17.2	16.0	8.6	9.5
F-G	70.4	77.5	15.4	5.1		
Mean	77.6	88.8	34.2	38.9	18.5	25.4
SD	6.6	6.8	18.5	24.6	6.6	12.3
SE	1.4	1.5	4.0	5.4	2.7	5.0

Location of agreement between 2008 maps

The level of detail of mapping also had a large impact on the pattern of agreement between the seven maps (Fig. 3.3).



Figure 3.3 The number of surveyors' maps which agree for each 10 m x 10 m grid square in terms of classification to (a) habitat, (b) community and (c) sub-community National Vegetation Classification type. Data were collected in a field trial carried out in Snowdonia, 2008.





All seven maps agreed on 40.2% of the study area at the habitat level but only 0.6% of the area at the community level (Table 3.7a). At the sub-community level, the four maps suitable for analysis at this level were in agreement over 3.8% of the area (Table 3.7a). Of the four maps included in the sub-community analysis, they were all in agreement over just 3.8% of the area. Comparison of agreement in Part A (homogenous patches of bracken and dry heath) with agreement in Part B (Bylchau Terfyn (a heterogeneous mosaic of wet heath, blanket mire and *M. caerulea* dominated mire) shows that neither Part A or B has consistently higher agreement (Table 3.7b). Although the index of agreement is higher in Part A for sub-communities, it is lower in Part A for communities and similar at the habitat level.

Table 3.7 The percentage of study area agreed on in (a) whole study area and (b) Part A and Part B of study area in none of the maps up to all seven maps in terms of National Vegetation Classification habitats, communities and sub-communities (note that sub-community figures only include the four maps with >90% of the study area mapped to sub-community). Data collected in a field trial carried out in Snowdonia in 2008.

	Percentage of area agreed on						
Number of maps in agreement	habitat	community	sub-community				
0	0	0.1	53.8				
2	0.2	6.9	27.3				
3	2.4	40.6	15.1				
4	11	31.5	3.8				
5	15.1	14.2	0				
6	31.1	6.1	0				
7	40.2	0.6	0				

(b)

(a)

	Percentage of area agreed on								
Number of maps in	habitat type		comn	nunity type	sub-community typ				
agreement	Part A	Part B	Part A	Part B	Part A	Part B			
0	0	0	0	0.1	17.4	63.0			
2	0.8	0	8.5	6.5	45.8	22.6			
3	6.4	1.3	29.4	43.4	27.4	12.0			
4	13.3	10.5	26.5	32.8	9.3	2.4			
5	16.3	14.8	21.2	12.4	0	0			
6	26.2	32.3	11.9	4.7	0	0			
7	37	41.1	2.4	0.2	0	0			
Index of agreement* (0-100)	21.2	22.4	13.2	54.9	56.9	21.8			

*The index of agreement is out of 100, where 0 = no agreement and 100 = total agreement.

Frequency of agreement and amount of confusion between NVC types in 2008 maps

Of the nine communities which comprised more than 5% in at least one map, M15 (*Scirpus-cespitosus-Erica tetralix* wet heath) and M25 (*Molinia caerulea-Potentilla erecta* mire), which made the largest contribution to area in most maps, had low levels of agreement, with less than 20% of their total area (across all maps) agreed on in more than three maps (Figure 3.4b). U20 (*Pteridium aquilinum-Galium saxatile* grassland), H10 (*Calluna vulgaris-Erica cinerea* heath) and M17 (*Scirpus cespitosus-Eriophorum vaginatum* blanket mire) had the highest proportions of agreement, with five maps agreeing on at least 10% of their total area. These three communities all have constant species which are readily distinguishable during the June-July study period (H10 with *C. vulgaris* and *E. cinerea* in flower, M17 with *Eriophorum vaginatum* in flower and U20 *Pteridium aquilinum* at its maximum height).

There were much higher levels of agreement in the habitat types, with all seven maps agreeing on 36.7% of the total area mapped as M (mires) and at least five surveyors agreeing on a quarter of the total area of H (heath) and U (acid grassland) (Fig. 3.4a). However, proportion of agreement at the sub-community level was extremely low; no more than four surveyors agreed on the area of any sub-community (Fig. 3.4c). H10a, U20a and M17a had the highest proportion of agreement, with over 10% of their area agreed on by more than three maps (Fig. 3.4c).



Figure 3.4 The proportion of area of National Vegetation Classification (a) habitats (M= mires and wet heaths, U= calcifugous grasslands and montane communities, H= dry heaths and CG= calcicolous grasslands), (b) communities and (c) sub-communities agreed on in 7 maps (all seven maps in agreement) to 1 map (no maps in agreement). Bars are ranked according to NVC types with highest amount of agreement at the left and least at the right. Data were collected in a field trial carried out in Snowdonia in 2008.

Heath, mire and grassland NVC vegetation types were all confused with each other (Table 3.8a). The NVC communities which were most frequently confused with each other were those within the same habitat vegetation type, such as the mire communities M25 and M15 and U4 and U5, both acid grassland communities.

Table 3.8 Matrices of confusion and similarity between pairs of National Vegetation Classification communities. (a) Amount of confusion amongst surveyors based on the mean proportion of total study area in agreement in 21 pair-wise combinations of seven NVC maps produced during a field trial in Snowdonia in 2008 (0 = no confusion, 100 = total confusion). Bold indicates pairs with levels of confusion \geq 30%. Figures in italics in the central diagonal indicate agreement between NVC types, with 0 = no agreement and 100 = complete agreement). (b) Amount of similarity based on co-efficients derived from MATCH (Malloch 1990) according to species composition and frequency (Rodwell, 1991b, 1992), (0 = no similarity, 100 = total similarity). Bold indicates pairs with levels of similarity \geq 60% constant species in common. M16 was only present in one map.

NVC community	36	H10	H12	M15	M16	M17	M25	U4	U5
Calluna vulgaris-Erica cinerea heath	H10	51							
Calluna vulgaris-Vaccinium myrtillus heath	H12	35	3						
Scirpus-cespitosus-Erica tetralix wet heath Erica tetralix-Sphagnum compactum wet	M15	20	1	39					
heath Scirpus cespitosus-Eriophorum vaginatum	M16	0	0	8	0				
blanket mire	M17	0	0	15	1	29			
Molinia caerulea-Potentilla erecta mire	M25	1	0	57	8	42	37		
Festuca ovina-Agrostis capillaris-Galium saxatile grassland	U4	8	0	27	1	3	23	11	
Nardus stricta-Galium saxatile grassland Pteridium aquilinum-Galium saxatile	U5	10	1	18	1	3	10	30	54
community	U20	8	1	25	0	2	18	35	32

(b)

(a)

NVC community		H10	H12	M15	M16	M17	M25	U4	U5
Calluna vulgaris-Erica cinerea heath Calluna vulgaris-Vaccinium myrtillus	H10								
heath	H12	60							
Scirpus-cespitosus-Erica tetralix wet									
heath	M15	57	47						
Erica tetralix-Sphagnum compactum									
wet heath	M16	42	35	62					
Scirpus cespitosus-Eriophorum									
<i>vaginatum</i> blanket mire	M17	52	43	76	64				
Molinia caerulea-Potentilla erecta mire	M25	33	24	52	42	37			
Festuca ovina-Agrostis capillaris-									
Galium saxatile grassland	U4	54	42	37	19	27	34		
Nardus stricta-Galium saxatile									
grassland	U5	65	56	53	34	45	39	63	
Pteridium aquilinum-Galium saxatile									
community	U20	53	50	38	21	31	32	71	63

(c)

NVC community		H10	H12	M15	M16	M17	M25	U4	U5
Calluna vulgaris-Erica cinerea heath	H10								
Calluna vulgaris-Vaccinium myrtillus heath	H12	25							
Scirpus-cespitosus-Erica tetralix wet heath	M15	53	18						
Erica tetralix-Sphagnum compactum wet									
heath	M16	29	21	68					
Scirpus cespitosus-Eriophorum vaginatum									
blanket mire	M17	43	13	75	75				
Molinia caerulea-Potentilla erecta mire	M25	42	0	70	38	60			
Festuca ovina-Agrostis capillaris-Galium									
saxatile grassland	U4	27	0	20	0	15	35		
Nardus stricta-Galium saxatile grassland	U5	25	0	18	0	13	33	74	
Pteridium aquilinum-Galium saxatile									
community	U20	29	0	23	0	18	38	68	63

Certain NVC communities were never confused, and these tended to be those from different habitat NVC types, for instance the dry heath communities H10 and H12 were not confused with the wet heath and mire communities M16 and M17. These communities are also most dissimilar in species composition, reflecting a general tendency for surveyors to confuse NVC types which are similar in species composition (Table 3.8a, b). However, some communities which are very similar in species composition were not often confused in the maps, particularly the mire communities M15, M16 and M17, which indicates that there was something distinctive about these communities. A regression of the MATCH similarity coefficients for pairs of communities against amount of confusion amongst the surveyors demonstrated a significant effect ($r^2 = 0.108$, p < 0.05) (Fig. 3.5a).

When the proportion of constant species in common between pairs of NVC communities is investigated, those which are most confused tend to be the communities with the highest proportions of constant species in common (Table 3.8a, c). This explains the high amount of confusion between M15 and M25, which have 70% of constant species in common. The proportion of constant species in common has a significant relationship with the amount of confusion amongst surveyors ($r^2 = 0.33$, p < 0.01) (Fig. 3.5b).



Figure 3.5 Relationship between proportion of confusion between pairs of NVC communities in the maps of seven surveyors and (a) similarity coefficients derived from match (y = 0.32x + 41.96, $r^2 = 0.108$), (b) proportion of constants in common based on floristic tables in MATCH (y = 0.35x + 1.56, $r^2 = 0.33$). Data were collected in field trials carried out in Snowdonia, 2008.

Surveyor experience, effort, survey conditions and method of assignment to NVC type amongst 2008 maps

Surveyor experience of vegetation mapping using the NVC ranged from 2 to 20 years, experience of surveying in upland habitats from 1 to 30 years, and cost per survey from £160 to £410 (Table 3.1). None of these three variables were

significantly correlated with amount of agreement or number of NVC types recorded per surveyor, except for upland habitat experience which was significantly correlated with mean area of agreement at the habitat level (Table 3.9). However, it was interesting that correlations with both NVC mapping and upland habitat experience were generally positive, with co-efficients consistently higher for upland habitat experience than for experience in NVC habitats. In addition, four of the seven surveyors stated that they were most familiar with surveying in upland habitats (Table 3.1) and, in general, these surveyors had higher levels of agreement and a higher number of NVC types than the other three surveyors. This implies that habitat specific knowledge is more important than experience at using the NVC system.

and cost charged per survey. Data were collected in field thats carried out in Showdonia, 2008.									
	experier	ice	experier	nce					
veriable	correlation	2	correlation		correlation				
vanable	coenicient	ρ	coefficient	р	coefficient	р			
agreement at habitat type	0.22	0.64	0.85 *	0.01	0.4	0.38			
agreement at community	0.36	0.43	0.49	0.26	0.74	0.06			
agreement at sub-community	-0.63	0.37	0.63	0.37	-0.07	0.88			
number of habitat types	0.24	0.61	0.4	0.37	-0.08	0.87			
number of communities	0.22	0.64	0.34	0.46	0.05	0.92			
number of sub-communities	0.29	0.53	0.51	0.24	-0.04	0.94			

Table 3.9 Spearman's rank correlation co-efficients between mean area of agreement per surveyor and number of types listed per survey for NVC major, community and sub-community types and number of years of surveyor NVC mapping experience, number of years surveying in upland habitats and cost charged per survey. Data were collected in field trials carried out in Snowdonia, 2008.

* Correlation is significant at the 0.05 level

The number of days between pairs of surveys varied from 1 to 38 (Table 3.2) and was not significantly correlated with agreement or number of NVC types. Time spent on the survey varied from 4.5 h to 10 h and length of route from 3.7 km to 9.6 km amongst the surveyors (Table 3.2). Neither time nor length of route was significantly correlated with agreement, however length of route was significantly positively correlated with number of sub-communities recorded (0.940, p = 0.002).

Finally, surveyors used a variety of methods to assign areas to NVC type (Table 3.1). Four surveys were carried out using formal identification tools (Rodwell, 1991b, 1992), whilst three surveyors relied heavily on their own experience supplemented in one case by their organisation's keys and in the other two by formal keys or tables. The latter three had the highest levels of agreement (mean habitat agreement 80.3%, community 37.7% and sub-community 19.5%, the highest number of NVC types (mean number of habitat types 3, community 13 and sub-community 19.7), the largest proportion of area mapped to sub-community level (mean 95.7%) and the largest number of polygons mapped (mean 191.3).

The two surveyors who mapped at the smaller resolution (onto OS maps) and had the smallest number of polygons in their maps also tended to have low numbers of NVC types compared with the surveyors who mapped at the larger resolution (Table 3.5, 3.6). However, there was no consistent pattern in pairwise comparisons of agreement between the two maps produced on OS maps (n = 1, habitat = 73.4%, community = 59.9%) and those produced on aerial photographs (n = 10, habitat = 80.4%, community = 44.4% and sub-community = 21.4%

3.3.3 Variation between 2008 maps and map produced in 1999

The map produced in 1999 had 76 polygons of 4230m² on average (Fig 3.6).



Figure 3.6 Map of National Vegetation Classification community produced during statutory mapping in Snowdonia in 1999.

The three major habitat types (H, M and U) in the 1999 map were present in each of the seven 2008 maps and although the 1999 map and all 2008 maps agreed that M was the dominant type, the 1999 map had a far higher proportion of type M than any

of the 2008 maps (Table 3.10). The 2008 maps generally had a higher number of communities (mean 10) than the 1999 map which only had five communities (H10, M15, M25, U4 and U20); these were present in each of the 2008 maps (with the exception of one map which did not include any H10). The 2008 maps which assigned >90% to sub-community type all had higher numbers of sub-communities (mean 16) than the 1999 map which had seven sub-communities listed (and only assigned 69.7% to sub-community). The sub-communities in the 1999 map (H10a, H10c, M15b, M25a, M25b, U4a and U4e) were each recorded in at least two 2008 maps, however there was little correspondence about the relative extent of sub-communities between the 1999 map and the 2008 maps.

Table 3.10 The areas of National Vegetation Classification (a) habitats (and non-NVC categories), (b) communities and (c) sub-communities in the map produced in 1999 as part of statutory habitat mapping of the Eryri SSSI, as a percentage of area mapped. Number of vegetation types recorded are also shown.

NVC habitat (code: description)	
H: heath	6.6
M: mire	85.2
U: acid grassland	5.8
Number of habitats recorded	3
Non-NVC categories recorded	
Boulders/Rock/Scree	2.3
NVC community (code: description)	
H10: Calluna vulgaris-Erica cinerea heath	6.6
M15: Scirpus-cespitosus-Erica tetrali wet heath	44.5
M25: Molinia caerulea-Potentilla erecta mire	40.7
U4: Festuca ovina-Agrostis capillaris-Galium saxatile grassland	2.8
U20: Pteridium aquilinum-Galium saxatile community	3
Number of communities recorded	5
NVC sub-community (code: description)	
H10a: Typical	3.8
H10c: Festuca ovina-Anthoxanthum odoratum	2.9
M15b: Typical	25.1
M25a: Erica tetralix	14
M25b: Anthoxanthum odoratum	1.5
U4a: Typical	0.2
U4e: Vaccinium myrtillus-Deschampsia flexuosa	2.8
Number of sub-communities recorded	7
Proportion mapped to sub-community	69.7

Mean spatial agreement between the 1999 map and each of the 2008 maps was 37.9%, with the maximum pair-wise agreement of 45.6% and the minimum of 22.8% (Table 3.11). This compares with higher agreement of 76.8% (range 86.6%-64.4%). at the habitat level and low agreement at the sub-community level of 16.5% (range 44.5%-1.7%), although this may be due to the fact that only 69.7% of the 1999 map was assigned to sub-community type. These levels of spatial agreement are very similar to the mean spatial agreement between pairs of 2008 maps (community = 34.2%, sub-community = 18.5% and habitat = 77.6%).

Table 3.11 Spatial concurrence between the 1999 map and each of the 2008 maps, assessed using percentage agreement. Sub-community figures only include the 1999 map and the four 2008 maps with >90% of the study area mapped to sub-community. Data were collected in Snowdonia during statutory mapping in 1999 and in field trials in 2008.

	NVC type			
Pair of maps	habitat	community	sub-community	
1999 Map-Map A	80.2	35.5	6.0	
1999 Map-Map B	81.6	40.7		
1999 Мар-Мар С	64.4	22.9	44.5	
1999 Map-Map D	86.8	45.6	13.7	
1999 Map-Map E	76.2	41.1		
1999 Map-Map F	69.8	38.9	1.7	
1999 Map-Map G	78.7	40.5		
Mean	lean 76.8		16.5	
SD	7.0	6.7	16.8	
SE	2.6	2.5	8.4	

When the proportions of each NVC community are correlated between the 1999 map and each 2008 map and between each pair of 2008 maps, the 1999 map is significantly correlated with all of the 2008 maps (p<0.01) (Table 3.12). However, not all pairs of 2008 maps are significantly correlated; out of 21 possible pairs of maps, 11 pairs are significantly correlated. There is no consistency in terms of the size of correlations involving particular maps, although pairs including maps E, F and G have the lowest correlations.

Table 3.12 Spearman's rank correlation co-efficients between proportions of different NVC, communities in pairs of maps. Data were collected during statutory habitat mapping in 1999 and in field trials carried out in Snowdonia, 2008. r = correlation co-efficient.

		Мар									
Мар		1999	Map A	Map B	Map C	Map D	Map E	Map F	Map G		
Map A	r	.695**	1								
,	p	0.001									
Map B	r	.612**	.509*	1							
	p	0.004	0.022								
Map C	r	.684**	.870**	.585**	1						
	p	0.001	0	0.007							
Map D	r	.674**	.635**	.922**	.721**	1					
	p	0.001	0.003	0	0						
Map E	r	.728**	.759**	0.191	.729**	0.277	1				
	p	0	0	0.421	0	0.237					
Map F	r	.762**	.672**	0.151	.703**	0.269	.839**	1			
	p	0	0.001	0.524	0.001	0.251	0				
Map G	r	.670**	.639**	.919**	.763**	.961**	0.289	0.35	1		
	p	0.001	0.002	0	0	0	0.217	0.13			

* Correlation is significant at the 0.05 level

** Correlation is significant at the 0.01 level

3.4 Discussion

This study has discovered a similar level of between-observer variation as that reported by Cherrill and McClean (1999a) who undertook a comparable trial of Phase I mapping. This study found a mean paired agreement of 34.2% (with a range of 5.4-69.6%) while Cherrill and McClean (1999a) report 25.6% (range 17.3-38.8%). This is surprising since the NVC is a finer-grained classification than Phase I and the maps were produced at a finer scale that in the Phase I trials, whereas previous work suggests that maps produced at coarser scales imply generalization and less precision (Kuchler, 1973; Millington and Alexander, 2000). Furthermore, the current study found that each 2008 map picked up large changes when compared with a map of the same area produced in 1999. These changes were of similar magnitude to the differences between 2008 maps, with mean paired agreement between the 1999 map and each 2008 map of 37.9% (range 22.8-45.6%). The fact that the mean paired agreement between the 2008 maps and between the 1999 and 2008 maps are so similar may imply that inconsistencies of these magnitudes are the norm in vegetation mapping, although further work would be required for corroboration. Basing management plans and monitoring on erroneous mapping could lead to mismanagement of vegetation and misleading conclusions about change.

Furthermore, this study found that agreement between the 2008 maps decreased as level of detail increased; from 34.2% for communities to 18.5% for sub-communities. A similar relationship was reported in Cherrill and McClean (1999a) and is presumably explained by the greater number of choices available as detail increases. The main sources of variation in vegetation maps have been reported to be classification of patch to vegetation type, heterogeneity within a patch, placement of boundaries for a given patch and resolution of mapping (Aspinall and Pearson, 1995; Cherrill and McClean, 1999a). From this study, it is hard to isolate the effects of these because an inaccurate boundary will give the impression of a misclassified vegetation type in the area between the inconsistently mapped boundaries, as will a large polygon in one map encompassing several smaller polygons of different vegetation types in other maps. Furthermore, in common with most sites to which NVC is applied our study site was predominantly composed of semi-natural vegetation which are spatially complex and defy classification into artificially

constructed types and have few linear features for spatial referencing (Averis *et al.*, 2004).

Since there are often few landmarks by which to orientate in upland areas, boundary location was expected to be a significant problem. When a 10 m wide strip around all boundaries was excluded and the maps were re-analysed, agreement was significantly increased (p<0.01 at the habitat level and p<0.05 at the community level). This is because the similarity in species composition and vegetation structure, and thus appearance, between many neighbouring vegetation types in this study meant that surveyors had to make seemingly arbitrary decisions about the location of the boundary somewhere in transition zone between vegetation types, something which has also been noted in other mapping studies (Castle and Mileto, 2003).

Reducing error arising from boundary location may be possible through quantification of location, but this requires sophisticated techniques not suitable for use in standard habitat mapping (Gosz, 1993; Kent *et al.*, 1997; Fortin *et al.*, 2000). Vegetation mapping in the Countryside Survey in 2007 was carried out using spatially referenced hand-held data imputers (Maskell *et al.*, 2008). There is also the fact that there are rarely discrete boundaries in nature and the cartographic 'model' of vegetation as abutted, homogeneous patches is not a good fit with reality on the ground. Pragmatically, the best that could be achieved on the ground is the use of GPS, although standard hand-held GPS units provide very variable accuracy (August *et al.*, 1994).

It may also be possible to categorise boundaries by type; Millington and Alexander (2000) noted that boundaries can be 'hard' where the transition between neighbouring vegetation types is crisp and distinct or 'soft, where the transitions are gradual. They point out that the boundaries are also sensitive to map scale; hard boundaries are easier to identify on fine scale maps but as maps become coarser in scale (i.e. show less detail), most boundaries become soft. Another way forward would be to change the representation of vegetation boundaries and to represent class membership as surfaces representing probabilities of class membership, similar to contour maps for altitude (Mark and Csillag, 1989). Since the boundaries of plant communities are in a constant state of flux (Kent *et al.*, 2006), it could be argued that

by the time the map is digitised, printed and in the hands of the end user, boundaries in areas of most rapid change may have shifted.

Given that the effect of removing boundary error on the overall correspondence between maps was relatively small (Table 3.5), discrepancies in assigned vegetation type are likely to be the main source of error. Visual inspection of the maps shows that surveyors frequently recognise the same ecotone and draw boundaries in similar locations, but assign different types to the vegetation on either side. Although hard to quantify, this is noticeable in the north-east of the study area, where vegetation occurred in visually homogeneous patches that were clearly visible on the aerial photograph. Here, all surveyors agreed on a general north-south boundary between heath and acid grassland, but disagreed on the actual vegetation classes each side of the boundary. Furthermore, the area chosen to represent more homogenous vegetation (Part A) did not have consistently higher levels of agreement between maps than the area with more heterogeneous vegetation (Part B).

Although the majority of communities were confused with each other, the most frequently confused communities were those which are ecologically related. There was often confusion amongst the mire/wet heath communities (M15 and M25 and M17 and M25) which have high species composition similarity coefficients with 70% and 60% of constant species in common respectively. The tendency to confuse communities which are similar in species composition and appearance has been found in other studies (Dargie, 1993; Hall, 1997; Cherrill and McClean, 1999b). The distinctiveness of species used to key out communities also has an impact on how often they are confused. The 'constant' species in the dry heaths, such as C. vulgaris, E. cinerea and Vaccinium myrtillus, tend to be large and easy to identify, whereas some of the constants in the mire communities, particularly the sphagnum mosses and T. cespitosum/E. vaginatum when not in flower, are inconspicuous and hard to distinguish (Rodwell, 1991b; Porley and Hodgetts, 2005). The mire communities (both in the lowlands and uplands) are considered difficult to classify and possibly extra time and sampling in these 'difficult' communities would help to increase consistency. Some of the surveyors recommend taking 'voucher' quadrats (a sample plot listing species and abundances in each vegetation type) for later identification to type. Target recording of vegetation patches with brief notes about species composition (possibly also noting their cover) can also be useful in

subsequently checking assignment to NVC type, particularly if every new community is noted. A more rigorous alternative would be to carry out complete floristic inventory of plots at a series of sample points over the site at a suitable intensity and formally assign an NVC type to each with the appropriate software. However, this would be so time-consuming as to make it impractical for widespread application.

The resolution of mapping is known to effect precision of vegetation maps, particularly at boundaries (Fortin, 1997, 1999), and in this study maps produced at a scale of 1: 5000 showed higher spatial agreement than maps produced at a scale of 1:10000 (although this could have been due to the ability to delineate areas of vegetation more easily on the aerial photographs than on the OS maps). Even where surveyors agree, such as in the patch of H10 (*C. vulgaris-E. cinerea*) heath in the north-west of the study area, some surveyors delineate patches of acid grassland within the dry heath matrix. Rodwell (2006) noted that small features are particularly hard to depict on coarse-scale maps; he suggested that the position and extent of flushes and springs are impossible to depict accurately at the scale of 1:10000 and may have to be shown notionally and used in conjunction with detailed description in accompanying text features.

The most appropriate resolution for a vegetation map depends on its purpose, e.g. extensive moorland is often mapped at a larger resolution than a 100 ha lowland grassland site. This may be partly due to the lower levels of variation on an extensive site, the costs involved and perhaps also its lower amenity value. Resolution may also reflect the particular background and perspective of the surveyor; those most familiar with mapping for the management of extensive sites apparently tend to map at a landscape scale, whereas ecologists involved in mapping small sites of special scientific interest prefer to work at a finer scale. Cost and tender instructions are also issues; the brief may specify mapping resolution or subcontractors may agree a daily fee to survey a set area at the most detailed resolution possible in the time available.

Surveyor effort was a factor in some of the general differences between maps, in that surveyors who took the longest route in their field survey produced a more detailed map. It is surprising that neither length of route or amount of time spent on the survey has no relationship with amount of agreement; this contrasts with studies of vegetation surveying (Nilsson and Nilsson, 1985), but agrees with specific mapping studies (Cherrill and McClean, 1999a). This study also shows that increased

experience generally results in increased consistency between observers, but this does not hold true for all surveyors and even then extensive experience only leads to a small increase in consistency (surveyors with more than 14 years experience of both NVC surveying and upland habitat surveying produced mean agreement of 85.0% at the habitat level, 46.8% for communities and 22.7% for sub-communities). Therefore, it could still be concluded that there are inherently high levels of variation in NVC mapping, which increased experience and effort cannot reduce significantly.

An important principle in the design of this study was to replicate the 'normal' behaviour of surveyors using NVC mapping, therefore the surveyors were not required to sample the vegetation. This does mean that reasons for confusion between mapped vegetation types are hard to determine. However, all the surveyors were aware that their survey was part of a field trial and that their maps were going to be compared with others for consistency. Potentially, this could bias surveyors to be more careful and precise than their standard practice, thus the low spatial agreement found is even more surprising, although it could also be that the production of more detailed than normal maps led to higher than normal inconsistency. The 1999 map used in this study was produced under very different conditions since it was part of statutory habitat mapping of large protected sites in Snowdonia. This may have meant that the surveyor was less careful and less detailed compared with surveyors in the field trial of 2008 which could contribute to the large differences between the 1999 map and each of the 2008 maps and may explain the low number of communities and sub-communities recorded in the 1999 map.

Several other differences in approach in the 1999 survey may have led to inconsistencies between the 1999 and 2008 maps. The 1999 map was drawn onto an aerial photograph taken seven years previously and in a different season to the survey. It was also not orthorectified and was blown up on a photocopier to a finer scale (1:4000 compared to the scale in 2008 of 1:5000). The surveyor was also asked to focus on habitats of particular conservation interest (wet and dry heaths) rather than the acid grassland. Although these factors may explain some of the differences, particularly the low proportion of U4 and U5 in the 1999 map compared with the 2008 maps (Table 3.5, 3.10), they do not explain the mean spatial agreement of only 37.9%. It is also interesting that the 2008 map produced by the same surveyor (map A) as the 1999 map had lower than average agreement with the

1999 map, indicating that repeat mapping by the same surveyor over time does not necessarily increase consistency. This is because peoples' skills and experience change over time, leading to different conclusions about the same landscape.

* The fact that repeat mapping by the same or different observers over time results in high disagreement illustrates the dangers in using mapping to detect change. If a conservation organisation commissioned repeat NVC mapping of a site, it is likely that large temporal change will be observed. Depending on the type of change, the organisation may conclude that their management is effective, that it is ineffective or they may be confused by the results. At the worst this could lead to the loss of important biodiversity and will certainly involve wasted resources, both for the cost of mapping and for subsequent decisions based on the spurious data. If this process were replicated at designated sites across the UK, this would have serious implications for the efficacy of conservation practice.¹

Recommendations have been made to increase the consistency of vegetation mapping, through the provision of unambiguous, standard methodologies, paired or team working, regular refresher courses and other procedures (Dargie, 1993; Wyn *et al.*, 2006; Brazier *et al.*, 2009 in prep). Whilst awareness of the need for rigorous standards in vegetation survey is welcomed, the subjectivity inherent in vegetation mapping should not be underestimated. More than half of the variation between two maps produced at the same time or repeated over time can be attributed to differing observer perceptions of the world. The utility of vegetation mapping is therefore restricted to site familiarisation and coarse inventory.

3.4.1 Limitations to this study

Although this trial did not include the surveying of quadrats in specific locations, this would be a useful way to explore reasons for discrepancies in vegetation type between maps. Floristic data from a series of quadrats could be assessed to evaluate the contribution of missed species, misidentified species and differences in estimation of abundance. It would also be interesting to see if not revealing to the surveyors that they are taking part in a field trial produces higher consistency, but this would not be professional.

3.5 Conclusions

Despite low spatial agreement, all the maps provided generally consistent information about the types of habitats and communities present and their gross distribution over the site as a whole. For most general survey and habitat inventory purposes this is perhaps not a problem as long as the associated uncertainty is acknowledged as differing perceptions of the same landscape will always lead to variation (Ruxton and Colegrave, 2006). However, for monitoring or surveillance, the level of error in NVC mapping will produce misleading conclusions, with more than 60% of 'change' apportionable to observer error. This could have major implications for the management of conservation sites across the UK. Although mapping systems such as the NVC were never intended to be used for monitoring (Rodwell, 1997), their ease of application means that they are used for this purpose (Appleby, 1991; Elzinga *et al.*, 2001; Stevens *et al.*, 2004b).

Given the limited resources available to conservation managers, the temptation to use mapping as a basis for management planning and monitoring is understandable, but where specific management objectives have been identified, then focusing resources on quantitative monitoring will often be more informative. Such methods have often been disregarded as being too expensive but perhaps it is time to reexamine them and work up methodologies which can combine the rigour of quantitative methods with the intuition of experienced surveyors to provide a robust basis for monitoring our changing environment.
Chapter 4

Consistency in habitat condition assessments

4.0 Abstract

Field assessments of site condition are frequently used as part of natural resource management. However, there are concerns about its subjectivity and risk of observer bias. This study provides the first test of the consistency of approaches used as part of Common Standards Monitoring (CSM), the statutory system for condition assessment of protected sites in the UK. Nineteen surveyors each used three different approaches to assess an area of fixed dune grassland on the coast of Anglesey in North Wales on the same day in summer 2008 and the resulting assessments of 'favourable' or 'unfavourable' condition were compared. The approaches produced very different conclusions due to discrepancies in sampling strategies and the arbitrary application of different thresholds. A further trial was carried out on the same area of fixed dune grassland, comparing assessments made by professional surveyors and volunteers using just one of the approaches used in CSM. Results showed that experience increases accuracy of assessments due to greater field skills, but that levels of consistency are similar regardless of experience.

This study notes that assessing quality of vegetation using current methodologies employed in the UK may provide an expert-based snapshot assessment of condition but does not provide a repeatable means of assessing change. Where it is used to provide a snapshot, it is recommended that approaches should always be samplebased, that guidance should specify sampling strategy and that experienced surveyors should be employed. Where it is necessary to assess change, diagnostic test methodology should be used with detailed and quantitative measurements from a small number of sites (identified as favourable and unfavourable by expert opinion) to validate more widespread snapshot condition assessments.

92

4.1 Introduction

4.1.1 Environmental management and condition assessment

Condition assessment is used in environmental management to ascertain the guality of a resource (Linke and Norris, 1993; Montgomery and Buffington, 1993; Parkes et al., 2003). Information from assessments is used to establish whether management has been effective, to provide evidence for development/funding applications, and to monitor change. Condition monitoring is designed to be rapid and low intensity and does result in less precision that quantitative monitoring (Goldsmith, 1991). Nevertheless, concerns have been raised that the results from qualitative methods such as condition assessment are too inconsistent (Gaston et al., 2006; Jackson and Gaston, 2008). Whilst all survey methodology is a compromise between accuracy/consistency of outcomes and cost effectiveness (Linke and Norris, 1993), it is still important that assessments undertaken at a local level are standardised and consistent to allow information to be scaled up and to monitor regional and global trends in resource quality and biodiversity (Balmford et al., 2005; Green et al., 2005). Condition assessment methods need to be empirically validated to check for inaccurate and inconsistent conclusions (Legg and Nagy, 2006; Jackson and Gaston, 2008).

This study focuses on the assessment of vegetation condition, for which no standard definition exists, although Gibbons and Freudenberger (2006) explain that it is a 'concept which reflects a desire to extend vegetation management from a concern about extent, type and configuration to one that also considers quality, health, function or viability.' Generally, estimates of quality require some position on what is a desirable state, this may be in terms of species richness, abundance, the presence of individual species assemblages or habitats (Firbank *et al.*, 2001). This is converted into a series of criteria through consultation with expert opinion about what constitutes 'good' or 'bad' condition for each resource, and vary from physiognomic aspects of vegetation to abundance of taxonomic groups indicative of ecosystem health (Oliver, 2002).

Vegetation condition can be assessed at a range of scales, from site to regional and broader according to programme objectives (Briggs and Freudenberger, 2006). Just at the site level, there are numerous methods employed to assess condition, ranging

from protocols to allow rapid assessment by non-specialists, which generally involve little species identification, to those that require trained specialists (Gibbons and Freudenberger, 2006). Protocols must be chosen according to objectives, time, expertise and other available resources. Whatever the method used, consistency of outcome is dependent on the rigour with which aspects are measured, which requires standardised and repeatable protocols.

4.1.2 Common Standards Monitoring

In this study, condition assessments used as part of a system of monitoring across designated conservation sites in the UK, Common Standards Monitoring (CSM) are investigated. In terms of CSM, condition assessment is the process of assessing that the habitat and species interests of a designated site are meeting the objectives for which the site was originally designated (Rowell, 1993b; JNCC, 2004c). The objectives list the attributes (characteristics of the interest feature that can be used to describe its condition) and associated targets. Typically, for habitat features, attributes include extent, floristic composition, vegetation structure, and physical characteristics. The same suite of attributes is used for each type of feature across the UK, with specific targets set at the site level to allow for geographical variation and local distinctiveness. The condition assessment monitors the features against the targets prescribed in the conservation objectives. After the attributes have been measured, it is possible to assign the feature to one of the agreed reporting categories (JNCC, 2008a). These can be broadly broken down into 'favourable' or 'unfavourable' condition, i.e. meeting targets or failing them (JNCC, 2008a).

CSM serves two main purposes. Firstly, it helps to guide management of the site by providing an early warning of whether all undesirable change is occurring. If it is decided that a feature is in unfavourable condition, further investigation should be made to ascertain the reasons why and corrective action taken. Secondly, it enables UK Government to undertake its national and international reporting commitments in relation to designated sites including the EU Habitats Directive which requires member states to monitor '...the conservation status of the natural habitats and species referred to in Article 2 (natural habitats and species of wild fauna and flora of Community interest), with particular regard to priority natural habitat types and priority

species.' (EEC, 1992). This indicates the effectiveness of current conservation action and investment, and identifies priorities for future action at a country level.

Central to CSM is the establishment of 'common standards' to ensure that consistent assessments are produced for each interest feature by different staff involved in monitoring across the UK. This was stipulated in the Environment Act of 1990 which charged the Joint Nature Conservation Committee (JNCC) with developing common standards for monitoring nature conservation in order to report at a UK and European level on the state of Sites of Special Scientific Interest (SSSIs). Although there is little published literature about CSM, questions have been raised about the extent to which the current application of CSM is providing a consistent and accurate representation of the condition of features and formal empirical validation of CSM has been recommended (Chapter 2, Jackson and Gaston 2008).

In order to investigate consistency in the application of CSM, it is important to untangle potential sources of inconsistency (Fig 4.1.). Sources 1 and 2 (inconsistencies arising from differences in sets of habitat guidance and from site-specific interpretation of guidance) will not be dealt with in this study, although a brief overview is provided here. The CSM protocols are differentiated by habitat, each of which has a set of recommended sampling strategies. This is confusing to the surveyor attempting to sample more than one habitat, and will have a large effect on the resulting consistency. For instance, the CSM guidance for woodland habitats recommends 'a structured walk around the site with a series of observation stops along the way' (JNCC, 2004b), whereas the CSM guidance for upland habitats permits a choice of sampling strategy (JNCC, 2005a). Thresholds for the feature to be considered in favourable condition also varies between habitats; in woodland habitats the whole area has to pass on all attributes to be scored as favourable condition, whereas in upland habitats 90% of the feature area must be favourable for the feature to be deemed in favourable condition.

95



Figure 4.1 Potential sources of inconsistency in the application of Common Standards Monitoring guidance.

The CSM guidance allows for site-specific interpretation of sampling strategy and thresholds. This is good in principle since, for instance areas of dry heath in Snowdonia in Wales and the Cairngorms in Scotland differ in scale, accessibility, species composition and ecology. However, these modifications rely on expert local knowledge and experience which is not always available, and also means that differences in opinions and changes in personnel could result in inconsistent conclusions.

Sources 3-5 (Fig. 4.1) will be addressed in this study. Source 3, the use of subjective measures (qualitative or semi-quantitative) rather than quantitative measurements is an important source of inconsistency. Many studies have investigated the consistency of semi-quantitative vegetation measures, mostly focusing on the visual estimation of plant cover or frequency (Hope-Simpson, 1940; Sykes *et al.*, 1983; Kennedy and Addison, 1987; Dethier *et al.*, 1993; Bråkenhielm and Qinghong, 1995; Klimes, 2003; Ringvall *et al.*, 2005; Vittoz and Guisan, 2007; Cheal, 2008). These have shown that such measures are subject to considerable variation, with

discrepancies of >20% both in repeated measurements by the same observer and measurements between different observers. In an interview survey of 60 conservation practitioners from across the UK, 69% of people involved in using JNCC's CSM guidance said that it was too subjective especially in the use of visual estimates of cover (Chapter 2). Furthermore, a trial of consistency using JNCC CSM guidance for woodlands, found that targets which involve a lot of interpretation, such as shrub cover or browsing levels, had the most variation between surveyors and that this could partly be attributable to differences in perception (Kirby, 2002).

Variation in experience and training is another important source of inconsistency; a variety of surveyors carry out CSM, ranging from dedicated agency staff through consultants to local volunteering groups. Surveyors receive varying quantity and quality of training, and although differences between observers (observer bias) is inevitable, they are likely to be higher for inexperienced/untrained observers. Although a relevant study (Cheal, 2008) found no relationship between cover estimations and recording experience, other studies have found that the use of experienced surveyors does reduce observer error (Scott and Hallam, 2003). The importance of training for consistency is particularly important for semi-quantitative measures, with the recommended system being to train people using measured samples and then refresh their memory at regular intervals, often using cards or photographs (Francini *et al.*, 2009). It has been shown that with careful attention to training, results using this system can produce high consistency (Haydock and Shaw, 1975; Bealey and Cox, 2004; Ross *et al.*, 2004).

A final source of inconsistency is the use of different approaches; this is very important because the UK statutory conservation agencies have adopted, modified or revised the JNCC CSM guidance (Chapter 2). A brief background to the development of CSM guidance is necessary to understand why alternative approaches have been developed. The original 'common framework' for monitoring protected sites in the UK (which provided the basis for the production of guidance for CSM) included a strong link between site-specific management and the condition of habitats (Rowell, 1993a, b; Rowell, 1996). However, when the standard CSM guidance was published, Countryside Council for Wales (CCW) staff felt that the intention in the framework set out by Rowell (1993a, b) to provide a link between monitoring and management was lost. They set out to develop an alternative approach to habitat monitoring using

97

methods combining field experience, site knowledge and site-specific quantitative measures (Brown, 2001; Hurford *et al.*, 2001; Hurford and Perry, 2001). This led to the 'Grid System' method now in use for designated conservation sites in Wales. This means that different statutory agencies across the UK use different sampling strategies, measurements and thresholds with the intention of producing comparable condition assessments. Although there are no known published studies exploring the effect of using different approaches to CSM, a study exploring the effect of using three different methods (sample plots, point-transects and a rapid visual assessment) to assess the conservation status of forests across Europe found that different methods produced contrasting results (Cantarello, 2007; Cantarello and Newton, 2008). Therefore, it is expected that the use of different approaches to CSM in the UK will produce inconsistent assessments.

4.1.3 Condition assessment methods used in this study

The current study investigates the condition monitoring method outlined in the JNCC CSM guidance (JNCC, 2004a) and compares it with two alternative approaches; the Grid System used widely across Wales by the Countryside Council for Wales (CCW) (Brown, 2001; Hurford and Schneider, 2006) and a new rapid assessment approach. The premise is that the Grid System approach is the most objective, unbiased method and therefore most the accurate of the three and thus provides a benchmark against which to test the other two methods. The following information outlines the three approaches and provides details about their specific application to dune grassland, the habitat used in this study. For further details of how the approaches were implemented in the field trials in this study see methods (section 4.2).

1. Standard (JNCC) CSM guidance

The JNCC CSM Guidance for sand dune habitats recommends the use of a 'W' shaped walk across the feature with at least 10 sample points of 2 m x 2 m placed by the surveyor (JNCC, 2004a). The shape of the 'W' is intended to ensure coverage of the extent of the feature and the points are chosen by the surveyor to 'represent' the variation in the habitat. Since observers can chose where to locate their plots and are likely to be familiar with what constitutes favourable condition, there is a possibility of

bias in sample location towards the 'good' patches (Robertson and Jefferson, 2000). Furthermore, the 2 m x 2 m plots each result in a large (4 m²) area of search which takes a relatively long time to search exhaustively and is likely to reduce efficiency and increase observer fatigue. The guidance provides lists of 'positive' and 'negative' indicator species and specifies how many of each should be present (and in what frequency) for the feature to be in favourable condition.

2. Grid System

CCW developed this approach under an EU LIFE Project on vegetation monitoring techniques (Brown, 2001; Hurford and Schneider, 2006). The method uses a systematic grid of points with a random start point; although not strictly true this form of sampling is conventionally considered as probabalistic with every point in the habitat having an equal chance of appearing in the sample (Brown, 2001). The approach requires site-specific prior knowledge of the spatial variation (known or expected) in the habitat in order to minimise the number of sample points required to achieve the target level of accuracy). When vegetation is sampled, the samples have a spatial relationship with each other, with samples close to each other more likely to be similar, a phenomenon known as spatial autocorrelation; 'everything is related to everything else but near things are more related than distant things' (Tobler, 1970; Kent et al., 2006). Spatial autocorrelation in vegetation has an effect on required sampling intensity in that habitats with high spatial variation require sample points spaced more closely than in more homogeneous habitats in which closely spaced points will waste resources as data points will contain similar and repetitive information (Dale and Fortin, 2002). By taking into account spatial variation in the habitat, the Grid System ensures that the number of samples will provide a sufficiently accurate estimate of the true population mean. The grid also ensures that the choice of sample point locations is objective and unbiased, with favourable and unfavourable aspects equally likely to be recorded. Furthermore, there are also a relatively large number of sample points used in this approach; between 30 and 50 sample points are placed across the grid, increasing the accuracy and consistency of the results (Grieg-Smith, 1957).

The area of each sampling unit across the grid also contributes to the accuracy of the results; this is also optimised according to the scale and arrangement of vegetation patches (Brown, 2006). As the size of the individual units of observation increases, a greater proportion of the spatial heterogeneity of the system is contained within a sample, and cannot be detected, while between-sample heterogeneity decreases (Wiens, 1989; Dungan *et al.*, 2002). The guidelines are designed to ensure that sampling area is kept to the minimum necessary, which is 0.5 m radius (~0.8 m²) in dune grassland (CCW, 2005b), resulting in a small area of search which minimises search time, reduces observer fatigue and increases consistency (Sykes *et al.*, 1983; Klimes, 2003; Archaux *et al.*, 2007). Attribute targets are set with reference to the area of sampling unit, using a size-dependant threshold for passing and failing individual points, and setting criteria for the proportion of passes across all sample points for the feature (in this case the habitat being assessed) to be deemed in favourable condition.

At each sample point the presence or absence of site-specific species indicator assemblages is recorded. These assemblages comprise a small number of species which define when a given habitat is in optimal condition (Hurford, 2006). For example, in dune grassland where optimal condition is associated with open grassland, the species assemblage comprises stress-tolerant species such as *Arenaria serpyllifolia* (Thyme-leaved sandwort) and *Euphrasia* sp. (Eyebright sp.) which decline in frequency as cover and height of grasses increases. When the quality of the vegetation is associated with a successional phase of development, the assemblage of species breaks down under pressure from competitive species, although this is not the case in upland heaths and woodlands. The purpose of these indicator assemblages is to provide an early warning of habitat degradation.

The selection of which species to include in these assemblages is critical and is carried out on a site by site basis by experienced monitoring ecologists in the CCW monitoring team. Point thresholds for favourable condition are based on the number of species in the assemblage that are present, for example five out of the seven species listed for dune grassland. The method becomes quicker as the surveyor becomes familiar with the appearance of the required habitat, or at least with the look of communities which contain the indicator species. It is contended that this system is objective and unbiased as opposed to the estimation of cover. Although this may

involve a large number of samples (e.g. 36 per 2500 m² of habitat) and thus be time consuming, due to the criteria of proportion of points to pass for the feature to be in favourable condition, the assessor can stop as soon as sufficient data is collected to make a judgement. For example, if the criteria is '70% of points in grid to pass for feature to be favourable condition', and the assessor has prior knowledge that the feature may be unfavourable, if they start in the worst area and find that more than 30% of total points in the grid fail, then sampling can stop (Brown 2006). This target threshold for proportion of sample points required to be in favourable condition for the feature to be deemed favourable is set by the CCW monitoring team and is specific to the feature on a particular site (Appendix 4.1). This relies on expert ecological and site knowledge to produce valid assessments.

3. Rapid assessment

The rapid assessment used in this study is based on a W- shaped walk in order to estimate the frequency of relevant species across the feature against the thresholds from the standard JNCC CSM guidance. This is intended to replicate a rapid assessment using expert opinion in which experts simply walk over the site using a mental checklist of criteria. Rapid methods are attractive due to ease of implementation and low resources required and are frequently recommended for site-based vegetation condition assessments (Sheil, 1995b; Gibbons and Freudenberger, 2006). However, studies looking at quick methods of visual estimation of plant cover and frequency have found problems with under-and over-estimations and significant differences between observers (Sykes *et al.*, 1983; Ringvall *et al.*, 2005).

4.1.4 Accuracy and consistency

Accuracy and precision are statistical terms used to describe the effectiveness of sampling designs. In any sampling exercise you wish to know the true population value of the parameter being sampled. However, it is not possible to obtain this value; what we obtain from sampling is an *estimate* of the true value. Accuracy and precision are ways of describing the relationship between estimates and the true value. Accuracy is the term used to describe how close the estimates are to the true value and precision is the term used to describe how tightly clustered the sample

estimates are around the average value as shown in Fig. 4.2 (Gotfryd and Hansell, 1985; Wong *et al.*, 2005).



Figure 4.2 Schematic illustration of terms precision and accuracy taken from Wong et al. (2005).

In this study, quantification of accuracy would require that the true value is known, which is impossible due to the subjective nature of condition assessments which ultimately rely on the judgements made when setting criteria for assessing conservation objectives. The accuracy of these criteria could be tested against a series of sensitive, quantitative measurements, but this would not provide information about the actual assessment methods. The best chance of being close to the truth in any of the assessment methods (i.e. being accurate) is when it yields precise results (with high agreement) and efforts have been made to reduce bias. However precision (which is taken as synonymous with consistency- the repeatability of methods at a point in time for the purposes of this study) does not rely on a 'truth'; observers might all agree on the assessment of a feature but all be wrong. Consistency can be measured according to how much agreement there is between different observers' assessments using any single condition-monitoring method.

The binary nature of condition assessment outcome into 'favourable' or 'unfavourable' has parallels with the assessment of diagnostic test results in medicine for which tests of accuracy have been devised (Altman and Bland, 1994b; Enøe *et al.*, 2001; Collaboration, 2008). For example, when investigating the presence of a disease, the results of a screening test (positive or negative, i.e. a binary classification) are compared to some absolute gold standard (actual diagnosis of positive or negative outcomes based on necropsy, biopsy or surgical inspection). The proportions of people with and without the disease who are correctly diagnosed by the screening test are calculated using a predictive matrix shown in Figure 4.3 (Hopley et al. 2001). From this matrix, two measures of the accuracy of the screening test can be calculated; sensitivity which is a statistical measure of how well a binary classification test correctly identifies the positive cases, and specificity which is a statistical measure of how well a binary classification test correctly identifies the positive cases.

		Condition (determine	d by "Gold standard")	
		True (Favourable)	False (Unfavourable)	
Sereening	Positive (Favourable)	Number of True Positives (TP)	Number of False Positives (FP) (Type I error)	Positive predictive value
test outcome	Negative (Unfavourable)	Number of False Negatives (FN) (Type II error)	Number of True Negatives (TN)	Negative predictive value
		Sensitivity (TP/(TP+FN))	Specificity (TN/(FP+TN))	

Figure 4.3 Relationships among terms in sensitivity/specificity matrix, taken from Hopley et al. (2001).

These accuracy measures are used in this study to judge the effectiveness of condition assessment methods, replacing positive and negative test results with observers' assessment of favourable and unfavourable condition respectively. Observers' assessments using a 'test' method will be compared with the same observers' assessments using a relatively unbiased method taken as the 'gold standard'. The proportion of observers who correctly assess the feature as being in favourable condition (true positive proportion, sensitivity) and proportion who correctly identify the feature as being in unfavourable condition (true negative proportion, specificity) will be calculated. High specificity is required when early

warning of degradation is required in order for appropriate changes in the management regime to be implemented. A specificity of 100% would indicate that all observers successfully recognize all unfavourable samples within a feature. This approach has been shown to be of use in conservation ecology, such as in the measurement of prediction success of presence/absence models for species and habitats (Fielding and Bell, 1997; Pearce and Ferrier, 2000) and the evaluation of methodologies for the identification of suitable restoration sites (Pakeman and Torvell, 2008).

Sensitivity and specificity can also be measures of consistency between observers using a single method when observers' overall assessments of condition are replaced with assessments of individual components, as long as they are also binary classifications. In the Grid System, assessments are made of favourable or unfavourable condition of every sample point across the systematic grid. Assessments of each sample point can be used in the matrix (Fig. 4.3) to calculate sensitivity and specificity for each combination of individual observers, which will measure how much agreement there is between observers' identification of favourable condition respectively.

The measures of sensitivity and specificity are dictated by the decision threshold or cut-off level, i.e. the measure is positive if the value was above some arbitrary cut-off and negative if below. In this study this is the criterion for favourable condition which in the Grid System takes the form of a required proportion of the feature with a set number of positive indicator species in order to classify the feature as favourable. If the threshold is varied over the spectrum of possible results, the sensitivity and specificity will move in opposite directions; as the number of positives increases the negatives naturally decrease (Zweig and Campbell, 1993). For each value of a threshold, there is a related sensitivity and specificity and it is only the entire spectrum of sensitivity/specificity pairs which provides a complete test of accuracy. Receiver Operating Characteristic (ROC) plots provide a view of this whole spectrum of sensitivities and specificities because all possible sensitivity/specificity pairs for a particular test are included (Altman and Bland, 1994a; Fielding and Bell, 1997; Zou et al., 1998). Obviously, a decision threshold must be chosen for a method to be used in condition assessment, but assessing method performance at a single threshold may result in misleading impressions about consistency or erroneous comparisons of accuracy since thresholds are often arbitrarily chosen. The ROC plot provides a comprehensive picture of the ability of a method to make the distinction being examined over all decision thresholds and thus provides a useful mechanism of assessment.

4.1.5 Objectives

This study aims to answer the following questions:

- 1. Do assessments of feature condition differ with the method used (Grid System, standard CSM or rapid assessment)?
- 2. What is the consistency of condition assessment amongst observers?
- 3. Are differences in assessed condition due to a) species frequency estimation versus sampling at points or b) size of sample unit?
- 4. What influence does the selection of attribute threshold have on favourable condition assessment?
- 5. Does observer experience reduce (a) accuracy and/or (b) consistency of condition monitoring assessments?

The study describes results from multi-observer field trials of three approaches to condition monitoring of a 50 m x 50 m area of fixed dune grassland in North Wales in May 2008. The number of favourable assessments using each method are calculated and diagnostic test methodology is used to draw comparisons. A trial comparing volunteer and professional assessments using a single approach to condition monitoring was also carried out. The aims are to (i) assess the consistency of results using different approaches to condition monitoring and (ii) explore the effect of observer experience on results of condition assessments. Implications for condition assessment and monitoring are discussed.

4.2 Methods

The study site was Aberffraw to Abermenai Dunes SAC (Special Area of Conservation) on Anglesey, an island off the North Wales coast, SH3569 (Fig. 4.4). Within this site an area of approximately 50 m x 50 m of 'fixed dune grassland' was identified and treated as though it was a separate small site for which assessments had to be made. The area was chosen in consultation with Clive Hurford, Dan Guest and Julie Creer from CCW. The habitat is classified as Phase I H6.5 Dune grassland, Annex I H2130 Fixed dunes with herbaceous vegetation ("grey dunes") and NVC types SD7 Ammophila arenaria–Festuca rubra semi-fixed dune community/SD9 Ammophila arenaria–Arrhenatherum elatius.



1 cm = 20 km

Figure 4.4 Location of study site in Wales used for field trials of condition assessments in 2008.

4.2.1 Field trials

Comparing condition assessment methods

Three methods for condition assessment were used. Firstly, the Grid System as developed as part of an EU LIFE Project (Brown, 2001; Hurford and Schneider, 2006) and detailed in the CCW Aberffraw to Abermenai Dunes SAC UK0020021 SAC monitoring report (CCW, 2005b). Secondly, following the standard JNCC CSM Guidance for sand dune habitats, specifically using the guidance for setting monitoring targets for fixed dune grassland (JNCC 2004a). Condensed versions of these two methods are provided in Appendix 4.1. Finally, a rapid assessment was

carried out, comprising a walk over the study area applying thresholds from the JNCC CSM guidance. A one-day trial was run at Aberffraw to Abermenai Dunes SAC on 29th May 2008 using 19 student volunteers from Bangor University. Each observer was asked to work individually in carrying out each of the three different methods across the chosen area (Fig. 4.5). All observers carried out the three methods in the same order (rapid assessment, standard CSM, Grid System) which was chosen to begin with the least detailed method and end with the most detailed method, such that potential bias resulting from increasing knowledge of the condition of the feature was minimised. Observers were also asked not to calculate overall condition assessment for any of the methods during the trial, again to minimise bias from prior expectations influencing observers' data collection. In order to investigate the effect of the size of the plot, observers were asked to record an additional 0.5 m radius plot at each sample point along the W walk using the standard CSM guidance. The sample points used in the Grid System were set up prior to the field trial by establishing a regular grid of 10 m spacing running north and east which filled the 50 m x 50 m area and placing numbered bamboo canes at every intersection (Fig. 4.5). All observers visited the same sample points over the grid.



Figure 4.5 Location of 50 x 50 m area of fixed dune grassland in Aberffraw to Abermenai Dunes SAC and detail of design of 'W' walk used in standard CSM approach and grid used in the Grid System in field trials of condition assessments in 2008. Crown Copyright Ordnance Survey 2009.

The instructions given to observers were as follows:

Rapid assessment: Spend 15 minutes walking over the area and record what proportion of the area each positive indicator species is present in.

Standard CSM guidance: Walk in a 'W' shape across the site, starting in the top left hand corner and finishing in the top right hand corner (Fig. 4.5). The length of the sides of the W should be roughly the length of the site. Stop and sample at 10 locations which you chose to build up a good picture of the condition of the area. At each stop note the presence of each of the positive indicator species in areas of 0.5 m radius and of 2 m x 2 m.

Grid System: At each plot over the grid (marked by numbered bamboo canes, Fig. 4.5), search an area of 0.5 m radius and note the presence of each positive indicator species, when 5 positive indicator species are found move onto the next plot.

Only one target, referred to in this study as 'positive indicator species', (which is referred to as 'typical species' in the standard CSM guidance where it forms part of the attribute of 'floristic composition', see appendix 4.1) was used in the trials, this was chosen in discussion with Clive Hurford and Dan Guest from CCW in order to be practical in the field with limited time and resources. The target for 'typical species' in the standard CSM guidance is the presence of eight typical species out of a possible list of 30-40 species; this was changed in the field trial to be comparable with the target for 'positive indicator species' in the Grid System. Both methods required the presence of five species out of a list of eight species, again chosen in consultation with CCW staff; they were mostly stress-tolerant (and competition intolerant) annual species indicative of early successional dune grassland. They are found in open, sandy areas and will begin to disappear during succession as these areas diminish and conditions become suitable for more competitive species. They comprise the 'site-specific indicator assemblage' used in the Grid System. In order to reduce observer bias, lower plants were excluded due to potential bias resulting from difficulties associated with their identification.

The following were chosen as positive indicator species: *Aira praecox* (Early hairgrass), *Arenaria serpyllifolia* (Thyme-leaved sandwort), *Cerastium* spp. (annual mouse-ears), *Erodium cicutarium* (Common Stork's-bill), *Erophila verna* (Common whitlowgrass), *Sedum acre* (Biting stonecrop) and *Viola tricolor* spp. *curtisii* (Dune pansy). Cerastium diffusum (Sea mouse-ear) and Cerastium glomeratum (Sticky mouse-ear) are annual mouse-ears and very difficult to distinguish between, so it was decided to record the presence of either or both as 'Cerastium spp.'. A particular problem with using annual species for monitoring purposes is that they can show large population changes from year to year due to responses to weather; this has been demonstrated for Cerastium spp. and Erophila verna (Hurford, 2006). Although this study uses these species in order to replicate current methodologies, this is a concern for interpretation of temporal change based on changes in such species.

There were several areas of difference in sampling strategy and target thresholds between the three methods (Table 4.1).

Second statement of the	table of the second		
	GRID SYSTEM	STANDARD CSM	RAPID ASSESSMENT
SAMPLE POINT LAYOUT	Grid with sample points marked by numbered bamboo canes placed at 10 m spacing	'W' walk with 10 sample points, chosen by each surveyor	No sampling. Walk over area chosen by each surveyor
SAMPLE POINT LOCATION	The same for all surveyors	Different locations chosen by each surveyor	n/a
NUMBER OF SAMPLE POINTS	36	10	n/a
AREA OF SEARCH	0.5 m radius	2 m x 2 m	n/a
MEASURMENT AT SAMPLE POINT	Presence/absence of specific species	Presence/absence of specific species	n/a
DATA COLLATION	By sample point	By species	By species
TARGET THRESHOLD (positive indicator species)	At least 5 positive indicator species present in at least 70% of sample points	At least 5 positive indicator species present at more than occasional* level over whole area	At least 5 positive indicator species present at more than occasional* level over whole area
How threshold is devised	Site-specific threshold decided by CCW monitoring team	Threshold specified in CSM guidance	Threshold specified in CSM guidance

Table 4.1 Sampling strategy and target thresholds used in field trials of three condition-monitoring methods (Grid System, standard CSM and a rapid assessment) at Aberffraw to Abermenai Dunes SAC in 2008. Further information is provided in appendix 4.1.

*Based on a version of the DAFOR scale which has been adapted to the particular characteristics of sand dunes:

DOMINANT: species appears at most (>60%) stops and covers no more than 50% of each sampling unit

ABUNDANT: species occurs regularly throughout a stand at most (>60%) stops and its cover is less than 50% of each sampling unit

FREQUENT: species recorded from 41-60% of stops

OCCASIONAL: species recorded from 21-40% of stops

RARE: species recorded from 1-20% of stops

Effect of experience

To test the effect of level of experience on consistency, a further one-day trial was run in the same 50 m x 50 m area of fixed dune grassland in Aberffraw to Abermenai Dunes SAC on 28th May 2008 using 22 observers; 6 conservation professionals (based on employment by a conservation organisation in a professional capacity) and 16 student volunteers from Bangor University. All observers carried out the same method (Grid System, Table 4.1, Fig. 4.5), recording the presence of positive indicator species in 0.5 m radius plots at 36 sample points marked out in the 50 m x 50 m area. All observers visited the same sample points and worked in pairs of the same category of observer as this was required by CCW who were involved in the field trial since this is their common practice in field trials of methodology. There were two mixed pairs (one student volunteer and one professional) but in each case the conservation professional carried out the survey with the volunteer acting as a data recorder, thus these pairs' results are counted as professional observers. This resulted in 7 pairs of volunteers and 4 pairs of professionals.

Training

To help to overcome the problems associated with using observers with little or no previous experience in vegetation recording, before all trials, an introduction to the site and the habitat was provided along with an indication of the appearance of favourable condition for fixed dune grassland. At least 2 hours of training in relevant species identification was also given; to ensure that the observers searched thoroughly for species, in the training fixed quadrats of 2 m x 2 m, 1 m x 1 m and 0.5 m x 0.5 m were marked out and subdivided into 16 cells. Observers were asked to search for the presence of each species in each cell and those who missed the presence of species were encouraged to look again at example specimens and develop a 'search image' for each species. In the training, bamboo canes marked at 0.5 m were used to demonstrate how to mark out an area of 0.5 m radius.

4.3 Results

4.3.1 Comparison of condition assessment methods

Do assessments of feature condition differ with the method used (Grid System, standard CSM or rapid assessment)?

The methods produced very different scores for the condition of the fixed dune grassland, with 17 observers (90%) assessing the feature as being in favourable condition using the standard CSM guidance compared with only 3 observers (16%) when the Grid System was used and 5 observers (26%) when the rapid assessment was used (Fig. 4.6).



Assessment method

Figure 4.6 Number of observers (out of 19) who assessed fixed dune grassland to be in favourable condition using three condition monitoring methods; Grid System, Standard CSM and rapid assessment. Data collected in a field trial at Aberffraw to Abermenai Dunes SAC in 2008.

However, this assessment has little meaning without exploring their relative accuracy. The difficulty is that with a qualitative measure such as condition, there is no objective truth to compare to. In this study, we have used the Grid System as the baseline for reasons outlined in the introduction. To investigate accuracy of the methods, sensitivity matrices comparing the standard CSM and the rapid assessment

as test methods against the Grid System as the gold standard were prepared. These consisted of True Positives, False Positives, False Negatives and True Negatives according to individual observer's assessments of condition (Fig. 4.3). Sensitivity matrices showed that the standard CSM guidance has a sensitivity of 1 i.e. none of the three observers who determined that the site was in favourable condition according to the Grid System determined that it was unfavourable according to the standard CSM guidance was used, a high proportion of observers (14 out of 19) missed that the site was unfavourable (according to the assessment using the Grid System) and instead assessed it as being in favourable condition.

Table 4.2 Sensitivity and specificity of site condition monitoring methods (standard CSM and rapid assessment) based on between-observer agreement with another method, the Grid System in field trials with 19 observers at Aberffraw to Abermenai Dunes SAC in 2008.

Sensitivity is the proportion of observers who correctly assess the feature as being in favourable condition and specificity is the proportion of observers who correctly identify the feature as being in unfavourable condition; power is the power of the test method to correctly identify favourable condition. TP= True Positive, FP= False Positive, FN= False Negative, TN= True Negative

GOLD STANDARD	TEST METHOD	ΤР	FP	FN	τN	sensitivity	specificity	power
Grid System	Standard CSM guidance(2 m x 2 m)	3	14	0	2	1.00	0.13	1.00
Grid System	Rapid assessment	0	5	3	11	0.00	0.69	0.00

In contrast, the rapid assessment had a sensitivity of zero, which means that none of the three observers who determined the condition of the site as favourable by the 'gold standard' Grid System determined it as favourable using the rapid assessment method. However, the specificity of the rapid assessment method was higher (0.7), since only 3 out of 14 observers missed when the site was in unfavourable condition.

What is the consistency of condition assessment amongst observers?

Figure 4.6 provides some indication of consistency of observers when using each method. The standard CSM method is the most consistent (with the ratio of favourable:unfavourable 17:2), followed by the Grid System (3:16) and the rapid assessment method (5:14). In order to understand the differences in outcomes, underlying reasons for the observers' assessments need to be explored

Are differences in assessed condition due to species frequency estimation versus sampling at points?

The low number of observers who assessed the site as favourable using the rapid assessment arises from an underestimation of the frequency of all the positive indicator species compared with sampling according to the Standard CSM (Fig. 4.7). This resulted in far fewer observers reaching the critical level of at least 5 species present in more than 40% of the area when the rapid assessment was used.



Figure 4.7 Mean frequency of positive indicator species estimated using a rapid assessment against that recorded using standard CSM guidance to assess condition of fixed dune grassland during field trials at Aberffraw to Abermenai Dunes SAC in 2008.

Looking at the species which were most underestimated in the rapid assessment (*Erophila verna*, *Aira praecox* and *Cerastium* spp.) they were three of the four least conspicuous of the species recorded according to an index of conspicuousness for each species, made up of the overall size of the plant, its colour, leaf area index and extent of flowering (Table 4.3). *Erophila verna* is slender-stemmed with small transparent seed-heads, *Aira praecox* is a tiny annual grass and *Cerastium* spp. comprises *Cerastium diffusum* and *Cerastium glomertum*, both annual mouse-ears which were desiccated at the time of the trials (Fig. 4.8). The species whose

estimated frequencies in the rapid assessment were closest to the recorded frequencies using the standard CSM guidance were the most conspicuous species, *Erodium cicutarium* and *Viola tricolor* s. *curtsii*, the largest two species recorded and both in full flower at the time of the trial.

Table 4.3 Index of conspicuousness (specific to the site and time of year) for species recorded in field trials at Aberffraw to Abermenai Dunes SAC in 2008. Information collected from the site, the Ecological Flora Database and the Wild Flower Key (Rose, 2006; Fitter and Peat, 2008).

HEIGHT (cm) 1=0.1-2, 2=2.1-4, 3=4.1-6, 4=6.1-8, 5=8.1+

COLOUR 1=very pale, 5=very bright

LEAF AREA (cm²) 1=0.1-1, 2=1.1-2, 3=2.1-3, 4=3.1-4, 5=4.1+

FLOWERING 1=no fresh flowers, old flowers still visible, 2=some fresh flowers and old flowers still visible, 3=less than half specimens with fresh flowers, 4=more than half specimens with fresh flowers, 5=all with fresh flowers.

SPECIES	HEIGHT	COLOUR	LEAF AREA	FLOWERING	CONSPICUOUSNESS (out of 20)
Erodium cicutarium	5	4	4	4	17
Viola tricolor s. curtsii	4	5	4	4	17
Sedum acre	1	5	1	3	10
Erophila verna	1	3	1	2	7
Aira praecox	1	3	1	1	6
Cerastium spp.	3	1	1	1	6
Arenaria serpyllifolia	25	1	1	1	8



Erodium cicutarium



Viola tricolor



Aira praecoxArenaria
serpvilifoliaImage: Construction of the series of the s

Figure 4.8 Photos of species recorded in field trials of methods to assess condition of fixed dune grassland during field trials at Aberffraw to Abermenai Dunes SAC in 2008. Photos taken by S. Hearn and C. Farmer (Farmer, 2008).

There was also more variation in their estimates of these positive indicator species (except for *Erophila verna*) when the observers used the rapid assessment method compared with the standard CSM guidance (Table 4.4). This is reflected in the mean difference from the group mean per observer.

Table 4.4 Difference in frequency of seven positive indicator species (actual and mean) estimated
using a rapid assessment method and recorded using standard CSM guidance to assess condition of
fixed dune grassland during field trials at Aberffraw to Abermenai Dunes SAC in 2008.

SPECIES	Ratio of mean frequency rapid	Actual frequer (%	ncy min-max	Mean difference in frequency from the group mean (%)		
	assessment: standard CSM guidance	rapid assessment	standard CSM guidance	rapid assessment	standard CSM guidance	
Aira praecox	1:2	0-80	40-100	18.0	12.4	
Arenaria serpyllifolia	5:7	0-100	30-100	23.3	13.6	
Cerastium spp.	6:9	0-100	50-100	20.5	10.7	
Erodium cicutarium	2:7	0-80	10-50	12.8	11.0	
Erophila verna	3:8	15-80	0-100	13.7	17.3	
Sedum acre	4:9	25-100	70-100	19.8	6.6	
Viola tricolor	5:8	15-100	40-90	22.9	10.9	

Are differences in assessed condition due to size of sampling unit?

The high number of observers who assessed the site as favourable using the standard CSM guidance as opposed to the Grid System, can partly be explained by the size of sampling unit. When data from the 0.5 m radius plots recorded by observers at each sample point along their particular 'W' walk are substituted for their original data recorded in the 2 m x 2 m plots, the number of favourable condition assessments decreases slightly from 17 to 14. The species area relationship means that observers recorded fewer species in the 0.5 m radius plots (mean 5.1 species per plot) than in the 2 m x 2 m plots (mean 5.8 species per plot) leading to lower overall frequencies using the 0.5 m radius plots (mean 58%) than the 2 m x 2 m plots (mean 71%). This means that fewer observers reached the target threshold of at least 5 species present in 40% of the sample plots. However, plot size exerts only a minor influence on condition assessments made using the standard CSM since there is a further difference of 11 condition assessments compared with the Grid System score of just 3 favourable assessments.

What influence does the selection of attribute threshold have on favourable condition assessment?

Thresholds are applied to determine how much of the feature (area in this study) must contain the specified number of positive species present for the feature to be deemed to be in favourable condition (Table 4.1). When thresholds for the three approaches are varied in 5% steps from 0 to 100%, outcomes change accordingly; the stricter (higher) the threshold, the fewer favourable condition assessments (Fig. 4.9). If the frequency threshold of 40% used in the standard CSM guidance and rapid assessment is applied to data collected using the Grid System, the number of favourable condition assessments increases dramatically from 5 (26%) to 15 (79%).



Figure 4.9 Number of observers who assess the area as being in favourable condition under a range of thresholds from 0 to 100% (in 5% increments) applied to data from field trials of three condition monitoring methods, the Grid System, standard CSM method and a rapid method carried out at Aberffraw to Abermenai Dunes SAC in 2008.

To further investigate the methods, sensitivity matrices (Fig. 4.3) were constructed comparing outcomes using the standard CSM guidance and the rapid assessment method with thresholds varied in 5% steps from 0 to 100%, against outcomes from the Grid System with a constant threshold of 70%. The range of sensitivity and specificity values for (i) the standard CSM method and (ii) the rapid assessment method were used to produce summary Receiver Operating Characteristic (sROC) plots using MetaDisc (Zamora *et al.*, 2006). The analysis added 0.5 to all zero values to facilitate analysis as MetaDisc is not able to cope with null values. The diagnostic odds ratio (dOR), a value representing the impact of any observer effect on the odds of the site being in favourable condition, changes with the threshold value and so an asymmetrical sROC curve (as opposed to a symmetrical sROC plot summarises the accuracy of the method as a single number, a perfect test will have an AUC close to 1 and poor tests have AUCs close to 0.5 (Hopley and Schalkwyk, 2001).

The resulting sROC curves (Fig. 4.10) show circles which are the pairs of specificity and sensitivity calculated for each 5% increment between 0 and 100%. There are 21 possible pairs of values (0%, 5%, 10% etc), but the graph shows fewer than this because some of the 5% increments result in repeated specificity and sensitivity values. There are more repeated specificity/sensitivity pairs (and thus fewer data points) for the standard CSM method than for the rapid assessment method. The middle curve is the line of best fit through the full set of 21 pairs of specificity and sensitivity values, the two outer curves indicate the 95% confidence intervals. These sROC curves demonstrate that although neither method is good at discriminating between favourable and unfavourable condition, the standard CSM method is twice as good with an AUC of 0.68, compared with 0.36 for the rapid assessment method.



Figure 4.10 Asymmetrical Summary Receiver Operating Characteristics plots of sensitivity against specificity calculated using thresholds from 0 to 100% (in 5% increments) for (a) the standard CSM method and (b) a rapid assessment method against the Grid System as the 'gold standard'. The circles are the pairs of specificity and sensitivity calculated for each 5% increment between 0 and 100%, the middle curve is the line of best fit and the two outer curves indicate the 95% confidence intervals and AUC=area under the curve (summarises the accuracy of the method as a single number, a perfect test will have an AUC close to 1 and poor tests have AUCs close to 0.5 (Hopley and Schalkwyk, 2001). Data were collected during field trials at Aberffraw to Abermenai Dunes SAC in 2008.

4.3.2 Effect of experience

Does observer experience increase (a) accuracy and/or (b) consistency of assessment?

a) Accuracy

The two groups of observer pairs drew different conclusions about the condition of the feature, with all seven volunteer observer pairs assessing the feature as unfavourable, (0:7 favourable:unfavourable) and the majority of the four professional observer pairs assessing the feature as favourable (3:1 favourable:unfavourable). Assuming that the site is in favourable condition according to the SAC monitoring carried out by CCW in 2005 (CCW, 2005b), none of the volunteer observer pairs assessed the site correctly, whereas 75% of the professional observers (3 pairs) correctly assessed the site as favourable. This is explained by consistent underrecording of species' frequencies by the volunteer observer pairs (by a mean of 11%, Fig. 4.11, Table 4.5).



Figure 4.11 Mean frequency of seven positive indicator species recorded by pairs of volunteer and professional observers using the Grid System method during a field trial at Aberffraw to Abermenai Dunes SAC in 2008.

Volunteer observers' under recording of species frequencies was most marked in *Erophila verna* and *Aira praecox* (Table 4.5) which is probably due to their inconspicuousness; they are both small annual species (<2 cm in height) which had already flowered at the time of the field trial, and have low scores for conspicuousness (Table 4.3). Frequencies from the two groups of observers were most similar for *Erodium circutarium and Viola tricolor* (Table 4.5), the largest and most conspicuous of the species (Table 4.3).

Table 4.5 Difference in mean frequency of positive indicator species recorded by volunteer and professional observers using the Grid system to assess condition of fixed dune grassland during field trials at Aberffraw to Abermenai Dunes SAC in 2008.

SPECIES	Conspicuousness: out of 20 (from Table 4.3)	Volunteer observers' mean frequency	Professional observers' mean frequency
Erodium cicutarium	17	4	8
Viola tricolor s curtsii	17	47	50
Sedum acre	10	81	91
Erophila verna	7	40	59
Aira praecox	6	63	87
Cerastium spp.	6	74	83
Arenaria serpyllifolia	8	58	68

b) Consistency

Looking at the consistency of recording, there are high levels of agreement *within* groups of volunteer and professional observers, both in overall condition assessment and in the proportion of sample points which are recorded as favourable (Fig. 4.12). The range of results from each group are also similar, with the seven pairs of volunteer observers recording between 43% and 67% (range of 24%) of the sample points as favourable and the four pairs of professional observers recording between 63% and 80% (range of 17%) as favourable. Due to the differences in number of pairs of observers between the two groups, this produces a slightly smaller standard error for volunteers compared with observers (Fig. 4.12).



Figure 4.12 Mean percentage of sample points assessed as being in favourable condition by pairs of volunteer (n=7) and professional observers (n=4) using the Grid System methodology during a field trial at Aberffraw to Abermenai Dunes SAC in 2008. Bars are \pm 1SE.

To investigate the variation within the groups of observers in more detail, agreement was calculated between observers within each group by drawing up matrices consisting of True Positives, False Positives, False Negatives and True Negatives, i.e. volunteer pair A's results for plots numbered 1 to 36 was compared with volunteer pair B's results for the same plots (Fig. 4.3 and 4.5). Meta analysis was performed by entering all sets of pairs of calculated values into the software MetaDisc (Zamora *et al.*, 2006) in order to produce pooled sensitivity and specificity for volunteers andprofessionals respectively and to compare heterogeneity. The analysis added 0.5 to all zero values to facilitate analysis as MetaDisc is not able to cope with null values. A low *p*-value for the heterogeneity chi-squared statistic (or a large chi-

squared statistic relative to its degree of freedom) provides evidence that the differences across the studies are greater than expected by chance alone. The metaanalysis of agreement between pairs of observers (treated as 'studies' in this analysis) within the two groups of observers displayed a slightly higher pooled sensitivity for professional observers (0.95) compared with volunteer observers (0.88) (Fig. 4.13, Table 4.6). This shows that professionals had higher levels of agreement than volunteers in the identification of sample points that pass. The heterogeneity chi-squared value within the volunteers' pairs of sensitivity figures (25.31) is higher than within the observers' pairs (15.46), reflecting the volunteers' larger range of sensitivity values. Taking sample size into account heterogeneity is within the range expected by chance alone for the volunteers but greater than expected by chance for the professional observers (p=0.01). This suggests that the four pairs of professional surveyors had a lot of variation in their results, although interpretation is limited by the small sample size.

Professional observers also displayed a slightly higher specificity (0.76) compared with volunteer observers (0.71), which shows that professionals had more agreement (although again not by much) in the identification of sample points which fail. The heterogeneity chi-squared value is again lower for the professionals than for the volunteers, and both *p*-values are within the range expected by chance alone; the large *p*-value for the volunteers (0.98) is due to the low heterogeneity relative to sample size.

Table -	4.6	Meta-analys	is of	measure	s of	sens	itivity	and	specificity	within	groups	of	volunte	er a	and
profess	siona	l observers	using	the Grid	Sys	tem a	pproa	ach to	o condition	assess	sment d	uring	g field	trials	s at
Aberffra	aw to	o Abermena	i Dun	es SAC ir	200	08.									

	VOLUNTEER OBSERVERS	PROFESSIONAL OBSERVERS
SENSITIVITY		
Pooled sensitivity	0.88	0.95
Heterogeneity chi-squared	25.31	15.46
df	20	5
p*	0.15	0.01*
SPECIFICITY		
Pooled specificity	0.71	0.76
Heterogeneity chi-squared	8.8	5.41
df	20	5
p*	0.98	0.37

*p-value of <0.1 indicates heterogeneity above that expected by chance alone



Figure 4.13 Pooled sensitivity and specificity according to agreement between (a) volunteer observer and (b) professional observers in condition assessment of plots using the Grid System in fixed dune grassland during field trials at Aberffraw to Abermenai Dunes SAC in 2008. Circles show sensitivity/specificity values for each unique combination of pairs of observers in volunteer (n=21) and professional groups (n=6), diamond shows pooled sensitivity/specificity respectively and dashed line shows 95% confidence interval.

In order to investigate the implications of differing observer assessments of condition for the detection of temporal change, the dataset was modelled to treat each pair of observers as a 'baseline survey' compared against other pairs as 'repeat' surveys. Calculation of the type I error rate for different groups of observers showed that repeat surveys within the group of volunteer pairs had the lowest type I error rate of 0 (Table 4.7) since pairs of volunteers consistently found the site to be unfavourable (and were all incorrect compared to the 2005 survey by CCW). Repeating baseline surveys carried out by professional observers with volunteer observers and vice versa results in a high type I error rate, of 0.75, due to the high incidence of false changes. Repeating baseline assessments carried out by pairs of professionals with other professionals also had a high type I error rate, of 0.5, implying that even professional observers will falsely detect a change in 50% of cases.

Table 4.7 Type I error rate between and within assessments by pairs of volunteer and professional observers using the Grid System approach to condition assessment during field trials at Aberffraw to Abermenai Dunes SAC in 2008.

Baseline assessment by	Repeat assessment by	Type I error rate
professionals	professionals	0.5
volunteers	volunteers	0
professionals	volunteers	0.75
volunteers	professionals	0.75

4.4 Discussion

4.4.1 Consistency of different approaches to condition assessment

This study suggests that the use of different methods for condition assessment produces different results, even when the objective, the area and the observers are the same. Observers were much more likely to assess fixed dune grassland as being in favourable condition using the standard CSM guidance compared with either the Grid System or a rapid assessment. Whilst studies comparing results from different vegetation sampling methodologies have come to similar conclusions (Kercher *et al.*, 2003; Vittoz and Guisan, 2007; Lavorel *et al.*, 2008), these discrepancies are particularly worrying since different approaches to CSM are currently used across the UK (Chapter 2). This will result in inconsistent assessments being used for purposes of comparison and since the Grid System approach is the main method used to assess condition of habitat features in Wales (Allen 2008, *pers. comm.*) this could result in Welsh sites appearing relatively worse off in pooled UK data. Looking at the figures from the first cycle of CSM reporting, the proportion of terrestrial habitat features (and selected habitat feature groupings) assessed as favourable condition in Wales are consistently lower than in the other UK countries (Table 4.8).

Table 4.8 Condition assessment data for selected habitats and all terrestrial habitats from England, Northern Ireland, Scotland and Wales from the first cycle of habitat condition assessments of designated conservation sites: Site of Special Scientific Interest (SSSI), Special Area of Conservation (SAC), Special Protection Area (SPA) and Area of Special Scientific Interest (ASSI) (EN, 2005; NIEA, 2008; SNH, 2008; WAG, 2008).

	Percentage of habitat features in					
	unfavourable					
	favourable	recovering	unfavourable			
ENGLAND (2000-2005)	condition	condition	condition			
Coastal	89)*	11			
Grassland	83	3*	17			
Heathland	73	}*	27			
Woodland, wood pasture and parkland	86	5*	14			
All habitats	45	24	31			
			N			
NORTHERN IRELAND (2002-2008)						
Coastal	38	2	61			
Grassland	35	3	63			
Heath and Upland habitats	45	5	50			
Woodland	4	22	73			
All habitats	36	8	56			
SCOTLAND (1998-2005)						
Coastal	66	1	17			
Lowland grassland	36	8	29			
Lowland heathland	43	0	18			
Upland	59	5	29			
Woodland	47	11	33			
All habitats	56	6	26			
WALES (2000-2006)						
Coastal	28	13	59			
Lowland grassland	0	19	81			
Lowland heathland	0	35	65			
Upland	35	6	60			
Woodland	26	21	53			
All habitats	25	17	58			

Habitat groupings are chosen to be comparable but there is some inconsistency in reported categories.

* 'Favourable' and 'unfavourable recovering' figures are reported as a single category. Years in brackets indicate date of reporting cycle.

It could also be that many sites in England, Scotland and Northern Ireland are reported to be in favourable condition when a more rigorous assessment method would show that they are unfavourable. Since condition is a value-laden concept, requiring data to be viewed through a pre-determined concept of what constitutes 'good', this study is not able to report on the relative accuracy of the different approaches. However, the Grid System uses a systematic sampling design, objective
measures and has a tight link with the specific site, habitat type and management practices; therefore it is most likely to be 'accurate' (Brown, 2001; Hurford and Schneider, 2006). Following this assumption, the low specificity and high type I error rate (probability of rejecting a true null hypothesis i.e. proportion of observers incorrectly assessing the feature as favourable) associated with the use of standard CSM guidance will lead observers to overestimate the amount of features in favourable condition whereas in contrast, observers using the rapid assessment are likely to underestimate the amount of features in favourable condition due to its low sensitivity and high type II error rate (probability of rejecting a true hypothesis i.e. proportion of observers incorrectly assessing the feature as unfavourable).

Moreover, the target frequency threshold is shown to exert a strong influence on final condition assessment, with the low number of favourable condition assessments using the Grid System approach compared with the Standard CSM guidance primarily due to differences in thresholds (it was harder to find 5 species together in 70% of the sample plots in the Grid System than for 5 species individually to be found at >40% frequency in the standard CSM method). Again, it is difficult to determine which threshold leads to the most accurate assessments. The thresholds in the standard CSM guidance are decided by experts and are intended to serve as a trigger mechanism so that, when changes that fall outside these thresholds are observed or measured, some further investigation or remedial action is taken (JNCC, 1998). In order to produce comparable results, thresholds in alternative protocols should lead to consistent representations of features (Jackson and Gaston, 2008). It could be that the threshold for the Grid System set by CCW for this site makes it too hard to find the feature to be in favourable condition.

The results of this field trial illustrate the choice between imposition of a standard, objective system (with little flexibility) versus a system with site specificity, flexibility and more reliance on expert opinion. The positive indicator species target in the standard CSM guidance is pre-determined to be at a low level (making it easy to pass the feature) whereas the threshold used in the Grid System approach is set specifically for the feature and site (making it more appropriate to the condition of the specific feature and harder to pass the feature). By specifying a threshold, the standard CSM guidance attempts to remove the decision from the individual surveyor and thus to reduce subjectivity; however this creates an inflexible and imprecise

system. Condition assessment is always value-laden; perhaps this needs to be accepted and used to advantage, as in the Grid System approach which allows the surveyor to set a site-specific threshold which is more likely to correctly identify favourable condition. This suggests that there is no substitute for experienced observers who can identify what constitutes favourable condition. However, setting thresholds on a site by site basis makes comparisons between sites and countries difficult.

It is interesting to explore reasons for the differences between methods. The finding that the rapid assessment resulted in a far smaller proportion of observers determining that the test area was of favourable status than when the Standard CSM method was used is perhaps unsurprising. Reasons are straightforward; since the data collation and threshold are the same, the only difference is in the sampling strategy. The rapid assessment relies on impressions gained whilst walking over the site, without the sort of detailed, ground-level searches forced by recording sample plots, so inconspicuous species are overlooked or misrepresented, a common problem in vegetation monitoring (Scott and Hallam, 2003). This finding is supported by studies looking at the assessment of plant diversity which found that rapid assessments tended to over-represent dominant species with large basal areas and underestimate small species such as mosses and lichens (Gaucherand and Lavorel, 2007; Lavorel et al., 2008; Giordani et al., 2009). English Nature's Validation Network also report similar findings in several studies, notably one in lowland heathland, where a rapid assessment under-assessed the number and abundance of negative indicator species, and ascribed this to the difficulty of viewing smaller broad-leaved species through and under an ericoid canopy (Bealey and Cox, 2004; Ross et al., 2004; Ross and Bealey, 2006). They show that this is particularly a problem in closed sward conditions, where the key species are small herbs at low frequencies and less readily picked up in qualitative assessments based on larger-scale, more rapid sampling. These reports recommend always sampling in plots, training in difficult species, carrying identification prompts and thorough searching; however the increase in time requirements must be considered.

Size of sampling unit between the Grid System and standard CSM method was shown to have only a minor effect; with the smaller plot size in the Grid System slightly reducing the number of species recorded at each sample point, from an average of 5.8 species in the 2 m x 2 m plots compared with 5.1 species in the 0.5 m radius plots) (Elzinga *et al.*, 2001). This is because the species area curve is saturated at a small area, therefore increasing the area leads to few new species being discovered. In addition, the lower species numbers recorded at each plot in the Grid System may be compensated for by the larger number of plots sampled.

It is also important that each method provides a consistent result between observers and although none of the methods produced universal agreement, the standard CSM guidance and the Grid System approaches produced higher agreement than the rapid assessment. The differences in sampling approach and subsequent binary classification of the area into 'unfavourable/favourable' obscures within group variation and makes comparative measures of consistency difficult. It could be that although 90% of observers agreed that the area was favourable using the standard CSM guidance, the threshold was set so low that there is a large amount of hidden variation. However, as previous studies have also shown (e.g. Elzinga *et al.* 2001), the variation of estimation of frequency was higher for every species using the rapid assessment method compared with sampling in plots.

Although these discrepancies between assessments using different methods has implications for the consistency of figures used to report against targets at a national level, it has been argued that the required amalgamation of site-based assessments results in information loss (Boyle, 2007; Jeeves, 2007). The fact that 25% of SAC habitat features in Wales have been recorded as in favourable condition compared with 45% of SSSI habitat features in England (Table 4.7) does not mean anything without knowledge of the relevant linkages between ecological and management factors. In an exploration of habitats on protected sites in England, Everett (2004) found that there was insufficient understanding of the ecology and management of upland habitats to know how to achieve favourable condition and sometimes even if the evidence base was sufficient, the target condition was unattainable, as in the case of salt marshes in the face of predicted sea-level rise. She concluded that this makes targets such as Natural England's aim of 'bringing into favourable condition, by 2010, 95% of all nationally important wildlife sites' (HM Treasury, 2004) both unachievable and meaningless. Together with the fact that there is often disagreement on what constitutes favourable condition, it seems that target setting and reporting is often a waste of valuable resources (Everett, 2004; Jeeves, 2007).

4.4.2 Effect of experience

Professionals and volunteers drew almost completely different conclusions about the condition of the feature, with the majority of professional assessments that it is in favourable condition being the most accurate in relation to the actual condition of the Aberffraw SAC as assessed by CCW monitoring ecologists (CCW, 2005b). This is because they were able to detect the presence of more species compared with volunteers. This has also been noted in other studies and is likely to be due to professional observers' training in species identification, experience of these particular species and retention of a 'search image' during the field trial (Haydock and Shaw, 1975; Scott and Hallam, 2003). Careful training in species identification and use of a species photo card during surveys could help to overcome this limitation in assessments by less experienced surveyors. It could also be that professional surveyors are able to compare the site with their knowledge of other sites and rank it accordingly, which means that their assessments correspond with the assessment made by CCW staff during the actual statutory monitoring of this site.

In terms of within-group consistency, however, there was no marked difference between the two groups, with professionals having only slightly more agreement in whether sample points were favourable/unfavourable than did the volunteers. In a sampling trial to test the effects of observer bias on the recording of species frequency, Hurford (2006) also found that groups of students and of professional ecologists both achieved high levels of within-group consistency. This indicates that if the measurement is kept simple and sufficient training provided, even inexperienced surveyors can produce consistent results, although as training or experience improved, some observers might improve their assessments faster than others leading to lower within-group consistency.

The current confusion between and within CSM methodology also raises concerns for detecting change, which requires reliable results in which the difference between observers/assessment methods is less than the differences over time. Even when just one of the methods (the Grid System) is modelled to represent re-surveying over time, professional observers falsely detect a change in 50% of cases. This suggests that the use of current approaches to CSM to assess change in feature condition over time will provide inconsistent and misleading information.

Although analysis in this study demonstrates the possibility of using diagnostic test methodology to evaluate alternative condition assessment methodologies, a thorough evaluation requires a verified accurate method, the development of which was outside of the scope of this project. Natural England's Validation Network has attempted to do this but has only investigated certain aspects of the standard CSM guidance rather than look at overall assessment of condition (Bealey and Cox, 2004; Ross *et al.*, 2004; Ross and Bealey, 2006). They collected data using (i) standard condition assessment protocols and (ii) quantitative techniques, for a series of attributes at the same locations, and used these to compare condition assessments at the attribute level. They concluded that the amount of agreement between methods varied according to the attribute, with the least agreement in attributes based on presence/absence of specific suites of species. They made no conclusions about the accuracy of standard CSM guidance.

4.4.3 Limitations to this study

The main limitation is that the comparison of methods here was only carried out on one site; a full evaluation of CSM approaches should be carried out over a range of habitats to encompass the spectrum of field methods, types of measurements and thresholds involved. A further limitation is the small sample size of observers; increasing this (particularly of professional observers) would give much more power to the overall outcomes, since errors and outliers will have less effect.

Finally, the study is limited by the use of only one target; this was chosen to be practical given the time and resources available but does not represent the full set of decisions which surveyors are normally required to make about a site. The more targets which are measured and the more decisions which have to be taken, the more complicated the monitoring becomes and the more likely it is that errors are made (Gibbons and Freudenberger, 2006). The process of stacking up targets so that a feature has to pass all targets at any one sample point or over the whole area also makes it progressively harder for the feature to be assessed as favourable.

4.5 Conclusions

This study has shown that the use of different approaches to Common Standards Monitoring, a widely used system in conservation in the UK, produces very different conclusions (even with the same observers in the same area on the same day). This has implications for reporting, evaluation of management practices and monitoring of change. The main explanations for these differences are discrepancies in sampling strategies and especially the application of different thresholds. This study concludes that where different approaches are used then pooling of results to report against targets at national and international levels should be avoided if possible or interpreted with caution.

Assessing quality of vegetation using current standardised methodologies employed in the UK may provide an expert based snapshot assessment of condition but does not provide a universal repeatable and reliable means of assessing change. Where it is used to provide a snapshot assessment, I recommend that approaches should always be sample-based as opposed to using rapid estimation of species attributes, that guidance should specify sampling strategy and that surveyors should be appropriately trained. Where it is necessary to assess change, detailed and quantitative measurements from a small number of sites (identified as favourable and unfavourable by expert opinion) should be used to validate more widespread snapshot condition assessments. Diagnostic test methodology, as used to measure the accuracy of screening tests in medicine, could provide an appropriate means of using the quantitative data to validate qualitative assessments.

Chapter 5

Quantitative vegetation monitoring; sampling for the detection of temporal change

5.0 Abstract

Vegetation monitoring aims to detect change in natural resources. However there is a concern that many programmes fail to meet their objectives due to a lack of focus on the magnitude of change to be detected and the power of the sampling to detect this change. There is also a need to take spatial pattern into account in designing monitoring. This study uses a systematic sampling design across a large, complex upland site in North Wales to provide objective and unbiased data to look at how spatial variation and sampling intensity effects sampling effort to detect simulated levels of change. Individual species patterns can be interpreted in relation to ecological and management factors however interpretation is limited for composite groups of species and 'communities' due to the responses of individual species cancelling each other out and the reduction of information to a single metric. There is a lack of predictability in spatial variation across spatial scales, which leads to an unpredictable relationship between sampling intensity and effort to detect simulated temporal change. It is also concluded that the detection of small changes over time (equivalent to those implicit in the objectives for current monitoring programmes) require unfeasible levels of effort for widespread use.

This study makes several recommendations. Firstly, monitoring programmes must include an *a priori* definition of meaningful potential change and ensure that the variables measured are sufficiently sensitive over the monitored time period and are indicative of the desired change. Secondly, since spatial variation is site-specific, optimal use of resources can be obtained by using prior knowledge and information from a pilot study, however this requires understanding of the metric used and how it is affected by spatial scale. Approaches should also balance Type I and Type II error rates carefully according to management needs considering the costs of either missing a change or falsely identifying one. Finally it is recommended that standardised protocols are used in quantitative approaches; these should include objective sampling strategies with unbiased measurements of variables of interest.

5.1 Introduction

5.1.1 The need for vegetation monitoring methods which work

Vegetation monitoring programmes aim to detect changes over time in specified characteristics of species and habitats, such as extent, abundance or condition. This requires two stages; an initial, reliable assessment of the particular characteristic followed by repeated assessments at later dates to allow change to be measured over time (Elzinga *et al.*, 2001). Reviews and studies have concluded that many monitoring schemes are unable to detect the level of change implicit in the objectives, largely because of a lack of explicit focus on the effect size (size of change) and power required (Green, 1979, 1989; Critchley and Poulton, 1998; Brown, 2001; Foster, 2001; Stefano, 2003; Legg and Nagy, 2006; Seavy and Reynolds, 2007). If these aspects are not made explicit in the objectives and subsequently in the design of monitoring, then it is unlikely that changes of ecological significance will be detected at an early enough stage.

It has been argued that these failures of monitoring programmes are due to a disparity between theory and application which is in turn explained by inadequate coverage of monitoring design in degree and post graduate courses and by a general lack of accessible, practical information on aspects of monitoring design (Legg and Nagy, 2006; Slater et al., 2006). A recent review of monitoring and surveillance the UK concluded that there is a disparity activities across between recommendations in the published literature and evidence about the ability of monitoring schemes to detect a change, either predicted or actual (Slater et al., 2006). This review noted that there is a lack of accessible information as to the most effective approaches to sampling design, data analysis and determining of trends. This was corroborated in my interviews with sixty conservation practitioners in the UK about conservation monitoring methods (Chapter 2). Although repeatability was the most commonly listed criteria for an effective monitoring method, only five interviewees mentioned the ability to detect change as a necessary criterion for an effective scheme and only one person stated the requirement of 'ability to detect change of a stated level'. Slater et al (2006) recommend that 'survey organisations should determine whether samples are adequate for detecting a magnitude of change in variables'.

Other reasons for the poor performance of many monitoring schemes are that scientists view it as routine practice more relevant to conservation management than academia and that it has long been organised adequately. Also, many relevant studies and papers are published in the grey literature to which scientists have limited access, for example the report by Slater et al in 2006 and internal papers published by the statutory bodies.

Recommendations to improve monitoring by focusing on the size of change to be detected are supported by a number of studies (e.g. (Green, 1989; Foster, 2001; Kirk, 2007; Nakagawa and Cuthill, 2007; Seavy and Reynolds, 2007) who urge *a priori* power analysis before commencing monitoring and surveillance programmes. *A priori* power analysis should be carried out at the planning stage of a monitoring project, so that the number of samples needed to detect a biologically important change is known. There is no general rule for what constitutes a 'biologically important' change but it must be described with a measurable value, either a quantity e.g. 20% cover or a qualitative state e.g. cover class. Specifying these quantities or states involves considering the species' possible and potential responses according to knowledge about their ecology, management schedule, time span of monitoring and resources available (Elzinga et al 2001). Although predicting populations' responses is challenging, Elzinga et al (2001) point out that this should not be an obstacle as without a measurable objective, there is no means to assess whether current management is effective.

Despite these recommendations, there are no accessible studies of environmental monitoring programmes which state their *a priori* effect size (change) or power. Even examples of monitoring schemes in the UK which have published *post hoc* levels of change detection and power analysis are infrequent and these tend to be confined to well-monitored animal populations, particularly birds, mammals and butterflies. Out of five single-species monitoring schemes reported by The Tracking Mammals Partnership, the dormouse is the only species for which the ability to detect change is reported (via the National Dormouse Monitoring Project), with sample sizes sufficient to detect changes in distribution of 25% over 25 years nationally and 50% distribution over 25 years regionally (JNCC, 2009c). None of the schemes report on survey power. The UK Butterfly Monitoring Scheme reports a similar capacity; changes of 50% abundance with a power of 80% over a 20-year period for 37 out of 51 species

(VanStrien *et al.*, 1997). Some schemes are able to detect a much smaller change, such as monitoring counts of *Rhinolophus hipposideros* (lesser horseshoe bats) across 79 roost sites in Wales which are able to detect a 5% increase or decrease in abundance over a 5-year period (Warren and Witter, 2002).

Information about the ability of vegetation monitoring schemes to detect quantitative change is harder to find; changes in individual plant species can be quantified through assessment of distribution or abundance at specific sites or across regions. such as data from the Countryside Survey (CS) in the UK which gives percentage change in abundance of common species such as stinging nettle and bramble across the UK however the power is not stated (Haines-Young et al., 2002). Countryside Survey data is also used to assess change in plant species richness in the UK biodiversity indicator 'Plant diversity in the wider countryside' (Partnership, 2009). Change in plant communities are harder to quantify because they are hard to define, consisting of varying proportions of associated species (Dale, 1994; Palmer and White, 1994; Kent et al., 1997; Kent et al., 2006). This means that most approaches to measurement of plant communities (and their change over time or space) are fraught with subjective decisions over placement of boundaries and assignation of vegetation type (Kent and Coker, 1992; Dargie, 1993; Cherrill and McClean, 1999a; Stevens et al., 2004b). The Countryside Survey data does provide assessments of change in extent of broad habitats and features, but again provides no indication of power or error.

While hard to interpret, ordination axis scores of vegetation data over time do offer a powerful way to detect change, for example through definition of a target and setting a limit for proximity to that target. However, multivariate approaches do not provide a basis to compare with any external standard. An alternative approach to quantified monitoring is to focus on qualitative change, for example the Red Data List for Vascular Plants in Great Britain (Cheffings and Farrell, 2005) which places species in standard IUCN threat categories (IUCN, 2009) by applying thresholds to trend data of varying quality. The UK biodiversity indicators 'Status of Biodiversity Action Plan Species/Habitats' are also based on change in categories of trend assessment such as 'stable', 'increasing' or 'decreasing' (UKBAP, 2008). Finally, the UK Common Standards Monitoring system (JNCC, 2006) allows change in the condition of plant communities to be assessed, between categories such as 'unfavourable' and

'favourable'. However these qualitative assessments provide a coarse measure of change with no indication of power or error. There is clearly a need for measurements of plant species and communities of sufficient precision and sensitivity to allow meaningful temporal changes to be detected with suitable power.

5.1.2 Detection of change: theory

Vegetation monitoring sets out to measure rate and direction of change over time and to assign explanations to the changes. Magnitude of change, power and acceptable error rates need to be established according to time spans, cost of missing a change and type of habitat sampled.

Power is the certainty of the estimation of a particular variable based on the sample observations i.e. in the case of monitoring the certainty of being right in observing that there has been a change (Legg and Nagy, 2006). It is frequently stated that statistical power is adequate if it is 0.8 or above (Di Stefano 2003), however the level of power depends on the objectives of the surveillance programme and the cost of being wrong in observing that there has been a change. Carrying out an a priori power analysis requires careful thought as power depends on several factors including effect size, error variance, sample size and the Type I error rate. Looking at these factors in relation to detecting change over time, effect size is an estimate of the difference between samples in years 1 and 1+x, which the sampling must be able to detect. The larger the effect, or the greater the change in the system, the easier the change will be to detect and the higher the power achieved by a given sampling regime. However, the size of the effect is usually unknown at the planning stage and is often small in size, thus the limits of acceptable change should be set and the sampling designed so that a change of that magnitude will be detected if it occurs (Legg and Nagy, 2006). This will ensure that sampling does not overlook changes of management importance, for instance if a conservation manager wants to be 80% certain (achieve a power of 80%) whether a change of at least 20% of cover of Erica tetralix has taken place but the current sampling can only detect a change of 40% of cover with a power of 80%, then the survey design is not working and cover changes will be too large by the time they can be detected. In this case, the easiest way to

achieve the necessary power is to increase the sample size as the larger the number of samples the higher the power of the test.

Finally, Type I error rate must be considered. By convention the Type I error rate, when the true null hypothesis is rejected and a false hypothesis accepted i.e. a false difference, is traditionally set at 0.05, which means that if the statistical test is repeated many times, using different random samples from the same population, about 5% of the outcomes would be 'significant' if the null hypothesis was correct (Foster, 2001). Increasing the acceptable Type I error rate can greatly increase the power of the test. The Type II error rate, the failure to reject a true null hypothesis and instead reject a true hypothesis, i.e. a missed difference, also needs to be considered... Since power is [1- Type II error rate], the higher the probability of making a Type II error the lower the power. When the conventional Type I error rate of 0.05 and power of 0.8 are used, the probabilities of making Type I and II errors are 5% and 20% respectively. This means that the cost of making a Type I error is four times more than the cost of making a Type II error is four times more than the cost of making a Type II error is four times more than the cost of making a Type II error is four times more than the cost of making a Type II error is four times more than the cost of making a Type II error is four times more than the cost of making a Type II error is four times more than the cost of making a Type II error is four times more than the cost of making a Type II error is four times more than the cost of making a Type II error is four times more than the cost of making a Type II error is four times more than the cost of making a Type II error is four times more than the cost of making a Type II error is four times more than the cost of making a Type II error is four times more than the cost of making a Type II error is four times more than the cost of making a Type II error is four times more than the cost of making a Type II error is four times more than the cost of making a Type II error is four times more than the cost of maki

However, in conservation management the cost of Type II errors may be greater than Type I errors (Brown, 2001). For instance, if cattle were used to maintain an area of wet heath, making a Type II error might mean failing to detect damage to the wet heath and keeping the cattle on too long, risking difficult to reverse damage to the habitat, whereas Type I errors may mean that damage is detected which is not actually happening, leading to the removal of the cattle and safeguarding of the resource (although there is a management cost). In this case, the type II error rate should be lower than the type I error rate, for example if the conventional error rates are swopped, the type II error rate would be 0.05 and the type I error rate 0.2. This means that there is only 5% probability of making a type II error (resulting in continued grazing and subsequent loss of the wet heath) and a 20% probability of making a type I error (resulting in removal of the cattle and safeguarding of the wet heath). This would require a very high power (95%) and possibly large sample size which would greatly increase monitoring expense. Thus Type I and Type II error levels need to be balanced, and should be determined on criteria external to the data, i.e. to minimise the costs of both kinds of error (monetary, social or aesthetic).

5.1.3 Importance of spatial pattern

Description of spatial pattern is intrinsic to vegetation monitoring, both to enable adequate sampling design and develop understanding of processes. There are a range of factors that cause spatial pattern in vegetation; environmental factors that are spatially heterogeneous such as geology and geomorphology, morphological factors, based on the size and growth patterns of the plants, phytosociological factors caused by species' relationships and the interactions between vegetation and grazing herbivores (Dale, 1999; Storch and Gaston, 2004). When vegetation is sampled, the samples have a spatial relationship to each other, with samples close to each other more likely to be similar; '*everything is related to everything else but near things are more related than distant things*' (Tobler, 1970; Kent *et al.*, 2006). This concept of spatial autocorrelation or lack of independence forms the basis of spatial pattern by creating a certain amount of predictability in the arrangement of plants and patches, often exhibited in periodicity of some kind such as groves of trees alternating with open grasslands across a landscape (Fortin *et al.*, 1989; Legendre, 1993; Koenig and Knops, 1998; Dale, 1999).

This pattern or patchiness of vegetation has been given a hierarchy of scale by landscape ecologists, taking the patch scale as the mosaic of patches which form the landscape in any local area (Wiens, 1989; Withers and Meentemeyer, 1999; Kent et al., 2006). These patches can be aggregated to give the landscape scale and disaggregated to give the individual species distribution at the plant scale. Ecological spatial processes such as competitive interactions or dispersal operate in many ways at different scales. For instance, it has been shown that at a fine scale, species shift in abundance over space and time in response to disturbance and that species can be characterised according to their ability to invade gaps, suppress other species and to persist when suppressed themselves (Grubb, 1982). These effects of competition and biological processes at a fine scale may override relationships at a landscape scale, such as between climate and vegetation (Wiens, 1989; Dungan et al., 2002; Fagan et al., 2003; Wagner, 2003; Kent et al., 2006). Several studies have investigated the scaling of species distribution patterns, to see whether self-similarity can be used to infer patterns at larger or smaller scales (Kunin, 1998; Kunin et al., 2000; Hartley et al., 2004). It has been concluded that domains of scale operate,

where processes within one domain can be predicted but that there are sharp and unpredictable transitions between domains.

Ability to detect spatial pattern is a function of extent, the overall area encompassed by a study and grain, the size of the individual units of observation (Wiens, 1989; Dungan *et al.*, 2002). When the scale of measurement of a variable is changed, the spatial variance changes, depending on whether the grain or the extent is altered. As grain size increases, a greater proportion of the spatial heterogeneity of the system is contained within a sample, and cannot be detected, while between-grain heterogeneity decreases. This means that different patterns will emerge at different scales of investigation and according to different sampling designs (Legendre *et al.*, 2002; Fagan *et al.*, 2003).

Various studies have investigated the effect of sampling design on pattern detected, and vice versa, finding that systematic sampling designs are most appropriate for detecting spatial autocorrelation, but only if the sampling step (interval between successive samples) is at the correct resolution and preferably varies (Cochran, 1977; Goslee, 2006). It has also been found that spatial pattern has a large effect on the analysis of vegetation data collected across a range of scales, largely due to spatial autocorrelation which restricts the assumptions of independent observations (Fortin *et al.*, 1989; Lennon, 2000; Legendre *et al.*, 2002). These studies have concluded that spatial structure of the sampling universe must be taken into account when choosing both sampling design and subsequent analysis.

5.1.4 Spatial pattern analysis

Spatial pattern analysis looks for evidence of regularity, randomness or clumping of a given vegetation variable within a particular area. When ecologists calculate mean values for a given variable over space, they implicitly interpolate known values to unmeasured points (Robertson, 1987). Parametric statistics provide estimates of variance about unbiased means (as long as assumptions about sample independence and normality are met) and are used to describe attributes and test hypotheses. However, the presence of spatial autocorrelation often violates assumptions about normality which leads to imprecise estimates of variance which differ substantially from overall population variance. Explicit analysis of spatial pattern

will mean that autocorrelation resulting from interacting environmental, species and management factors can be taken into account in the interpretation of ecological processes and the effect of treatments.

There are two main ways of exploring spatial pattern of vegetation: firstly the use of metrics to quantify spatial pattern and secondly geostatistics or spatial statistics. Landscape metrics measure aspects of pattern, such as degree of isolation or connectivity of patches (Schumaker, 1996; Gustafon, 1998; Kent, 2007). Spatial statistics are generally used (i) to identify the spatial scales over which patterns (or processes) remain constant and (ii) to interpolate or extrapolate point data to infer the spatial distributions of variables of interest (Turner *et al.*, 2001). Both approaches are strongly influenced by classification system (if used) and the spatial scale the data was collected over.

There are a wide variety of methods available for geostatistics and spatial statistics which use point data for a variable and assume that it is spatially continuous (Dale, 1999). Blocked-quadrat variance methods allow spatial patterns to be identified from transects or grids of contiguous sampling plots (Fortin and Dale, 2005). Adjacent plots are aggregated into blocks of increasing size and the variance amongst blocks is plotted against block size so that a peak of variance indicates the size of a regular pattern (Dale, 1999; Schlup and Wagner, 2008). These methods are used to explore the spatial structure of individual species and pairs of species. Other techniques commonly used to explore the structure of spatial variation and suggest suitable sampling schemes are geostatistical estimation and interpolation (Isaaks and Srivastava, 1989; Clark and Harper, 2000; Webster and Oliver, 2001). The methodology was developed in mining geology and has been used in many disciplines including soil science, farming and ecology (Journel and Huijbregts, 1978; Kent and Coker, 1992). Data from sampled locations are used to calculate spatial variance and this is used to estimate predictions for unsampled locations within the same area. Geostatistical estimation utilises the variogram, which provides an unbiased description of the scale and pattern of spatial variation of a given property. It does so by summarising the way in which the variance of the property changes as the distance and direction separating sample points varies (Oliver and Frogbrook, 1998). The variogram is defined by:

In this function, Z(x) and Z(x+h) are the values of Z, the property of interest, at any two places x and x+h separated by h, a vector having both distance and direction and known as the lag. The symbol E denotes the expectation. The semivariance, v, at a given separation is half the expected squared difference between values at that separation. The variogram is the semivariance plotted against a series of lag distances, and can be unbounded indicating that the full range of variation present at this level of resolution has not been encompassed (Fig 5.1a). The slope shows that the semivariances increase as the separating distances become larger, describing how the property is more similar at places close to each other than at places further away. The variogram can also be bounded, reaching an asymptote known as the sill variance (Fig 5.1b). Where the variogram reaches the sill is known as the range of spatial dependence; points separated by distances less than the range are spatially dependent, those separated by distances greater than the range are spatially uncorrelated. Bounded variograms suggest that the property displays a patchy distribution and that all of the variation at this level of resolution has been encompassed by the sampling. This variogram also meets the ordinate at a positive value known as the nugget variance, which encompasses any measurement error and unresolved spatial variation (variation over less than the smallest sampling interval). Variograms can also be pure nugget, where there is no apparent spatial dependence in the data (Fig 5.1c). For continuous properties, such as vegetation abundance or height, this usually means that the sampling has failed to resolve the variation at the scale of investigation.



Figure 5.1 Forms of variogram (a) unbounded (b) bounded (c) pure nugget (Oliver and Frogbrook, 1998).

Kriging is the method of estimation in geostatistics; there are several types, of which ordinary kriging is most commonly used (Frogbrook, 2000). This involves using weighted linear combinations of the data within a specified area. The weights are derived from the variogram and allocated so that estimates are unbiased and the estimation variance is minimised. Kriged estimates are then used for interpolation and mapping of variables for use in planning and management. The sampling intensity for interpolation should relate to the range of spatial variation at a site in order to optimise sampling effort (Frogbrook, 1998). This can be achieved by matching the sampling intensity to the structure of the spatial variation; if the variogram is bounded then the sampling interval should be about half of the range of spatial dependence and can be determined according to the precision required.

Kriging is effective for individual plant species, however it is problematic with more than one variable, although co-kriging can look at the pattern of pairs of species by looking at the effect of scale on covariance (Webster and Oliver, 2001). Multivariate datasets such as those containing information about plant communities are hard to analyse through kriging. This is the case with spatial analysis in general; plant communities are less well investigated, and work has tended to focus on relationships between species richness and resource availability, biomass, and productivity (Grieg-Smith, 1986; Mittelbach *et al.*, 2001; Cornwell and Grubb, 2003). Vegetation science tends to concentrate on true gradients in space i.e. the

differentiation of communities along environmental gradients using discrete or continuous approaches, and to assume that observations are spatially independent (Antoine *et al.*, 1998; Schlup and Wagner, 2008). There is little integration of plant community spatial pattern analysis with standard non-spatial methods for detecting community structure, such as classification and ordination (Wallace *et al.*, 2000; Wagner, 2003, 2004; Wagner and Fortin, 2005; Kent *et al.*, 2006). One approach is multi-scale ordination which partitions the variance produced in an ordination according to spatial intensity of the scales sampled across (Hoef and Glenn-Lewin, 1989; Wagner, 2003, 2004; Couteron *et al.*, 2005; Kent, 2006).

5.1.5 Spatial pattern, sampling design and monitoring

Monitoring has tended to use classical sampling approaches, primarily simple random sampling or stratified random sampling (Cochran, 1977; Minium et al., 1999; Brown, 2001). These approaches ignore spatial pattern, unless it is incorporated into sampling stratification. However, the sampling approach of a monitoring programme effects the spatial variation picked up which has an impact on the ability to detect temporal change (Section 5.1.3) (Brown, 2006). This is because an increase in the variance increases the effect size which decreases the power to detect a specified change. The use of a systematic, regular grid of points as a sampling layout for monitoring has been suggested as an alternative to classical sampling where spatial pattern is of interest (Brown, 2001). This systematic sampling provides an unbiased estimate of the mean, with every point in the sampling universe having an equal probability of appearing in the sample, as long as the start point is randomly chosen (Cochran, 1977). Knowledge of spatial variation can be used to optimise sampling efficiency by using an extensive sampling design to identify areas of high variability for more intensive sampling (Smartt, 1978; Caldas and White, 1983). This does however rely on the spatial variability between scales being predictable and means that there will be an interruption to field work between phases of the project.

Sampling design also includes error which increases the effect size in monitoring and thus decreases the power to detect a temporal change. The two most important sources of error are sampling error and observer error, either of which might stay constant over time or can change between sampling periods (Thompson, 1999;

Ferretti, 2009; Gottardini *et al.*, 2009). Sampling error comes from sampling design which fails to ascertain the parameters of the variable of interest; increasing the number of samples generally reduces sampling error (Legg and Nagy 2006). Observer error is produced during measurement and classification and depends on several factors including the extent to which a standard protocol is available (and adhered to), the objectivity of measures used, the experience and training of the surveyor and the quality assurance procedures implemented (Sykes *et al.*, 1983; Milberg *et al.*, 2008).

The type of vegetation measurement also has a large effect on the estimation of variance and thus on the temporal change which can be detected (Green, 1979; Bonham, 1989; Bråkenhielm and Qinghong, 1995; Critchley and Poulton, 1998; Elzinga et al., 2001; Carlsson et al., 2005; Scott and Smart, 2006). Frequency measurements are recommended as quick, consistent and objective (Bonham, 1989; Bullock, 1996; Ejrnaes and Bruun, 2000; Ringvall et al., 2005; Ramsay et al., 2006) although they have limited sensitivity to change (Bonham 1989). Visual estimation of cover is also fast and easy but has been shown to be subject to considerable observer error (Sykes et al., 1983; Kennedy and Addison, 1987; Cheal, 2008) whereas cover estimated through frequency of species at points or sub-plots within larger plots takes longer but is consistent between observers and robust to cover changes during the growing season (Bonham, 1989; Ejrnaes and Bruun, 2000; Vittoz and Guisan, 2007). Structural measurements such as height also provide additional sensitivity but take more time and are prone to observer error (Stewart et al., 2001). All vegetation sampling relies on the identification skills of observers, with small and fine-leaved species (particularly lower plants) producing high discrepancy between observers (Sykes et al., 1983; Lepš and Hadincová, 1992; Scott and Hallam, 2003). Studies have also shown that between observer error is higher for species with low frequency and cover (Milberg et al., 2008).

5.1.6 Objectives

The objectives are to answer the following questions:

- 1. Can a systematic sampling strategy be used to produce ecologically meaningful predictions of (a) taxa spatial distribution pattern and (b) plant community distribution pattern?
- 2. Is spatial variation predictable as sampling intensity increases?
- 3. Are relationships between sample size/sampling effort and sampling intensity consistent using presence/absence and mean height data?
- 4. Do (a) sampling intensity and (b) spatial heterogeneity affect the sample size/sampling effort required to detect change of a specified magnitude?
- 5. How does sampling intensity affect the quality of information recorded?
- 6. Are there key thresholds in magnitude of temporal change and magnitude of power required to detect that change?
- 7. What sample size/sampling effort would be required to detect temporal change in plant communities of management and ecological importance, using (a) quantitative monitoring and (b) qualitative monitoring?

This study uses quantitative inventory data from an upland protected site collected using an extensive, systematic sampling approach to identify two areas of contrasting spatial heterogeneity. These two areas are sampled with increasing intensity and data explored to see whether spatial variation remains predictable. The relationship between spatial pattern of vegetation and measured environmental variables is also explored. Data is used to simulate consistent levels of temporal change across all sampling intensities and in the two areas of contrasting spatial heterogeneity; sampling effort to detect these changes are derived. The influence of type of measurement, magnitude of change and level of power on sampling effort required to detect simulated temporal change is investigated. Possible reasons for the interaction of sampling intensity with spatial pattern and other factors are discussed in relation to spatial variability at the site. The sampling effort to detect a change in condition category using qualitative monitoring is also carried out. Recommendations for quantitative and qualitative monitoring are made.

5.2. Methods

5.2.1 Data collection and collation

Quantitative sampling

The study site is Hafod y Llan (SH6252, altitude 100 m-1085 m), a 1043 ha upland estate which is located in Snowdonia National Park (SNP) in North Wales and forms part of the Eryri Special Area of Conservation (SAC), Eryri Site of Special Scientific Interest (SSSI) and Eryri National Nature Reserve (NNR). The site is on the south facing side of the valley and has a cool, oceanic climate with average rainfalls of 2400mm pa recorded (Williams 2004, *pers. comm.*). Soils are predominantly acidic podsols and there are areas of exposed slate bedrock. Habitats across the site range from woodlands in the lower valley to a mosaic of grassland and wet/dry heath habitats on the open moor through to juniper heath towards the summit of Snowdon; classifications and priority status for the main habitats are given in Table 5.1.

 Table 5.1 Classification of habitats across Hafod y Llan, Snowdonia National Park in terms of National Vegetation Classification (NVC) and prioritisation at UK and European levels.

National Vegetation Classification (NVC) community	UK level priority? (BAP)	European level priority? (Annex I)
W17 Quercus petraea-Betula pubescens-Dicranum majus woodland	Atlantic oak woodland	Old sessile oak woodlands with <i>llex</i> and <i>Blechnum</i> (91AO)
W8 Fraxinus excelsior-Acer campestre- Mercurialis perennis woodland	Mixed ash woodland	No
H10 Calluna vulgaris-Erica cinerea heath	Subalpine dry dwarf- shrub heath	European dry heaths (H4030)
H12 Calluna vulgaris-Vaccinium myrtillus heath	"	u
H18 Vaccinium myrtillus-Deschampsia flexuosa heath	ű	u
M15 <i>Scirpus cespitosus-Erica tetralix</i> wet heath	Wet heath	Northern Atlantic wet heaths with <i>Erica tetralix</i> (H4010)
H15 Calluna vulgaris-Juniperus communis ssp. nana heath	Alpine juniper heath	Alpine and boreal heaths (H4060)
M17 Scirpus cespitosus-Eriophorum vaginatum blanket mire	Blanket bog	Blanket bogs (H7130)
M25 <i>Molinia caerulea-Potentilla erecta</i> mire	No	No
U4 Festuca ovina-Agrostis capillaris- Galium saxatile grassland	No	No
U5 Nardus stricta-Galium saxatile grassland	No	No
U21 Cryptogramma crispa-Deschampsia flexuosa	Siliceous scree	Siliceous scree of the montane (H8110)

In 1989, Hafod y Llan was bought by the National Trust (a large UK environmental charity) who are using it as a flagship site to demonstrate sustainable farming such as keeping the traditional Welsh hill farming practice of hafod a hendre (summer grazing on the mountain at high altitude and winter grazing lower in the valley), whilst retaining and enhancing the conservation interest of the site (NT, 2000). When they bought it, the long history of sheep grazing had degraded the conservation status of the habitats, particularly the areas of heath, through the preferential grazing of heath species and bilberry. It is a useful site to explore monitoring methodologies as management is in place in order to restore the habitats from 'unfavourable' to 'favourable' condition (NT, 2004). Since 2002, the National Trust have halved the number of ewes to 1550 in order to allow the heath habitats to recover and introduced 30 welsh black cattle to manage the spread of competitive species unpalatable to sheep such as *Molinia caerulea* (purple moor grass) and *Nardus stricta*.

In summer 2006, a grid of 400 m spacing with a random start point was created across Hafod y Llan and 5 m x 5 m sample plots were established at each grid intersection; each sample plot was divided into 25 1 m x 1 m sub-plots and measurements were taken in each of these sub-plots (Appendix 5.1). In summer 2007, two areas of 100 ha were chosen to be contrasting in their level of spatial variation in species composition and structure based on visual exploration and analysis of data from the 2006 plots. One area is homogenous acid grassland and the other is a heterogeneous mosaic of dry heath, wet heath, blanket bog, bracken, mire grassland and acid grassland (Fig. 5.2).

In each 100 ha area a grid of 200 m spacing was created and a 2 m x 2 m sample plot was established at each grid intersection; each sample plot was divided into sixteen 50 cm x 50 cm sub-plots and measurements taken in each of these sub-plots. The size of plot was changed from 5 m x 5 m to 2 m x 2 m following exploration of the data from 2006; plots were designed to ensure comparability of data. A grid of 100 m spacing was established over the central 25 ha of each of the two 100 ha areas and 50 sample plots established at each intersection as for the 200 m grid (the 100 m grid in the homogenous area had to be offset in order to avoid old quarry workings). Additional 2 m x 2 m plots of 50 m spacing were added to the central 6.5 ha of the two 100 m grids and subsequently 2 m x 2 m plots at 25 m spacing were

added to the central 1.56 ha of the 100 m grids; bad weather limited the number of plots at 50 m and 25 m spacing which could be recorded. Plots which fell on the same sample point as a plot from a different sized grid were not re-recorded. This resulted in a total of 164 plots recorded and when plots from one grid scale were substituted for a plot in a different grid scale, this produced 62 plots at 400 m spacing, 50 plots at 200 m spacing, 50 plots at 200 m spacing (Fig. 5.2).



Figure 5.2 Layout of sample plots in Hafod y Llan, Snowdonia National Park in 2006/7.

A regular, systematic grid was chosen for several reasons; no previous knowledge of the spatial distribution of the data are required, it provides an unbiased estimation of the parameters of interest, because each plot in the geographic distribution area has a *priori* the same probability of being included in the sample and it ensures that sampling is evenly spaced across the study area (Fortin *et al.*, 1989). Practitioners claim that an inventory with a systematic design is resource efficient since plots are fast to establish and measure because their locations are based on fixed bearings and distances (Wong *et al.*, 2005). Although regular grids will give biased results if the spacing happens to coincide with periodic patterns in the habitat, such patterns are rare in semi-natural habitats such as this site (Brown, 2001).

Species inventory was restricted to those likely to increase/decrease in response to grazing management at the site, which aims to increase the extent and improve the condition of existing areas of UK/European priority habitats present on the site (Table 5.1). Lists of species were drawn up according to the NVC communities comprising the priority habitats (Rodwell, 1991b, 1992) and species present on the site noted, except for the woodland habitats as no plots were located in the woodland areas. From these lists, species were chosen to be large species such as *Calluna vulgaris* (ling heath), *Erica cinerea* (bell heath) and *Erica tetralix* (cross-leaved heath) and/or groups of species forming dominant components of the communities (Table 5.2). This is partly because they are key structural components of these communities and partly because they are easy to identify. In addition, species known to be increasing in distribution and abundance at the site (NT, 2004) were recorded; *Molinia caerulea* (purple moor grass), *Nardus stricta* (mat grass), *Juncus effusus* (soft rush), *Juncus squarrosus* (heath rush) and *Pteridium aquilinum* (bracken).

(i) Heath-forming species	(ii) Grasses, sedges and ferns	(ii) Bryophytes		
Calluna vulgaris (ling heath)	Sedges (all species)	Carpet forming		
Erica cinerea (bell heath)	Grasses (all species excluding M.	bryophytes (all species)		
<i>Erica tetralix</i> (cross-leaved heath)	Caerulea and N. stricta)			
	Molinia caerulea (purple moor grass)	Polytrichum spp.		
Vaccinium myrtillus (bilberry)	Nardus stricta (mat grass)	Sphagnum spp.		
<i>Juniperis communis nana</i> (prostrate juniper) <i>Ulex gallii</i> (western gorse)	Juncus effusus (soft rush)			
	Juncus squarrosus (heath rush)			
	Pteridium aquilinum (bracken)			
	Blechnum spicant (hard fern)			
	Cryptogramma crispa (parsley fern)			

Table 5.2 Taxa (species or groups of species) measured at sample plots across a grid at Hafod y Llan, Snowdonia National Park in 2006/7. Taxa in bold are those used in analysis.

This resulted in species in three groups of taxa: (i) heath-forming species; (ii) grasses, sedges and ferns and (iii) bryophytes (Table 5.2). Species known to be difficult to identify, such as 'sedges', 'grasses' and sphagnum mosses were recorded as composite groups to reduce sampling time and eliminate problems of misidentification of species (Scott and Hallam, 2003). Two further groups of bryophytes were recorded, *Polytrichum* mosses and carpet-forming bryophytes which encompasses any mosses or liverworts which are low growing and form a thin mat over the ground, including *Hylocomium splendens, Hypnum jutlandicum, Pleurozium schreberi* and *Rhytidiadelphus loreus/squarrosus*.

Presence/absence of these taxa were noted in each sub-plot and height recorded in each sub-plot by using a plastic ruler to measure height of each taxon once in the middle of each sub-plot, or of the individual/patch nearest to the centre (Appendix 5.1). Any plants inside the boundary of the sub-plot were included. In addition, a number of environmental variables were recorded at each sample point in order to explore their relationships to the vegetation. Altitude, aspect and slope angle were recorded at each sample point (Table 5.3); since Hafod y Llan encompasses a wide range of each of these three variables they are likely to have a considerable impact on the plant communities. Furthermore, frequency of grazing, bare rock and bare soil were recorded in each sub-plot to allow species response to these factors to be explored (Table 5.3, Appendix 5.1).

Environmental variable	Details (recorded at sample point unless stated)
Altitude	Extracted from the 10 m resolution Digital Elevation Model (Landform Profile DTM 1: 1000 ¹)
Aspect	Measured in degrees using a compass
Slope	Measured in degrees using a clinometers
Frequency bare rock	Proportion of sub-plots with bare rock present
Frequency bare soil	Proportion of sub-plots with bare soil present
Frequency of grazing	Sum of proportion of sub-plots with evidence of grazing/browsing and proportion of sub-plots with evidence of dung

Table 5.3 Environmental variables collected at sample points and in plots in Hafod y Llan, SnowdoniaNational Park in 2006/7, for details of measurements see Appendix 5.1.

¹ Crown copyright/database right 2008. An Ordnance Survey/EDINA supplied service

Data were made comparable from all plots on the grid by using measurements from the four 1 m² sub-plots nearest to the south-western corner of the 5 m x 5 m plots and measurements for all sub-plots from the 2 m x 2 m plots (since both areas are 4 m²). Species composition (presence/absence of taxa in 4 m²) and vegetation structure (average height of taxa in 4 m²) were calculated at the plot level. Plots from the 400 m grid were used to create grids of greater spacing; 3 plots at 1600 m spacing and 16 plots at 800 m spacing. Data analyses such as ordination tend to be sensitive to the presence of species that occur only at a few locations, which means that rare species may have a large influence on the analysis (ter Braak, 1995; Legendre and Gallagher, 2001). Therefore taxa that occurred at less than ten sample plots in the data set were excluded from the analysis, which resulted in 14 taxa (Table 5.2). Absences were treated as equal to zero height throughout the analysis.

Qualitative monitoring

In summer 2008, additional condition monitoring using Common Standards Monitoring (CSM) methodology (JNCC, 2005b) was carried out in approximately 5 ha of degraded wet heath in the south west of Hafod y Llan. The same 5 ha area had been sampled in 2003 by ecologists from the National Trust using the CSM Guidance for Upland Habitats amended for the Eryri SAC (NT, 2004; Ardeshir, 2005; JNCC, 2005a). The same sampling methodology and criteria were used again in 2008, which involved establishing 55 temporary sample plots of 1 m radius at 15 m intervals over a regular grid. According to the CSM guidance, for the wet heath habitat to be in favourable condition, at least 70% of sample plots must reach a series of predetermined targets (Appendix 5.2). In both 2003 and 2008 the area was found to be in 'unfavourable condition'; in 2003, 4% of sample points passed and in 2008, 25% of sample points passed. Sample points mainly failed due to low levels of positive indicator species and high levels of *Molinia caerulea* (purple moor grass), a negative indicator.

5.2.2 Data analysis

Spatial pattern

Individual taxa

The first objective was to assess the spatial variability in the dataset through geostatistics. Mean height data as an indication of abundance was used to explore spatial pattern in Genstat (Genstat, 2005). Geostatistical analysis such as kriging must be based on a minimum of approximately 100 data points (Webster et al., 1992; Frogbrook, 2000; Fortin and Dale, 2005). The dataset in this study contains 164 plots, with a maximum of 64 at any single sampling intensity, therefore spatial pattern could not be analysed separately at each sampling intensity and instead plots from all sampling intensities were combined for spatial analysis. Geostatical analysis also requires the data to be normally or near normally distributed (Webster and Oliver, 2001), and so descriptive statistics were calculated, primarily the skewness value, which indicates the extent to which a given property deviates from a normal distribution. There are no set values to indicate when data are skewed, but a rule of thumb is to regard data with a skewness value of greater than 1 as positively skewed and data with a value less than -1 as negatively skewed (Webster and Oliver, 2001). Histograms and box-plots were also plotted to allow the statistical distribution to be examined. Data were positively skewed for all taxa due to absences and many low height values. Therefore the scale of measurement was transformed to produce a new scale with a distribution closer to normal. For this, a constant of 0.1 was added to all values in order to remove zeros from the data and log10 was used. Only taxa with a skewness value of between -1.2 and 1.2 in their transformed data were investigated for spatial structure. Data were also checked for significant spatial trend, which is when there is a trend with say increase in the mean in either a east/west or north/south direction (Cressie, 1993). There is no set value for defining the presence of trend in the data and a threshold value, where trend accounts for no more than 20% of the variation, was applied and taxa over this threshold excluded from further analysis (Frogbrook, 2000). There were insufficient data to check for anisotropy which is when variation varies with direction or zones (Zimmerman, 1993), so all data were assumed to be isotropic (Frogbrook, 2000).

The next stage was to compute semivariances and form experimental variograms for each taxa, choosing the lag distance and the maximum lag interval to produce the variogram which best represented the spatial structure of each taxa. The lag distance generally corresponds to the smallest sampling interval and the maximum lag interval represents the distance before which the number of comparisons starts to decrease markedly or where the variogram becomes eratic (Frogbrook *et al.*, 2002). A range of models were fitted to the semivariances for each taxa (power function, circular, spherical, penta-spherical and exponential) and the best fitting model chosen as the one that minimises the residual sum of squares (Webster and Oliver, 2001).

Using the appropriate variogram model parameters and data, values were predicted at 50 m intervals on a square grid by ordinary kriging. Each prediction was made using a minimum of three and a maximum of 20 data points (Webster and Oliver, 2001). The search radius was equal to the range of spatial dependence as all of the models were bounded. These predictions were then back transformed, contoured and mapped in MapInfo (MapInfo, 2004). The kriging variances were computed and back transformed at the same time as the kriged predictions and these were also mapped; these are guides to the reliability of the predictions and will show if there are areas where sampling should be increased to improve the predictions (Frogbrook, 2000).

Community

Information on the spatial pattern of individual taxa does not capture the essential features of the spatial pattern of the whole community (Dale, 1999), and so spatial pattern at the community level was also investigated. There are ways of combining standard analyses of plant communities through multi-scale ordination which partitions the variance produced in an ordination according to spatial intensity of the scales sampled across (Wagner, 2004). However, this requires at least 100 plots at each intensity (Fortin and Dale 2005) and so could not be used with this dataset. Instead ordination analysis was performed using a matrix of site by species' mean height data for all taxon recorded in more than ten plots (Table 2, section 2.1.1). Data were entered into PCOrd (McCune *et al.*, 2002) and ordinated using indirect gradient analysis techniques (IGA), where multi-species samples are ordered based on their

overall species composition. Principle Components Analysis (PCA) and Detrended Correspondence Analysis (DCA) were used to summarise the species x object data into a single metric per sample point which would provide an indication of variability within the dataset. Both PCA and DCA aim to construct hypothetical variables that give the 'best fit' to the data according to an assumed linear or unimodal response model (Birks, 2008). PCA provides the solution for linear responses and assumes that there is a normal multivariate distribution in the data whereas DCA provides the solution for unimodal responses and does not allow the means and variance to be independent (VanGroenewoud, 1992; Palmer, 2009). Both techniques were performed in order to check if one was more suitable (explained more of the variation in the dataset) and the correlation of PCA and DCA sample point scores on axes 1 and 2) were checked using a Spearman's Rank two-tailed test in SPSS (SPSS, 2003). The sample point scores for PCA and DCA were strongly correlated; correlation coefficient = -0.925 (p<0.001) for axis 1 and -0.936 (p<0.001) for axis 2. DCA was chosen for the remainder of the analyses since it copes better with species data showing non-monotonic responses.

Axes 1 and 2 of the DCA ordination of the sample point scores explained the majority of the variance for both presence/absence (axis 1 = 44.1%, axis 2 = 35.8% and axis 3 = 20%) and height data (axis 1 = 58.6%, axis 2 = 27.2% and axis 3 = 14.3%). Because the analysis aims to capture total variation amongst the dataset, no single axis was most important, therefore plot scores for the two axes explaining the majority of the variation (axes 1 and 2) were combined using the following formula:

($(\sum axis 1 \text{ scores } x \text{ eigenvalue for } axis 1) + (\sum axis 2 \text{ scores } x \text{ eigenvalue for } axis 2))$

 $((\sum axis 1 + axis 2 \text{ scores}) x (\sum eigenvalues for axes 1 and 2))$

[2]

Since the axes scores are independent (and may therefore represent different aspects of variation in the dataset), it was important to carry out sensitivity analysis. Spearman's Rank correlation was carried out between axis 1 scores and the combined scores and axis 2 scores and the combined scores in SPSS (SPSS 2003). The sample point scores for axis 1 and axes 1 and 2 combined and axis 2 and axes 1 and 2 combined were strongly correlated for both presence/absence (correlation)

coefficient = 0.634 (p<0.001) and -0.660 (p<0.001) respectively) and height data (correlation coefficient = 0.742 (p<0.001), and -0.692 (p<0.001) respectively). This implies that the variation resulting from the combination of axes 1 and 2 reflects the variation in each axis independently. These combined DCA scores for the presence/absence and mean height data were used in several analyses described below. The combined scores for mean height were interpolated in MapInfo. Height data was chosen rather than presence/absence for this analysis in order to show more variation in measurements. The Inverse Distance Weights option was chosen (where the interpolated value is the weighted means of the neighbouring measured values).because as in kriging, it assumes that the value at the location of interest will be more similar to neighbouring points than to those further away.

Combined axis 1 and 2 ordination scores of height data were also used to test for differences in variance and mean of the 400 m sample points between the subjectively chosen spatially homogenous and spatially heterogeneous areas using an independent t-test in SPSS (SPSS, 2003). Following this, ordination scores for sample plots from the homogenous and heterogeneous areas were each partitioned into 400 m, 200 m and 100 m spacing and the variance (standard deviation) compared.

Relationship to measured environmental variables

To investigate the relationships between the measured environmental variables and the plant community structure, Canonical Correspondence Analysis (CCA) was performed in PcOrd (McCune *et al.*, 2002) using all 14 taxa in the main matrix and the six environmental variables in the second matrix (Tables 5.2 and 5.3). In order to make the measurements of aspect meaningful, the angle in degrees was converted to radians and cosine transformed to an index of exposure where due north is -1 and due south is 1 (south is given the maximum value as it is assumed to be receive maximum sun). CCA is a direct gradient analysis in which species composition is directly and immediately related to measured environmental variables (Palmer, 1993). In CCA, the correlation between species scores and sample scores is maximised, whilst the sample scores are constrained to be linear combinations of explanatory environmental variables (Jongman *et al.*, 1995; Palmer, 2009). It is

based on the assumption that species data is a unimodal function of position along environmental gradients (ter Braak, 1995). The default option for standardising column and row scores was chosen (centering and normalizing, where site scores are rescaled such that the mean is zero and the variance is one). For the scaling of ordination scores, the optimising columns (species) option was chosen, and for the graphing of scores, the plot scores as linear combinations of environmental variables was chosen.

Investigation of power to detect temporal change

The combined scores from the DCA ordination of presence/absence and mean height data (as described previously) were used to calculate mean and variance (standard deviation, SD) for sample points at each successive sampling intensity; this is the baseline from which possible temporal changes were modelled. The mean of the baseline data for each sampling intensity was increased in 2% increments from 2% to 20% and 20% increments from 10% to 130%. Each new set of means was a modelled repeat survey of vegetation indicating a change in the vegetation (assuming that the same sample points were re-sampled and assuming that amount of change was consistent between sample points). The baseline and modelled repeat group means were entered into a standard effect size calculator (Network, 2008) along with sample sizes and variance in order to generate effect size statistics for the differences between the baseline and repeat (Glass, 1977; Kirk, 2007). The variance for the control group is used as the variance for the repeat which assumes that the modelled change data had the same variance as the baseline data and thus that any differences are due to sampling variation (McGaw and Glass, 1980). The analysis also assumes a re-sampling of the same sample points (permanent sample plot) approach, using a re-allocation approach will normally require more samples to achieve a given power (Green, 1989). Assuming a Type I error of 0.05 and a power of 0.8 (since these are standard in conservation monitoring), sample sizes required to detect the changes were calculated using Gpower software (Faul et al., 2007). Sample sizes were converted into days of effort according to how many sample points could be recorded in a day at each sampling intensity (Table 5.4). Following this, the process of calculation of effect size and effort was carried out for

presence/absence data from sample points in a) the spatially homogenous and b) the spatially heterogeneous areas.

 Table 5.4 Number of sample points recorded per day at various sampling intensities; based on field methodology used in Hafod y Llan, Snowdonia National Park in 2006/7.

Sampling intensity (distance between sample points, in meters)	Number of sample points recorded per day			
1600	10			
800	15			
400	20			
200	25			
100	30			
50	50			
25	50			

The quality of information collected at each sampling intensity was explored using the Habitats of Wales Phase I data survey 1979-1997 (CCW, 2005a). The sample points in this study were overlaid in MapInfo on the Phase I data and the Phase I habitat at each point noted. The sample points were then re-sampled ten times at each intensity using the appropriate number of samples which could be recorded in a day (Table 5.4). A random start point for each 'day' of sampling was chosen for each of the ten 'runs' of re-sampling. The average number of Phase I habitats recorded in a day was then calculated.

Qualitative monitoring

In order to investigate the amount of effort to detect a shift in the wet heath habitat from the 2003 and 2008 'unfavourable condition' assessments to a favourable condition, change had to be predicted. This involved manipulating the 2008 data by increasing cover of positive indicator species such as *Drosera* spp. (sundew) and decreasing cover of negative species such as *M. caerulea*, until 70% of sample points passed the criteria and the area would thus be scored as being in 'favourable' condition. This represents the changes which the management regime of reduction in sheep numbers and introduction of cattle grazing are intended to bring about to this area of wet heath (NT, 2004). This resulted in data for three time points (2003, 2008 and 'favourable') for which mean and variance (standard deviation, SD) was

calculated. Effect sizes were then generated in a standard effect size calculator (Network, 2008) for the change in the mean and variance between i) 2003 and 2008, ii) 2003 and 'favourable' and iii) 2008 and 'favourable'. These effect sizes were used to calculate sample sizes required to detect the three sets of change using Gpower software, assuming a Type I error of 0.05 and a power of 0.8 (Faul *et al.*, 2007). This analysis used a re-allocation of sample point approach and assumed that amount of change was consistent between sample points.

5.3. Results

5.3.1 Spatial pattern of vegetation and relationship to environmental variables

Individual taxa spatial pattern

Transformation of taxa data resulted in normal distributions for six taxa, which were analysed for spatial structure (Table 5.5). None of the data showed 'significant spatial trend' (trend which accounts for >20% of the variation in the data). The experimental variograms were computed using a range of lag distances between 200 m and 400 m (Table 5.5, Fig. 5.3).

Table 5.5 Information about taxa: percentage of sample points found in, skewness, lag distance used to compute variogram and model parameters. Data collected from Hafod y Llan, Snowdonia National Park in 2006/7.

ΤΑΧΑ	% sample plots recorded in	Lag distance (m)	Skewness of data	MODEL	Parameters		
					C ₀ ¹	C²	a³
Calluna vulgaris	43	400	0.4476	circular	0.4078	0.4719	1101.3
Molinia caerulea	58	200	1.0760	circular	0.3622	0.7510	601.9
Nardus stricta	64	400	-0.3827	circular	0.1613	0.6144	680.5
Vaccinium myrtillus	32	250	1.1609	circular	0.2000	0.4202	1327.0
All sedges	72	200	-0.6016	circular	0.4267	0.4837	1178.3
All carpet forming bryophytes	57	400	0.3596	circular	0.3987	0.4834	1243.0

¹ c₀ is the nugget variance, ² c is the sill of the variance and ³ a is the range of the spatial dependence.

The variograms for all six taxa are bounded (Fig 5.1b), which means that the spatial variation at this level of investigation has been encompassed by the sampling. The range of spatial dependence is shortest for *M. caerulea* and *N. stricta* at 601.9 m and 680.5 m respectively and longest for *V. myrtillus* at 1327.0 m. The variance that is unresolved is reflected in the nugget variance, which accounts for between 25% (*N. stricta*) and 80% (carpet forming bryophytes) of total variance (Fig. 5.3).

The maps of kriged predictions are drawn across the whole grid (Fig. 5.4). The maps all show a patchy distribution across the area, with *C. vulgaris* and *V. myrtillus* having patches of highest height (>7 cm) towards the north of the area. *M. caerulea* has highest values (>16 cm) in the south, and *N. stricta* has several patches with high values (>16 cm) in the north east. The sedges have patches of high values (>8 cm) in the north and the west and carpet forming bryophytes have low values (<0.5 cm) in the middle of the site which increase outwards.

Calluna vulgaris

Vaccinium myrtillus



Figure 5.3 Experimental variograms and fitted models for height of six taxa sampled in Hafod y Llan, Snowdonia National Park in 2006/7. Symbols are the experimental semivariances and the line is the fitted model.




Molinia caerulea



All sedges



Vaccinium myrtillus







All carpet forming bryophytes



Figure 5.4 Maps of kriged predictions for height in cm for six taxa sampled in Hafod y Llan, Snowdonia National Park in 2006/7. Black outline shows site boundary.

The pattern of kriging variance is similar for all six taxa, so only *C. vulgaris* is shown (Fig 5.5). Kriging variances are smallest in the most intensively sampled areas in the west and south of the site, and largest at the edge of the site where there are fewer samples from which to predict. The patch of high values to the north-east of centre is the result of two missing values in the dataset, due to inaccessible sample points.



Figure 5.5 Map of kriging variance for *C. vulgaris*; data collected in Hafod y Llan, Snowdonia National Park in 2006/7. Black outline shows site boundary.

Multi-taxa spatial pattern

Interpolation of combined ordination scores was carried out and shows an overall pattern of high values at lower altitudes along the south east boundary of the site, moving to lower values at higher altitudes (Fig 5.6).



Figure 5.6 Map of interpolated axis 1 and 2 sample point scores from a DCA ordination of all height data. Crosses show sample points at 400 m, 200 m and 100 m spacing (50 m and 25 m not shown), black outline is the site boundary and black lines within site show compartment boundaries. Data collected in Hafod y Llan, Snowdonia National Park in 2006/7.

The patches of lowest variation just north of the centre match up with the areas with highest kriged predictions for heights of *C. vulgaris* and *V. myrtilllus*. These areas are dominated by dwarf shrubs which are likely to out-compete the other vascular and non-vascular taxa comprising the dataset.

The areas selected for sampling as contrasting in variation of vegetation had significant differences in the variances of their combined axis 1 and 2 DCA scores based on height data (p<0.001), indicating that the two sample areas had not been drawn from populations with equal or similar variance. The mean values were also significantly different (p<0.001), with the spatially 'homogenous' area having a lower mean combined axis 1 and 2 score (155) than the spatially 'heterogeneous' area (235). Looking at the estimated variance (standard deviation) for combined axes 1 and 2 DCA ordination scores, the variance of the sample points at 400 m intensity is lower in the homogenous area than in the heterogeneous area, whereas the variance of the sample points at 100 m is far higher in the homogenous area than in the heterogeneous area than in the homogeneous area than in the homogeneous area than in the homogeneous area than in the heterogeneous area than in the homogeneous area than in the heterogeneous area than in the homogeneous area t



Figure 5.7 Variance in combined axes 1 and 2 DCA ordination scores and sampling intensities (400 m, 200 m, 100 m distance between points) using presence/absence data from a spatially homogeneous area and a spatially heterogeneous areas in Hafod y Llan, Snowdonia National Park in 2006/7. There are insufficient sampling locations to include sampling intensities of 1600 m, 800 m, 50 m and 25 m.

Relationship to environmental variables

The CCA ordination scores for species and bi-plots scores for environmental variables were used to pick out the main relationships between the 14 taxa height data and six environmental variables (Fig. 5.8). The environmental variables are represented as lines radiating from the centroid of the ordination, the longer the environmental line, the stronger the relationship of that variable with the community (ter Braak, 1995) The position of species points relative to the environmental lines can be used to interpret their relationships. Only environmental variables with a correlation of at least 0.2 with either axis are included in the ordination graph in order to eliminate uninformative environmental variables.

The proportion of variation in the species data explained by the measured environmental variables is low; axis 1 explains 10.9%, axis 2 7.2% and axis 3 2.3% (Table 5.6), thus there are other physical, chemical soil or hydrological variables which may explain more of the variation, but which were not measured in this study. However, correlations show that the first axis is primarily an altitude gradient (-0.67) and the second axis primarily a slope gradient (0.6) (Table 5.6).

	Axis 1	Axis 2	Axis 3
Percentage variance of taxa data explained	10.9	7.2	3.2
Intraset correlations			
Altitude	-0.67	0.06	0.09
Aspect	0.05	-0.17	-0.05
Slope	0.01	0.6	-0.17
Frequency bare rock	0.1	-0.07	-0.21
Frequency bare soil	-0.13	0.37	0.2
Frequency of grazing	-0.35	-0.32	-0.15

Table 5.6 Percentage of variation explained and intraset correlations (correlation between the ordination axes and environmental variables) from the Canonical Correspondence Analysis of taxa height data and environmental variables collected at Hafod y Llan, Snowdonia National Park in 2006-7.











Figure 5.8 Canonical Correspondence Analysis of taxa height data and six environmental variables collected at Hafod y Llan, Snowdonia National Park in 2006-7. Ordination diagrams of: (a) Axis 1 and 2; (b) Axis 1 and 3 and (c) Axis 2 and 3. Environmental variables are represented by lines and taxa by crosses (sample points are omitted for ease of interpretation).

(a)

The ordination shows that height of C. vulgaris and E. cinerea are strongly positively associated with slope angle and frequency of bare rock on axes 1 and 2, and closer inspection of the patches of high height values show that they coincide with steep slopes and cliff areas with high frequencies of exposed rock (Fig. 5.4, Fig. 5.8). Height of V. myrtillus is most strongly positively associated with altitude on axes 1 and 2, and rock frequency on axis 3 which also explains the patches of high value kriged predictions for V. myrtillus which are in the areas above 600 m (Fig. 5.4). Height of *N. stricta* and the sedges are both positively associated with altitude (axis 1) and grazing (axis 1 and for sedges also 2 and 3), which explains their appearance at higher elevations. There seems to be some evidence of positive association of Sphagnum spp. with grazing on all axes and Polytrichum spp. and grazing for axes 1 and 2. There are also negative associations; *Pteridium aguilinum* is negatively associated with altitude on axis 1 and both Juncus spp. are negatively associated with slope on axis 2. Finally, the carpet forming bryophytes and all grasses are not strongly associated with any of the measures of environmental variables, suggesting that either the sampling did not encompass the variables most closely associated with them, or that the contrasting environmental responses of the individual species cancelled each other out.

5.3.2 Detection of change

Simulated temporal change

There is no consistent relationship between number of samples and effort (number of days) to detect a change and sampling intensity. Instead, number of samples/days required peaks at a sampling intensity of 100 m, dropping sharply as sampling intensity is both decreased and increased (Fig. 5.9). Presence/absence and height data both have a similar relationship between sampling intensity and number of samples/days, although their magnitude is much higher using height measurements. For instance, to collect sufficient data at a sampling intensity of 100 m to detect a change of 10%, it would take 1472 samples (49 days) using presence/absence data compared with 45000 samples (1500 days) using height data (Fig. 5.9).



Figure 5.9 Sample size required to detect change over a range of 10-90% across different sample intensities (1600 m, 800 m, 400 m, 200 m, 100 m, 50 m, and 25 m distance between points) using combined axes 1 and 2 scores from DCA ordination of (a) presence/absence sample size and (b) vegetation height data. Sample size is converted into number of days of sampling for (c) presence/absence and (d) vegetation height. Data collected in Snowdonia National Park in 2006/7.

Data from sampling in the two areas differing in spatial variability display very different interactions to each other. Data from the homogenous area follows a similar relationship to that seen using data from all sample points in that number of samples and effort (number of days) is much higher at 100 m sampling intensity compared with 200 m and 400 m (Fig. 5.10). Data from the heterogeneous area shows a very different pattern, with the largest number of samples/highest effort (number of days) required to detect a change using data from the lowest sampling intensity of 400 m and decreasing with higher intensity.



Figure 5.10 Sample size required to detect change over a range of 10-90% across different sample intensities (400 m, 200 m and 100 m distance between points) using combined axes 1 and 2 scores from DCA ordination of presence/absence data for (a) homogenous and (b) heterogeneous area. Sample size is converted into number of days of sampling for (c) homogenous and (d) heterogeneous area. There are insufficient sampling points to include sampling intensities of 1600 m, 800 m, 50 m and 25 m.

The sample size and effort (number of days) required to detect a simulated temporal change of over 30% in combined axes 1 and 2 scores of presence/absence data at all sampling intensities is very low (Fig. 5.9). Amount of samples and effort measured in days increases as magnitude of simulated temporal change decreases; the smaller the change the more samples/sampling effort is required. Sample size and effort required increases particularly sharply to detect a change below 10%; for example a sampling intensity of 50 m requires 37 samples (1 day) to detect a change of 10%, 121 samples (2.5 days) to detect a change of 6% and 942 samples (19 days) to detect a change of 2% (Fig. 5.11).



Figure 5.11 Sample size required to detect change over a range of 10-90% at a sampling intensity of (a) 50 m and (b) 400 m distance between points (chosen to be illustrative of the relationship across sampling intensities) using combined axes 1 and 2 scores from DCA ordination of presence/absence. Sample size converted into number of days of sampling is also shown for (c) 50 m and (d) 400 m. Data collected at Hafod y Llan in Snowdonia National Park in 2006/7.

Increasing the power also has a substantial effect on number of samples and effort (days) required to detect a given level of change; the higher the power the larger the samples/effort required, with the samples/effort increasing sharply over a power of 0.9 (Fig. 5.12). This is particularly noticeable for the 100 m sampling intensity, with other intensities following the same pattern but with a lower magnitude.

(a)

(b)



Figure 5.12 (a) Sample size and (b) sample size converted into number of days required to achieve a given power to detect a 20% change at different sample intensities (1600 m, 800 m, 400 m, 200 m, 100 m, 50 m, and 25 m distance between points) using combined axes 1 and 2 scores from DCA ordination of presence/absence data recorded at sample points in Hafod y Llan in Snowdonia National Park in 2006/7.

172

The quality of information collected at different sampling intensities also needs to be balanced against sample size and effort. Using the mean number of Phase I Land Cover types which could potentially be recorded per day, the plots at 100 m intensity record the highest mean number of habitat types at 4.1 per day (and the largest standard error) while the 25 m plots record the lowest at 1.8 per day (Fig. 5.13).



Figure 5.13 Mean number of Phase I land cover types recorded per day according to re-sampling points from different sample intensities (1600 m, 800 m, 400 m, 200 m, 100 m, 50 m, and 25 m distance between points). Phase I data recorded during Habitats of Wales Phase I survey 1979-1997, data across various sampling intensities recorded in Hafod y Llan in Snowdonia National Park in 2006/7. Bars show +/- 1 SE.

Qualitative monitoring

The effect size for the change recorded in the 'unfavourable' wet heath between 2003 and 2008 is 0.1, classed as 'small' (Glass, 1977; Cohen, 1988). Using the data manipulated to make the wet heath into 'favourable' condition, the effect size between 2003 and 'favourable' is 0.35 and between 2008 and 'favourable' is 0.19 (these are still 'small' effect sizes according to Glass (1977)). Concentrating on the change between the 'unfavourable' wet heath in 2008 and the manipulated 'favourable' wet heath, the power to detect that change using 55 sample points is only 0.26. If a Type I error rate of 0.05 is assumed and sample size converted to effort based on 55 sample points taking one day, achieving a power of 0.7 will require

262 samples (equivalent to 5 days) and a power of 0.8 will require 201 samples (6 days) (Fig. 5.14).



Figure 5.14 (a) Sample size and (b) number of days required to detect a change from unfavourable condition wet heath in 2008 to favourable condition wet heath using a range of power levels (given a Type I error rate of 0.05). This is based on data collected in Hafod y Llan, Snowdonia National Park in 2008.

174

5.4 Discussion

The spatial pattern of individual taxa and the plant community (based on a single metric) were successfully described through prediction using data from a quantitative and systematic sampling strategy. The kriged distributions of height of *C. vulgaris, V. myrtillus, M. caerulea* and *N. stricta* match their ecological preferences mediated by management actions. Despite the recent reduction in grazing, the species' responses to the long history of grazing at the site are evident with *C. vulgaris* and *V. myrtillus* confined to steep rocky slopes at high altitudes in the middle and north of the site which are less accessible to sheep. These species are palatable to sheep and decline under heavy grazing (Jones, 1967; Anderson and Yalden, 1981; Pakeman *et al.*, 2003). Conversely, *M. caerulea* and *N. stricta* are less preferred by sheep (Berendse, 1985; Welch, 1986; Hulme *et al.*, 1999) and are tallest on the accessible gentler slopes. The tall *M. caerulea* in the south west of the site could also be due to the long term practice of burning in that compartment (NT, 2004), which is known to lead to replacement of *C. vulgaris* with *M. caerulea* (Hobbs and Gimingham, 1987).

The reduction in grazing at this site is expected to increase the heather communities and reduce the abundance of the less preferred grass species, however such high altitude heather moorland is slow growing and relatively resistant to change (Milne and Hartley, 2001). The changes are also likely to take place in a heterogeneous pattern over the site due to the spatial variability both in habitat preference and grazing intensity shown by herbivores such as sheep (Palmer *et al.*, 2004; Hartley and Mitchell, 2005). Palmer *et al.* (2004) noted the effect of positive feedback loops whereby sheep are attracted to areas with abundant palatable grasses and any heather present is replaced by grasses. Site based factors such as aspect, altitude and soil nutrients also play a role in the competitive balance between heather and grasses and these again exhibit spatial variation. The positive association found between *C. vulgaris* and slope and frequency of bare rock could indicate that steep, rocky areas are likely to exhibit a faster heather recovery from high grazing pressure.

Soil fertility also influences rate of change in heather cover in response to changes in grazing, with the competitive balance between *C. vulgaris* and grasses and *M. caerulea* and *E. cinerea* strongly affected by nitrogen levels (Berendse, 1985; Hartley and Mitchell, 2005). On Hafod y Llan, the dominant geology is rhyolite, over which

thin acid soils have developed (NT, 2000). These soils such as brown podzols, peaty gleys and peat rankers have generally low ph and low fertility. Nitrogen deposition rates in this area of Snowdonia are 20-25 kg ha⁻¹ yr⁻¹ which is a relatively high rate for the UK (Stevens *et al.*, 2004a). High nitrogen deposition in comparable low fertility upland areas has been linked to initial increases in *C. vulgaris* growth with subsequent negative long term effects through the combined effects of out-competition by grasses, climatic stress and grazing pressure (Carroll *et al.*, 1999). In addition, it has been found that high nitrogen deposition may change nitrogen limited ecosystems into phosphorous limited ones which could favour species such as *M. caerulea* that are better adapted to P limitation (Kirkham, 2001). The complexity of species interactions with each other and the environment means that shifts in species composition over time are difficult to predict.

Although the kriged distributions for individual species are ecologically meaningful, the distributions for the composite groups of all sedges and all carpet forming bryophytes are hard to interpret, which is likely to be because the responses of individual species are cancelled out. In a study manipulating nutrients and grazing, Hartley and Mitchell (1995) found that individual sedge and grass species showed contrasting responses, for example *Carex echinata* and *Carex nigra* were common in ungrazed plots whereas *Carex binervis* and *Carex panicea* were common in grazed plots.

Whilst the interpolation of ordination scores allowed the variance in the dataset to be displayed spatially, it did not appear to relate to the distribution of plant communities across the site or to environmental variables. This could be because the conversion of information about many taxa into a single metric is reductive and it is difficult to know what it means in relation to the pattern of vegetation (Legendre and Fortin, 1989). One way to overcome this could be to combine a categorical approach with spatial statistics, for instance to identify plant communities at a series of points and interpolate between. This would allow probability of occurrence to be calculated and change to be quantified. It has been noted that although these approaches (categorical and point data) are rarely combined (Gustafon, 1998), they may provide a useful way forward; in this case the disadvantages are the time involved and reliance on subjective assignation of plant community type.

All of the kriged predictions and estimation variances are dependent on the interaction of the sampling intensity with the spatial variation at the site, as previous studies have demonstrated (Frogbrook, 2000). The interaction between spatial variation and sampling intensity is not obvious for any taxa/composite group or ordination scores since the areas sampled with increasing intensity do not encompass more variation than the whole site sampled at 400 m intensity. The fact that estimation variance increases markedly as sample spacing increases means that the predications between the points in the grid spaced at 400 m apart are less reliable than those in the two areas sampled more intensively (Robertson, 1987). Investigation of predictions and estimation variance for a number of different sampling intensities has been used in determining samples for estimating soil nutrients in precision farming and minerals in mining (Isaaks and Srivastava, 1989; Oliver and Frogbrook, 1998). In these approaches, estimation variance is partitioned by sampling intensity and the most efficient regime chosen according to an acceptable threshold. This has potential for demonstrating optimum sampling regimes for vegetation monitoring, however the most precise estimates of spatial autocorrelation and thus predictions of distribution are obtained from large sample sizes (Tobin, 2004). There are insufficient sample points at each intensity to explore precision of estimates in the current study; furthermore, due to the low resources available for conservation and the large number of protected sites requiring monitoring, it is unlikely to be a feasible method.

The fact that variance in ordination scores did not increase consistently as sampling intensity increased (from 100 m to 400 m between sample points) in the areas of contrasting spatial heterogeneity implies that spatial variability is not predictable across spatial scales. This is interesting since studies of fractals in ecology assert that for many phenomena, the amount of resolvable detail is a function of scale, with increasing resolution revealing variation that previously passed unnoticed (Mandelbrot, 1983; Hastings and Sugihara, 1993; Harte *et al.*, 1999). However, other studies have also reported a lack of predictability across spatial scales and suggest that the most valuable aspect of the model of fractality may be in allowing us to measure departures from it, for example, by finding where the scaling changes or breaks down altogether (Halley *et al.*, 2004; Joseph and Possingham, 2008). For instance, landscape ecologists have shown that the connectivity of landscapes and

their sensitivity and importance of landscape pattern to key species is scale dependent, peaking at certain scales (Keitt *et al.*, 1997). Halley *et al.* (2004) recommend the development of more techniques to test for significant departures from fractal structure in order to identify and quantify the scaling properties of nature.

The lack of predictability across the specific sampling distances of 100 m and 400 m in this study is also supported in a study by Hartley et al. (2004) who found that it is difficult to predict the distribution of certain plant species across Britain between the scales of 200 m and 1000 m, and hazarded that this is due to anthropogenic disruption within these scales. They speculate that it is this range of scales over which human influence is strongest causing disruption to process and pattern; above and below this range species still exhibit their natural patterns. Although the site in the current study is composed entirely of semi-natural habitats, there is still a history of human influence through grazing management, quarrying and recreation.

The unpredictability in spatial variation across scales reveals the complex spatial pattern of the site. While spatial variation could not be explored fully due to lack of sufficient sample points at each sampling intensity, samples at 100 m distance required the largest sample size and most effort to detect the simulated changes across the whole site. Effort to detect a change can be explained by variance in the data. The higher the variance or 'noise' present in the data, the harder it is to pick out a trend, thus a larger sample size (and effort) is required (Grieg-Smith, 1957). It could be that the distance of 100 m records more of the variation in vegetation present on the site compared to the other sampling intensities. If the mosaic of patches of acid grassland, dry heath, wet heath, blanket bog and bracken are spaced at roughly 100 metres apart then neighbouring sample points at the 100 m intensity will each be located in patches of different vegetation types and the data will contain high variance and require large monitoring effort.

It would be easy to conclude from this that the 100 m sampling should be avoided in favour of the least intensive and least effort sampling intensity of 1600 m. However, this does not take the quality of information recorded at the various sampling intensities into account. While the 100 m intensity requires the most effort, it also results in the highest number of Phase I Land Cover types recorded in a day. The sample points spaced closer together than 100 m (50 m and 25 m) take less effort to detect a change, however since they frequently fall within a single vegetation type

they also provide less information (fewer Phase I habitats). Sample points further apart than 100 m (400m, 800 m and 1600 m) again take less effort to detect a change however, they also provide less information about the habitats present on the site since the distance between samples is sufficient that elements of the vegetation community are repeated (Dale, 1999).

This has important implications for the design of sampling for vegetation monitoring programmes. In this case, a sampling regime of one plot in each vegetation patch (i.e. 100 m spacing) may be the most powerful for detecting change. This will require more effort than sampling at finer or coarser resolutions but may be what is required given the spatial pattern present at the site. Clearly this is a small-scale study and findings cannot simply be scaled up to apply to whole landscapes, due to unpredictability of spatial heterogeneity which leads to allometric scaling (Hobbs, 2003). A pilot study conducted only in the spatially 'homogenous' area would provide very different assumptions about the whole site compared to pilot data from the 'heterogeneous' area. Analysis of spatial pattern at a number of scales across sites of varied vegetation and management would allow the development of optimal sampling to be explored more fully. This still requires a better understanding of the sensitivity of the metric to change at different scales as Turner *et al.* (2001) note that understanding is often limited about what level of change in any single metric constitutes an ecologically important change.

Therefore the choice of variable is an important aspect of sampling design. The larger magnitude of effort required to detect a given level of change in the height measurements is likely to be because this variable is more quantitative and has higher levels of variance due to the additional information (Stewart *et al.* 2001). The binary measurement of presence/absence of bracken has far less quantitative information than the continuous height measurements of 0-180 cm. This means that a parameter with less variation such as presence/absence rather than height will require less effort to detect a prescribed level of change but may not provide sufficient information for conservation monitoring. For instance, grazing management may successfully achieve conservation objectives through greatly reducing the dominance of bracken without actually eliminating it and early detection of a decline in height or cover of heather would be a useful warning long before it becomes absent.

Simulated levels of change to determine the sample size and sampling effort required to detect them with a given power produces valuable results. It is obvious that the smaller the change the harder it is to detect and the more samples (effort) are required. If 50% of species disappeared from a given area it would require far less effort to detect it than if only 10% were lost. However, this study underlines the fact that the effort required increases markedly at the smallest levels of change; changes below 10% are extremely hard to detect and it takes twice the effort to detect a change of 10% than a change of 20%. This shows that it is expedient to consider the level of change which is possible and necessary to observe. If it is vital to be able to detect a small change (below 20%) then other ways of increasing the effect size could be investigated. One option is to increase the sensitivity of the indicators; the best are those which closely reflect the processes of change (Legg and Nagy, 2006).

In terms of power, it is interesting that sample size and sampling effort required does not increase markedly with increase in power to detect a change up to the highest levels of power, except over 0.9. Thus in conservation management, even with limited resources, it makes sense to set the power high, (up to 0.9) as for a small amount of extra sampling effort conservation organisations will ensure that a potentially damaging change in the resource is not overlooked.

However, both the trial of the quantitative sampling across the whole site and the repeat of the qualitative monitoring of 5 ha of wet heath on the site demonstrated that achieving a sufficiently high power to detect levels of change implicit in current conservation monitoring requires unfeasible levels of effort. This study found that condition monitoring as used across the UK achieves a power of only 0.26 to detect a potential change from 'unfavourable' to 'favourable' condition; this means there is a 74% chance of missing the shift in condition which could lead to a management regime being continued for too long. This level of risk is unacceptable in the management of designated sites across the UK since it could lead to many sites being mis-managed, with serious consequences for biodiversity.

Finding a solution to this problem requires either more resources or a change in methodology. In terms of resources, a reasonable power of 0.8 requires much larger sample sizes; six staff days (to record 201 samples) is an unfeasible amount of time for a resource-poor conservation agency to devote to the monitoring of 5 ha of wet heathland. Previous studies imply that other monitoring programmes may face similar

problems; analysis of Countryside Survey data from across the UK in 2000 showed that detecting change in landscape elements with sufficient confidence (set at a coefficient of variation of 10%) would require a ten-fold increase in the current sample size of 569, up to 6700 squares (Clare and Howard, 2000). Given the limited resources of conservation agencies, resources should be targeted at producing methodologies (and refining existing ones) with the aim of increasing power. Suitable methods should measure a small amount of the right variables; those which are closely aligned to the desired change and thus more likely to demonstrate a difference.

5.4.1 Limitations to this study and recommendations for the future

The main limitation of this study is the fact that it is conducted only on a single site, although the relationships demonstrated between spatial variation and sampling effort and between change/power and sampling effort can be used to inform sampling approaches more generally. A further limitation is the predictive approach used to derive effect sizes; actual changes from year to year are unknown.

In terms of possible improvements to the sampling, several recommendations can be made:

(a) use a simpler system for sample point placement by using a known start point and pacing out subsequent sample locations along a given bearing; this would be time efficient and allow more sample points to be measured

(b) record species present in 1 m² at the sample point, again this would increase sample size

(c) avoid grouping species, instead carry out a full species inventory at each plot in order to allow more variation in species to be investigated

(d) record more samples at each intensity, particularly at the finer scales

(e) use appropriate statistical and multivariate methods to detect change, and

(f) do a power analysis to demonstrate the efficiency and capability of the chosen monitoring system.

5.5 Conclusions

This study has provided valuable answers to the questions raised about the potential of quantitative sampling to describe single and multi-taxa distributions and to detect change across various scales. The sampling strategy is objective and unbiased both in the location of sample points and in the measurements at these points, which means that the data collection is repeatable over time and space both between and within surveyors. Individual species patterns can be interpreted in relation to ecological and management factors however interpretation is limited for composite groups of species and 'communities' due to the responses of individual species cancelling each other out. The lack of predictability in spatial variation across spatial scales leads to an unpredictable relationship between sampling intensity and effort to detect simulated temporal change. Furthermore it is clear that the choice of variable has an impact on variation in data and that the use of a single metric makes interpretation of ecologically important change difficult. It is also concluded that the detection of small changes over time (equivalent to those implicit in the objectives for current monitoring programmes) require unfeasible levels of effort for widespread use.

This study makes several recommendations. Firstly, monitoring programmes must include an *a priori* definition of meaningful potential change and ensure that the variables measured are sufficiently sensitive over the monitored time period and are indicative of the desired change. Secondly, since spatial variation is site-specific, optimal use of resources can be obtained by using prior knowledge and information from a pilot study, however this requires understanding of the metric used and how it is affected by spatial scale. Approaches should also balance Type I and Type II error rates carefully according to management needs considering the costs of either missing a change or falsely identifying one. Finally it is recommended that standardised protocols are used in quantitative approaches; these should include objective sampling strategies with unbiased measurements of variables of interest. It is acknowledged that resources for such approaches will not permit universal application; instead they could be carried out at selected sites for specific purposes.

Chapter 6

General discussion: implications of this study for monitoring change across complex sites and landscapes

6.1 Why is so much vegetation monitoring in the UK a 'waste of time'?

This study highlights the lack of available knowledge amongst practitioners in the UK concerning sample design and statistical analyses for vegetation monitoring (Chapter 2). When interviewed, the majority of practitioners were able to list many criteria for effective monitoring, agreeing with published literature that the objectives should be clearly stated and that the monitoring should be repeatable with standard protocols and appropriately experienced observers (Pereira and Cooper, 2006). However, the criteria that they listed focused on the logistics of field implementation, with few practitioners mentioning magnitude of change, sample size or power analysis; this supports more general claims about a lack of practitioner knowledge of these issues (Foster, 2001; Legg and Nagy, 2006).

Since there is an array of literature detailing monitoring protocols including sample design (Grieg-Smith, 1957; Cochran, 1977; Elzinga et al., 2001), measurements in plots (Bonham, 1989; Sutherland, 2000) and statistical analysis (Green, 1979), along with advice about how to conduct monitoring (Goldsmith, 1991; Vos et al., 2000), it is not that guidance is not available, it is just not used. There are a number of reasons for this, including a lack of time to interpret the information, fear of statistics and inertia to change methodology (Legg and Nagy, 2006). The diversity of methods is very large due to the variety of objectives, vegetation types and spatial and temporal scales (Hockings et al., 2000; Hockings, 2003; Green et al., 2005; Pereira and Cooper, 2006), such that practitioners are left genuinely confused and tend to continue with existing methodology. Another problem is that although methods to monitor single species are relatively straightforward, tracking change in many species or in plant communities is complex and generally involves some information loss during measurement and analysis. Vegetation monitoring is also often planned and carried out by people who are skilled in plant identification, but who do not necessarily have an appropriate understanding of sampling design, which results in good species inventory but poor detection of change (Stohlgren, 2007).

The resistance of most practitioners to the change and development of existing monitoring methods has been noted as a further reason for the current lack of appropriate habitat monitoring methods in place (Hurlbert, 1984). There is a tendency to use field mapping of vegetation to provide spatial information on extent of habitats

or communities at a single point in time and use plots (often permanently located) in presumed homogenous stands across the site to monitor changes (Bakker *et al.*, 1996; Bakker *et al.*, 2002; Lengyel *et al.*, 2008a). This means that monitoring data is biased towards monitoring the changes in quality within the patches sampled as opposed to changes in spatial aspects such as extent and distribution of patches of different vegetation types. Furthermore, neither approach provides good quality data for detecting change since vegetation mapping is unsuitable for monitoring purposes (Chapter 3) and plots recorded in homogenous patches of vegetation tend to be small in size and number and result in biased plot location, a lack of replication and low power (Elzinga *et al.*, 2001). In the light of general misapprehension amongst conservation practitioners concerning vegetation monitoring, it is not surprising that there is an issue over the choice of methodology to monitor habitats across Hafod y Llan which in UK terms is a large and complex site.

Although stronger links between researchers and practitioners is often recommended in order to improve the quality and validity of monitoring (Lindenmayer, 1999), current practice is still insufficient since practitioners are unsure which sampling method to use in a given situation (JNCC, 2008c). Courses on sampling design and statistical analysis in environmental degree and post-graduate programmes should be supported by ongoing knowledge transfer between universities and conservation organisations. Specific training should also be provided according to practitioners' needs, for example monitoring to meet reporting requirements versus monitoring for management for conservation managers in the UK.

6.2 Implications of using qualitative methods for vegetation monitoring

Qualitative methods offer what appears to be a pragmatic solution to the complex problem of vegetation assessment for resource-limited conservation organisations. They can provide useful overviews of sites and landscapes, giving an indication of type, extent and condition of species and communities. When land managers need to know the spatial distribution of habitats on protected sites, Phase I or NVC mapping is usually commissioned. When national targets for the extent of priority habitats in favourable condition need to be reported, condition assessments are essential. However, the interviews reveal strong agreement among practitioners that the current use of qualitative methods is not effective for monitoring (Chapter 2) and this is validated by the multi-observer field trials of the National Vegetation Classification (NVC) system and Common Standards Monitoring (CSM) (Chapters 3 and 4).

Vegetation mapping is a coarse tool and is primarily used in planning, environmental impact assessment and management; it has been argued that repeat mapping of plant communities could be used to monitor management impacts (Dargie, 1993). In the case of Hafod y Llan, comparison of 2006 maps with those prepared in 1999 could be used to detect change in the area classified as Festuca-Agrostis-Galium grassland U4 to Calluna vulgaris-Erica cinerea dry heath H10 (Rodwell, 1991b; NT, 2000) (Chapter 3) in response to reduction of the high intensity grazing (Averis et al. 2004; Fig 1.1; Fig. 1.2). However, there are fundamental problems in the use of this type of mapping for monitoring; firstly mapping is often not sample-based and so the probability of making a type I or II error i.e. allocating a patch to the wrong NVC type, and the power to detect change cannot be calculated. Secondly, the extent of a change from one vegetation type to another cannot be quantified, since there is no way of knowing where a patch lies within the entire range of any category. Finally, this study and that of Cherrill and McClean (1999b) show that discrepancies in boundary delineation and assignment to vegetation class using vegetation mapping are so high that discrepancies between observers would be significantly greater than the temporal change expected from the management interventions in Hafod y Llan.

The low repeatability between observers in vegetation mapping is explained by the fact that the assignment of a patch to a particular vegetation type is largely a judgement of personal opinion. Any system of classification simplifies reality to some extent and requires a decision based on the best available knowledge. Although the concept of stable, homogenous areas forming a tessellated landscape has served a useful role in increasing knowledge about plant and environment interactions (Daubenmire, 1968), it has long been argued that 'patches' are not a good model of reality and attempting their delineation and assignment to community is subjective (Gleason, 1926). Assignment of a patch to a plant community depends on the surveyor's ability to identify and estimate abundance of several vascular and non vascular plants which requires high levels of botanical expertise and even then different surveyors can interpret the same landscape in different ways. A technological fix is not really possible as even the use of computer software to

ordinate quadrat data into vegetation types results in high levels of error (Palmer, 1992). Using vegetation mapping for monitoring also assumes that patches of a certain vegetation type will change as a whole; this is an implicit assumption in theories of vegetation succession and chronosequence studies (Clements, 1936).

The recognition of boundaries between vegetation types or plant communities is particularly problematic as surveyors seldom agree on the extent of patches in the absence of a distinct boundary such as a fence line. Furthermore, boundaries themselves often represent ecotones and are interesting habitats in their own right (Franklin, 1995). In terms of detecting change across landscapes, ecotones often contain the most spatially and temporally dynamic vegetation and are recommended as foci for monitoring (Caldas and White, 1983).

Direct field-based assessments of vegetation condition can also be a useful tool where resources are limited; it is growing in popularity and is widely used at the landscape scale in North America and Australia (Oldham *et al.*, 1995; Parkes *et al.*, 2003). These assessments involve value judgements about the quality of habitats or species and there are a range of methods available, varying in the attributes which are assessed and level of detail (Briggs and Freudenberger, 2006). Whilst rapid quality evaluation of protected areas has been proposed in some European countries, the UK is unusual in Europe in having implemented condition assessment as the main monitoring method on protected sites, in the form of Common Standards Monitoring (CSM) (Svátek and Buček, 2007; Lengyel *et al.*, 2008b). Although CSM is perceived outside the UK to be a rigorous, effective method (Teder *et al.*, 2007), the practitioner interviews conducted in the current study raised issues of inconsistency of outcomes, and these concerns were supported in results from the field trials (Chapters 2 and 4).

Any attempt to develop rapid tools for use by non-specialists relies on standardised, detailed protocols; a problem with CSM is that it attempts to combine fixed procedures for some aspects and flexible interpretation of guidance for others, notably the choice of field method (Chapter 4). Even if the same criteria are applied using a range of field methods leads to inconsistent assessments and frustrates experienced surveyors who feel they can recognise 'good' and 'bad' quality habitats, yet find that applying the prescribed criteria sometimes does not score them in this way. Gibbons and Freudenberger (2006) note that it is essential to match condition

187

assessment/monitoring methods with management objectives, time and expertise of the surveyor. Since the standard CSM method often fails to do this, both since it is used by observers with a wide range of experience, and because the link between management and monitoring is lost, a lot of time and resources are being spent collecting potentially misleading information (Chapter 4; Everett, 2004).

The frustration experienced by monitoring ecologists involved in the use of the standard UK CSM methodology led to the development of alternative methods intended to overcome the original problems (Hurford et al., 2001; Hurford and Perry, 2001). Although they may well be better, the fact that the alternative methods and the original approaches are used concurrently in different regions produces misleading information at the UK level, due to differences both in sampling methodology and especially in target frequency thresholds for certain attributes (Chapter 4). For example, reported figures for Indicator 21, 'Condition of features on Natura 2000 sites' in Wales (WAG, 2008) and Indicator H2 Condition of Sites of Special Scientific Interest (SSSIs) in England' (DEFRA, 2009a) do not quantify their error and their consistency is not assessed when they are amalgamated into the UK Indicator 4, 'UK Priority Habitats data' (DEFRA, 2009b). This situation could be resolved by better communication amongst the organisations involved and through validation of the various approaches across a range of habitats in a range of sites. This is currently lacking for CSM but is essential to allow confidence in the information currently used for reporting against targets and measuring management effectiveness.

One possibility might be to use experts and ask them to make an explicit judgment about condition without collecting any data. When the Countryside Council for Wales urgently needed to know the condition of features on protected sites across Wales, they asked a panel of experts to work through a series of sites and make assessments based on their knowledge and experience (Rod Gritten 2007 *pers. comm.*). It may be that this opinion is actually more consistent and accurate than field-based assessments using CSM, but it is difficult to repeat through time due to landscape amnesia and also levels of consistency are unknowable without reference to a quantitative control. What is needed are ways to record expert knowledge to provide a baseline for temporal comparison perhaps as standardised photos of examples of priority habitats in favourable and unfavourable condition (Haydock and Shaw, 1975). This could be supported with regular training and quality assurance, and use of photo monitoring such as in Fig. 1.2 to cross-reference across sites. Potentially, these rapid assessments based on benchmark photos could be validated by a series of detailed, quantitative assessments of condition made at a smaller number of sites.

CSM guidance was based on a great deal of work by experienced ecologists who recognised that field assessment requires time and effort and in the original formulation, the guidance included detailed assessments which have since been reduced into simplistic condition categories. Even improving the consistency of condition assessments will still only allow crude measures of change between condition categories even if these evolve beyond the dichotomous 'favourable' and unfavourable'. Other intermediate categories were originally intended to be used in CSM, but are generally not used due to the subjectivity involved, for example in the judgement of when an area is 'favourable recovering' (Jackson and Gaston, 2008). It is also argued that target setting and reporting is often a great waste of valuable resources (Hare et al., 2007; Jeeves, 2007) since the link between condition and the ecology and management of sites is lost. In some cases, favourable condition is unattainable due to anthropogenic or natural pressures and in other cases there is insufficient understanding of the ecology of priority habitats to know how to achieve favourable condition through prescribed management interventions (Everett, 2004). This makes targets such as Natural England's aim of 'bringing into favourable condition, by 2010, 95% of all nationally important wildlife sites' (HM Treasury 2004) both unachievable and futile.

Similar to vegetation mapping, quantifying the spatial extent of changes is difficult unless sample data is collected which can be used to calculate the power and probability of making type I and II errors. When these calculations are possible, it is clear that confidence in condition assessments is extremely low; Chapter 5 shows that, using current CSM methods in an area of degraded wet heath on Hafod y Llan, the power to detect a change from unfavourable to favourable condition is only 0.26 i.e. there is a 74% chance that the assessment of favourable condition is incorrect. Obviously, sample size could be increased to augment the power and increase the reliability of the assessment, or more sensitive measurements taken but these measures will increase the time investment in the method which would negate its value as a supposedly cheaper surrogate for more quantitative methods.

189

6.3 Implications of using quantitative methods for vegetation monitoring

Although there is a tendency to view quantitative methods as more reliable than their qualitative counterparts, this is often based on insufficient understanding of techniques (Chapter 2). Quantitative methods will only detect the levels of change necessary to meet management objectives with appropriate power when they are explicitly designed to do so. The magnitude of change and power are rarely stated in reports of conservation monitoring, either *a priori* or *post hoc*, and sample size calculations are invariably ignored. Quantitative monitoring could be vastly improved by stating an *a priori* definition of meaningful potential change, and ensuring that the variables measured are sufficiently sensitive over the monitored time period and are indicative of the desired change. Sample size calculations should also be carried out to ensure that adequate power is reached, and that Type I and II error rates are balanced according to management needs considering the costs of either missing a change or falsely identifying one.

It is likely that if these calculations were performed for existing monitoring programmes, it would reveal that much current monitoring would require much larger sample sizes than currently used to achieve adequate power to detect the size of change implicit in objectives. For instance, the systematic sampling system implemented in Hafod y Llan in this study demonstrated that detecting changes of 30% (similar to the size of change implicit in the objectives of many conservation programmes) requires up to 6.5 days of sampling effort (Chapter 5). As it is, we simply do not know the error or power associated with most monitoring and consequently reports of biodiversity at regional, national and international levels which are often taken literally should instead be viewed as no more than a rough guide. There is little evidence of questioning of reported figures such as the Welsh Indicator 19a 'Trends in BAP species', even though the data included is of widely varying quality, or the Living Planet Index, despite the >25% missing values in its time series and lack of differentiation between the quality (confidence levels) of the datasets used (WAG, 2006; Collen *et al.*, 2009).

The fact that quantitative monitoring design and analysis has tended to overlook the effect of spatial pattern has also had a large effect on detecting change. In a review of habitat monitoring schemes across Europe, Lengyel *et al.* (2008a) found that in over half of the schemes, spatial variation in habitats is either not monitored or

monitored by unspecified methods. They concluded that this is either because schemes operate at small spatial scales or because co-ordinators 'could not decide on or did not think it important to record their spatial method'. Vegetation patterns can be discerned at multiple spatial scales, with species responding individually to environmental and management factors and to competition with other species, inclusion of spatial variation is essential to detect change in range, area or fragmentation of habitats and species (Arita *et al.*, 2002; Legendre *et al.*, 2004; Chiarucci *et al.*, 2008). This results in unpredictability through scales and means that sampling at any single scale is likely to miss important spatial variation at other levels of resolution (Chapter 5) and that results from any single site cannot be safely extrapolated to wider areas (Hobbs, 2003). Furthermore, the more spatial variation there is in a given sampling domain, the harder it is to detect change over time, particularly if the 'noise' within data at one point in time is more than the change in data between two times.

It has been suggested that multi-scale and nested sampling designs should be used, both to quantify pattern at single sites and across landscapes and to ensure compatibly with data from other monitoring schemes (Stohlgren *et al.*, 1995; Critchley and Poulton, 1998). For example, the United States Department of Agriculture (USDA) Forest Health Monitoring Programme uses a nationwide systematic array of sample plots which provide a large, unbiased sample of the nation's forests, with measurements of different groups of species in 168 m², 17 m² and 1 m² nested plots (Ritters *et al.*, 1992). Plots are recorded annually and provide information about condition and trends in forest ecosystems along with estimated variance at multiple scales, with replicates at each scale. As part of this national forest survey, sampling with partial replacement has been shown to provide cost-efficient quantification of changes in time and space (Scott, 1998).

Several other studies have also shown that the scale of measurement and sampling design affects results and that information from multiple scales is needed in the study of complex environments (Dutilleul, 1993; Stohlgren *et al.*, 1997). For example, a study conducted in a forest plantation in North America found that multi-scale sampling allowed complex spatial patterns between soil, forest floor, and plant community variables to be investigated (Lister *et al.*, 2000). A limitation to multi-scale sampling is the cost; generally more plots are used and macroplots are larger,

making the sampling and the analysis complicated and time-consuming (Critchley and Poulton, 1998; Stohlgren *et al.*, 1998). A typical monitoring scheme may place several 2 m x 2 m plots in a patch of lowland grassland and record cover of all species, whereas a multi-scale scheme could measure fewer nested plots for comparable costs. Cost needs to be weighed up against data value; the single size plots might cover a greater geographic range and more rare habitats but may be too small to capture locally rare plants and would not be able to assess within-site spatial variation in species richness, cover or frequency, or species interactions.

Since results from Chapter 5 demonstrate that spatial variance is not predictable through scales and that different species display very different spatial distributions, a multi-scale and nested design could be a valuable approach in the UK. This design could be used to build national datasets from site assessments, with the higher costs justified through the provision of precise, quantitative data against which more widespread, simpler qualitative assessments can be validated. Diagnostic test methodology, as used in medicine to test the accuracy of screening tests (Chapter 4), could be a useful approach to validate the accuracy and consistency of the qualitative assessments against the detailed quantitative data.

Monitoring methods are often overly influenced by common, dominant species, with interpretation assuming that changes in them are representative of changes in the community as a whole. This has partially resulted from the prevailing paradigm of the existence of stable communities identified by just a few dominant species (e.g. as reflected in the NVC approach), and partially because common species are easier to identify, measure and locate and more is known about them. The use of indicator or keystone species have been suggested to improve cost-efficiency of monitoring (Simberloff, 1998; Carignan and Villard, 2002; Diekmann, 2003), although neither approach has fully been investigated in the UK. Different species respond in different ways to the same environmental factors and plant ecologists have long noted the importance of measuring individual plant species and understanding their ecology in order to interpret the trends detected (Gleason, 1926). Ideally, methods need to be able to pick up change in locally rare species (although this requires more focused effort) alongside change in more widespread, common species (Green and Young, 1993; Rosenzweig, 1995); it cannot be assumed that trends in a few indicator species represent whole community change.

192

The use of *aspects* of communities as a surrogate for records of a large number of single species has also been proposed several times (Noss, 1990; Goldsmith, 1991; Ritters *et al.*, 1992; Noss, 1999; Gray and Azuma, 2005). Surrogate indicators can be used in both quantitative and quantitative surveys and take far less time than a full species enumeration and may offer a significant saving if the strength and quality of the relationship between the surrogate and the community is known (Allen *et al.*, 2003). Indicators also have the mixed blessing of easy communication to policy makers and the public (Schiller *et al.*, 2001; Turnhout *et al.*, 2007).

The fact that there is little standardisation of quantitative techniques for monitoring vegetation over large areas reflects the problems involved in their development. Measuring the change in the area of dry heath shown in Fig. 1.2, along with all of the other changes occurring at the site over the same period is never going to be easy involving as it does the quantification of spatial and temporal changes at individual and community level as well as independent changes in environmental conditions and management interventions. The development of field and modelling techniques to detect and quantify patterns in space and time and reveal underlying mechanisms is seen by some as the 'Holy Grail' of vegetation monitoring (Stohlgren, 2007). The multi-scale and nested approaches which have been suggested involve teams of plant taxonomists, remote sensing specialists, spatial modellers, data managers, computer programmers and landscape ecologists in order to separate responses to short term threats from changes in response to longer term stresses (Likens, 1991).

Finally, a single monitoring system will never meet everyone's needs. For the Joint Nature Conservation Committee in the UK, a qualitative index may be all that is required in order to report against national targets, while at the site level a resource-strapped manager needs to know the most cost-effective interventions to improve biodiversity value and to be able to spot a change, for instance picking up the spread of an invasive species when it is small enough to be controlled. Different methods are required for these different purposes, and conservation managers need to be enabled to take effective decisions based on good understanding of available monitoring methodologies, and to have the confidence to adapt their practice over time.

6.4 Conclusions and recommendations for practice and future research

This study found that conservation practitioners are aware of several problems in qualitative methods widely used in site survey and monitoring in the UK. My field trials of mapping using the National Vegetation Classification (NVC) and of Common Standards Monitoring (CSM) supported these concerns and furthermore showed that these systems have fundamental flaws when used for monitoring. This study concludes that vegetation mapping such as that using the NVC should not be used for monitoring purposes and where it is used for site evaluation this should be with full acknowledgement of the inherently subjective and uncertain nature of the maps produced. Furthermore, the implications of inconsistency in the current application of CSM for reporting against targets should be recognized and although widespread condition assessment could be used to provide a snapshot of condition (possibly using standardised condition photographs), these field-based condition assessments should not be used for monitoring temporal change without some quantifiable controls. It is also concluded that the tendency among practitioners to view all quantitative monitoring methods as reliable is not based on evidence but rather an erroneous assumption that quantification is good which is based on a lack of understanding of what makes an effective quantitative method.

This study recommends that national monitoring in the UK should consist of a combination of linked qualitative and quantitative methods. At the widespread level, qualitative methods should be used to provide a coarse assessment of the condition of all sites of conservation interest. This qualitative method should use the rationale behind current Common Standards Monitoring but reduce the amount of detail involved through the use of photographs illustrating example condition categories. By reducing the detail and accepting that this is only a coarse measure of condition, it will be possible to measure a large number of sites. Key aspects of the approach are regular surveyor training in the recognition of habitat condition, and also regular moderation between surveyors to ensure consistency. This could be facilitated through the use of habitat experts who focus on two or three specific habitats and moderate their assessment at a national level.

As part of this approach, quantitative methods should provide a means of validation for the qualitative assessments. This should consist of a series of detailed, robust measurements of a range of habitats across a sub-sample of conservation sites. Sampling should enable the assessment of condition to enable validation of the qualitative aspect, and it is recommended that diagnostic test methodology be used as a verification method (as used in medicine to determine the accuracy of screening tests). This aspect of the quantitative approach should therefore remain consistent over time, with other aspects evolving over time according to changes at national, regional and local levels (e.g. resulting from climate induced change at a national level or particular species abundance response at a local level). A sub-set of sites should also remain fixed over time (possibly stratified by habitat and condition) with additional sites dropping in and out. The quantitative methods should comply with recommendations made previously in this thesis, namely that a priori power analyses should be used to demonstrate the efficiency and capability of the chosen monitoring system, appropriate statistical and multivariate methods be used to detect change and spatial scale and spatial variation be considered. As in the qualitative assessments, key to this approach is the use of well trained surveyors and an effective system of moderation and quality assurance. Whilst it will require large resource input, it is anticipated that the savings from simplification of current Common Standards Monitoring will facilitate this.

These recommendations involve substantial changes to existing methods which will require widespread discussion in order to reach consensual agreement at a national level; this is vital to ensure consistency and comparability of reported figures. Since the qualitative methods are based on existing CSM, there is already a common ground for discussion along with the benefit of future assessments retaining comparability with previous figures. Quantitative methods should also make use of existing site-based monitoring such as the Environmental Change Network of measurements. Further research into quantitative methods will also be required; this should focus on the development of a multi-scale and nested quantitative sampling design which could be used to address monitoring and surveillance questions and be used to test spatiotemporal models of vegetation change at landscape scales in the UK. Along with ensuring comparability at a UK level, it is also important to consider how monitoring compares throughout Europe and even internationally, particularly in relation to reporting against common biodiversity targets. Projects such as EuMon (EuMon, 2010) and Ebone (EBONE, 2010) have been set up to ensure consistent

monitoring across Europe; it is important to collaborate in these dialogues in order to share knowledge and ideas.

The development and reporting of figures should be as transparent as possible, and both qualitative and quantitative figures included for comparison. The effectiveness of this approach will be tested through (a) the ability of the quantitative methods to detect change over time at various scales (substantiated by detailed research) and (b) validation of qualitative assessments via diagnostic test methodology using the quantitative sampling.

At a local level e.g. Hafod y Llan, it is anticipated that the qualitative and quantitative assessments will be complementary, with widespread qualitative assessments carried out across the site (as outlined above). Resulting assessments should be used to identify areas of particular concern in which detailed quantitative monitoring is implemented, using the savings in resources from the reduced effort condition monitoring. Statutory conservation organisation monitoring teams should lead this monitoring and carry out all of the qualitative and quantitative monitoring used for reporting and diagnostic test validation. However, local site managers should be involved in the monitoring and provided with the assessment results in order to inform their site management and further monitoring. Where managers decide to carry out additional monitoring, basic advice and training on sampling design and statistical analysis should be made available; this could be through a handbook or via training sessions specifically for conservation managers. Training should stress the importance of considering meaningful *a priori* change, Type I and II error rates, spatial scale and heterogeneity of particular site/landscape.

References

- Alexander, M. and Rowell, T.A. 1999. Recent developments in management planning and monitoring on protected sites in the UK Parks. Parks 9, 50-55.
- Allegrini, M.-C., Canullo, R. and Campetella, G. 2009. ICP-Forests (International Cooperative Programme on Assessment and Monitoring of Air Pollution Effects on Forests): Quality Assurance procedure in plant diversity monitoring. Journal of Environmental Monitoring 11, 782–787.
- Allen, D. 2006. Surveillance. In: C. Hurford and M. Schneider, eds. Monitoring Nature Conservation in Cultural Habitats: a Practical Guide and Case Studies. Springer, Dordrecht, The Netherlands; pp. 25-42.
- Allen, R.B., Bellingham, P.J. and Wiser, S.K. 2003. Developing a forest biodiversity monitoring approach for New Zealand. New Zealand Journal of Ecology 27, 207-220.
- Altman, D.G. and Bland, J.M. 1994a. Diagnostic Tests 3: Receiver operating characteristic plots 7. British Medical Journal 309, 188.
- Altman, D.G. and Bland, J.M. 1994b. Statistics Notes: Diagnostic tests 1: Sensitivity and specificity. British Medical Journal 308, 1552.
- Anderson, P. and Yalden, D.W. 1981. Increased sheep numbers and the loss of heather moorland in the Peak District, England. Biological Conservation 20, 195-213.
- Angold, P.G., Gurnell, A.M. and Edwards, P.J. 1996. Locational errors in maps from environmental surveys and their implications for information extraction. Journal of Environmental Management 47, 341-354.
- Antoine, G., Jean-Paul, T. and Felix, K. 1998. Predicting the potential distribution of plant species in an alpine environment. Journal of Vegetation Science 9, 65-74.
- Appleby, C.E. 1991. Monitoring at the county level. In: B. Goldsmith, ed. Vegetation Monitoring. Chapman and Hall, London; pp. 155-173.
- Archaux, F., Berges, L. and Chevalier, R. 2007. Are plant censuses carried out on small quadrats more reliable than on larger ones? Plant Ecology 188, 179-190.
- Archaux, F., Gosselin, F., Berges, L. and Chevalier, R. 2006. Effects of sampling time, species richness and observer on the exhaustiveness of plant censuses. Journal of Vegetation Science 17, 299-306.
- Ardeshir, D. 2005. Condition Assessment of 'European Dry Heath', 'Northern Atlantic Wet Heath' and 'Blanket Bog' Habitats at the Eryri Candidate Special Area of Conservation (cSAC). CCW Environmental Monitoring Report, No. 26. Countryside Council for Wales, Bangor.
- Arita, H.T., Rodriguez and P. 2002. Geographic range, turnover rate and the scaling of species diversity. Ecography 25, 541-550.
- Aspinall, R.J. and Pearson, D.M. 1995. Describing and managing uncertainty of categorical maps in GIS. In: P. Fisher, ed. Innovations in GIS 2. Taylor and Francis, London; pp. 71-83.
- August, P., Michaud, J., Labash, C. and Smith, C. 1994. GPS for environmental applications: accuracy and precision of locational data. Photogrammetric Engineering and Remote Sensing 60, 41-45.
- Averis, A., Averis, B., Birks, J., Horsfield, D., Thompson, D. and Yeo, M. 2004. An Illustrated Guide to British Upland Vegetation. Joint Nature Conservation Committee, Peterborough.
- Bacaro, G., Baragatti, E. and Chiarucci, A. 2009. Using taxonomic data to assess and monitor biodiversity: are the tribes still fighting? Journal of Environmental Monitoring 11, 798–801.
- Bakker, J.P., Marrs, R.H. and Pakeman, R.J. 2002. Long-term vegetation dynamics: Successional patterns and processes. Introduction. Applied Vegetation Science 5, 2-6.
- Bakker, J.P., Willems, O.H. and Zobel, M. 1996. Why do we need permanent plots in the study of long-term vegetation dynamics? Journal of Vegetation Science 7, 147-155.
- Balmford, A. 1998. On hotspots and the use of indicators for reserve selection. Trends in Ecology and Evolution 13, 409.
- Balmford, A., Bennun, L., Brink, t., B., Cooper, D., Cote, I.M., Crane, P., Dobson, A., Dudley, N., Dutton, I., Green, R.E., Gregory, R.D., Harrison, J., Jaarsveld, v., A.S., Kennedy, E.T., Kremen, C., Leader-Williams, N., Lovejoy, T.E., Mace, G., May, R., Mayaux, P., Morling, P., Phillips, J., Redford, K., Ricketts, T.H., Rodriguez, J.P., Sanjayan, M., Schei, P.J. and Walther, B.A. 2005. The convention on biological diversity's 2010 target. Science 307, 212-213.
- Balmford, A., Green, M.J.B. and Murray, M.G. 1996. Using higher-taxon richness as a surrogate for species richness. I. Regional tests. Proc. R. Soc. Lond. B. 263, 1267-1274.
- Balmford, A., Green, R.E. and Jenkins, M. 2003. Measuring the changing state of nature. Trends in Ecology and Evolution 18, 326-330.
- Barbour, M.G., Burk, J.H., Pitts, W.D., Gilliam, F.S. and Schwartz, M.W. 1999. Terrestrial Plant Ecology. Benjamin/Cummins, Menlo Park, California.
- Bealey, C.E. and Cox, J. 2004. Validation Network Project. Upland habitats covering: blanket bog, dry dwarf shrub heath, wet dwarf shrub heath and *Ulex gallii* dwarf shrub heath. English Nature Research Report No. 564. English Nature, Peterborough.
- Belsky, A.J. 1985. Long-term vegetation monitoring in the Serengeti National Park, Tanzania. Journal of Applied Ecology 22, 449-460.
- Berendse, F. 1985. The effect of grazing on the outcome of competition between plant species with different nutrient requirements. Oikos 44, 35-39.
- Berti, G., Isocrono, D., Ropolo, L., Caranci, N., Cesare, M.R., Fossad, V. and Caramiello, R. 2009. An experience of data quality evaluation in pollen monitoring activities. Journal of Environmental Monitoring 11, 788–792.
- Birks, H.J.B. 2008. Ordination- an ever expanding tool-kit for ecologists? The British Ecological Society Bulletin 39, 31-33.
- Bonham, C. 1989. Measurements for Terrestrial Vegetation. John Wiley and Sons, Chichester.
- Boyle, D. 2007. Can you count a better environment? ECOS: A Review of Conservation 28, 1-6.
- Bråkenhielm, S. and Qinghong, L. 1995. Comparison of field methods in vegetation monitoring. Water, Air and Soil Pollution 79, 75-87.
- Brazier, D.P., Birch, K., Lindenbaum, C. and Wyn, G. 2009 in prep. Intertidal Phase 1 Mapping Validation Exercise. CCW Science Report. Countryside Council for Wales, Bangor.
- Briggs, S.V. and Freudenberger, D. 2006. Assessment and monitoring of vegetation condition: Moving forward. Ecological Management and Restoration 7, 74-76.
- Brown, A. 2001. Habitat Monitoring for Conservation Management and Reporting. 3: Technical Guide. LIFE- Nature project No. LIFE95 NAT/UK/000821 Integrating

monitoring with management planning: a demonstration of good practice on Natura 2000 sites in Wales. Countryside Council for Wales, Bangor.

- Brown, A. 2006. Strategic Sampling. In: C. Hurford and M. Schneider, eds. Monitoring Nature Conservation in Cultural Habitats: a Practical Guide and Case Studies. Springer, Dordrecht, The Netherlands; pp. 43-56.
- Brown, A. and Rowell, T.A. 1997. Integrating monitoring with management planning for nature conservation: some principles. Natur und Landschaft 72, 502-507.
- Brunialti, G., Giordani, P., Isocrono, D. and Loppi, S. 2002. Evaluation of data quality in lichen biomonitoring studies: the Italian experience. Environmental Monitoring and Assessment 75, 271-280.
- Brunner, R.D. and Clark, T.W. 1997. A practice-based approach to ecosystem management. Conservation Biology 11, 48-58.
- Bryman, A. 2001. Social Research Methods. Oxford University Press, Oxford.
- BSBI 2009. BSBI maps scheme. Available via http://www.bsbi.org.uk/html/atlas.html. Accessed 10/07/2009. Botanical Society for the British Isles, London.
- BTO 2009. Introduction to the Breeding Birds Survey 2009. Available via http://www.bto.org/bbs/take_part/introduction.htm. Accessed 10/07/2009. British Trust for Ornithology, Thetford.
- Bullock, J. 1996. Plants. In W. J. Sutherland, ed. Ecological Census Techniques: A Handbook. Cambridge University Press, Cambridge; pp. 186-213.
- Bunce, R.G.H., Barr, C.J., Gillespie, M.K., Howard, D.C., Scott, W.A., Smart, S.M., Poll, H.M.v.d. and Watkins, J.W. 1999. ECOFACT 1: Vegetation of the British countryside - the Countryside Vegetation System. NERC/Centre for Hydrology and Ecology.
- Bunce, R.G.H., Metzger, M.J., Jongman, R.H.G., Brandt, J., Blust, G., Elena-Rossello, R., Groom, G.B., Halada, L., Hofer, G., Howard, D.C., Kovar, P., Mucher, C.A., Padoa-Schioppa, E., Paelinx, D., Palo, A., Perez-Soba, M., Ramos, I.L., Roche, P., Skanes, H. and Wrba, T. 2008. A standardised procedure for surveillance and monitoring of European habitats and provision of spatial data. Landscape Ecology 23, 11-25.
- Bussotti, F., Cozzi, A., Cenni, E., Bettini, D., Sartia, C. and Ferretti, M. 2009. Quality assurance and measurement errors in monitoring tree crown conditions in Italy. Journal of Environmental Monitoring 11, 769-773.
- Caldas, F.B. and White, R.J. 1983. Sampling for vegetation survey: A generalized flexible model for sample location. Journal of Biogeography 10, 11-28.
- Cantarello, E. 2007. Towards cost-effective indicators to maintain Natura 2000 sites in a favourable conservation status. PhD thesis. University of Padova, Italy.
- Cantarello, E. and Newton, A.C. 2008. Identifying cost-effective indicators to assess the conservation status of forested habitats in Natura 2000 sites. Forest Ecology and Management 256, 815-826.
- Carignan, V. and Villard, M.-A. 2002. Selecting indicator species to monitor ecological integrity: a review. Environmental Monitoring and Assessment 78, 45-61.
- Carlsson, A.L.M., Bergfur, J. and Milberg, P. 2005. Comparison of data from two vegetation monitoring methods in semi-natural grasslands. Environmental Monitoring and Assessment 100, 235-248.
- Carroll, J.A., Caporn, S.J.M., Cawley, L., Read, D.J. and Lee, J.A. 1999 The effect of increased deposition of atmospheric nitrogen on *Calluna vulgaris* in upland Britain. New Phytologist 141, 423-431.

- Castle, G. and Mileto, R. 2003. Using the National Vegetation Classification (NVC) in woodland survey- 6000ha on. In: E. Goldberg, ed. National Vegetation Classification - Ten years' experience using the woodland section. Joint Nature Conservation Committee, Peterborough; pp. 76-86.
- CCW 2005a. Habitats of Wales Phase I Data, Lowlands and Uplands: 1979-1997. Countryside Council for Wales.
- CCW 2005b. Y Twyni o Abermenai i Aberffraw/Aberffraw to Abermenai Dunes SAC UK0020021. Fixed dunes with herbaceous vegetation ("grey dunes") H2130. SAC Monitoring Report. Countryside Council for Wales, Bangor.
- CCW 2009. Meta-data information for arctic-alpine plant monitoring for effects of climate change. Available via http://www.ccw.gov.uk/publications--research/library--information-servic.aspx. Accessed 24/06/2009. Countryside Council for Wales, Bangor.
- CEFAS 2009. International Bottom Trawl. Available via http://www.cefas.co.uk. Accessed 10/07/2009. Centre for Environment, Fisheries and Aquaculture Science, London.
- CEH 2009. About the Countryside Survey. Available via http://www.countrysidesurvey.org.uk. Accessed 10/07/2009. Centre for Ecology and Hydrology, Lancaster.
- Cheal, D. 2008. Repeatability of cover estimates? Ecological Management and Restoration 9, 67-68.
- Cheffings, C.F. and Farrell, L. 2005. The Vascular Plant Red Data List for Great Britain. Species Status 7: 1-116. Joint Nature Conservation Committee, Peterborough.
- Cherrill, A. and McClean, C. 1995. An investigation of uncertainty in field habitat mapping and the implications for detecting land cover change. Landscape Ecology 10, 5-21.
- Cherrill, A. and McClean, C. 1999a. Between observer variation in the application of a standard method of habitat mapping by environmental consultants in the UK. Journal of Applied Ecology 36, 989-1008.
- Cherrill, A. and McClean, C. 1999b. The reliability of 'Phase 1' habitat mapping in the UK: the extent and types of observer bias. Landscape and Urban Planning 45, 131-143.
- Chiarucci, A., Bacaro, G. and Rocchini, D. 2008. Quantifying plant species diversity in a Natura 2000 network: Old ideas and new proposals. Biological Conservation 141, 2608-2618.
- Chiarucci, A., Maccherini, S. and Dominicis, V.D. 2001. Evaluation and monitoring of the flora in a nature reserve by estimation methods. Biological Conservation 101, 305-314.
- Clare, T. and Howard, D. 2000. Estimating change in landscape elements using different sample sizes. In: Quantitative Approaches to Landscape Ecology, proceedings of the ninth Annual Conference of the IALE (UK) held at the University of Wales, Bangor 7th-9th September 2000, pp. 61-69.
- Clark, I. and Harper, W.V. 2000. Practical Geostatistics 2000. Geostokos (Ecosse) Limited, Scotland.
- Clarke, R. (Ed.), 1986. The Handbook of Ecological Monitoring. Clarendon Press, Oxford.
- Clements, F.E. 1936. The nature and structure of the climax. Journal of Ecology 24, 252-284.
- Cochran, W.G. 1977. Sampling Techniques. John Wiley and Sons, New York.

- Cohen, J. 1988. Statistical Power Analysis for the Behavioural Sciences. Lawrence Erlbaum Associates, New Jersey.
- Collaboration, T.C. 2008. What is a Cochrane diagnostic test accuracy review?. Available via http://www.cochrane.org/. Accessed 10/10/2008. The Cochrane Collaboration, Oxford.
- Collen, B., Loh, J., Whitmee, S., Mcrae, L., Amin, R. and Baillie, J.E.M. 2009. Monitoring change in vertebrate abundance: the Living Planet Index. Conservation Biology 23, 317-327.
- Contamin, R., Ellison, A.M. and Aaron, M. 2009. Indicators of regime shifts in ecological systems: What do we need to know and when do we need to know it? Ecological applications 19, 799-816.
- Cornwell, W.K. and Grubb, P.J. 2003. Regional and local patterns in plant species richness with respect to resource availability. Oikos 100, 417-428.
- Couteron, P., Ollier and S. 2005. A generalized, variogram-based framework for multi-scale ordination. Ecology 86, 828-834.
- Cressie, N.A.C. 1993. Statistics for Spatial Data. John Wiley and Sons, Chichester.
- Critchley, C.N.R. 2000. Ecological assessment of plant communities by reference to species traits and habitat preferences. Biodiversity and Conservation 9, 87 105.
- Critchley, C.N.R. and Poulton, S.M.C. 1998. A method to optimise precision and scale in grassland monitoring. Journal of Vegetation Science 9, 837-846.
- Dale, M.B. 1994. Do ecological communities exist? Journal of Vegetation Science 5, 285-286.
- Dale, M.R.T. 1999. Spatial Pattern Analysis in Plant Ecology. Cambridge University Press, Cambridge.
- Dale, M.R.T. and Fortin, M.-J. 2002. Spatial autocorrelation and statistical tests in ecology. Ecoscience 9, 162-167.
- Dargie, T. 1993. Repeat Phase 1 Habitat Survey and the Detection of Vegetation Change. English Nature Research Report No. 51. English Nature, Peterborough.
- Daubenmire, R.F. 1968. Plant Communities: A Textbook of Plant Synecology. Harper and Row, New York.
- de. Gruijter, J., Brus, D., Bierkens, M.F.P. and Knotters, M. 2006. Sampling for Natural Resource Monitoring. Springer-Verlag, Berlin, Heidelberg.
- DEFRA 2009a. England Biodiversity Strategy Indicators. Available via http://www.defra.gov.uk/wildlifecountryside/biodiversity/indicator.htm. Accessed 07/07/2009. Department for Environment, Food and Rural Affairs, London.
- DEFRA 2009b. UK Biodiversity Indicators in Your Pocket 2009. Avaiable via http://www.defra.gov.uk/. Accessed 07/07/2009. Department for Environment, Food and Rural Affairs, London.
- Dethier, M.N., Graham, E.S., Cohen, S. and Tear, L.M. 1993. Visual versus randompoint percent cover estimations: objective is not always better. Marine Ecology Progress Series 96, 93-100.
- Devillers, P., Devillers-Terschuren, J. and Ledant, J.P. 1991. CORINE Biotopes Manual: Habitats of the European Community. Office for Official Publications of the European Communities, Luxembourg.
- Diekmann, M. 2003. Species indicator values as an important tool in applied plant ecology- a review. Basic and Applied Ecology 4, 493-506.

- Dungan, J.L., Perry, J.N., Dale, M.R.T., Legendre, P., Citron-Pousty, S., Fortin, M.-J., Jakomulska, A., Miriti, M. and Rosenberg, M.S. 2002. A balanced view of scale in spatial statistical analysis. Ecography 25, 626-640.
- Dutilleul, P. 1993. Spatial heterogeneity and the design of ecological field experiments. Ecology 74, 1646-1658.
- EBONE 2010. EBONE / European Biodiversity Observation Network. Available via http://www.ebone.wur.nl/UK/. Accessed 24/01/2010.
- EC 2009a. Biodiversity. Available via http://www.ec.europa.eu/ environment/nature/natura2000. Accessed 10/07/2009. European Commission, Brussels.
- EC 2009b. Natura 2000 network. Available via http://www.ec.europa.eu/ environment/nature/natura2000. Accessed 10/07/2009. European Commission, Brussels.
- ECN 2009. The UK Environmental Change Network. Available via: http://www.ecn.ac.uk/. Accessed 12/10/2009.
- EEC 1992. Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora. Committee of the Communities of the European Union, Brussels.
- Ejrnaes, R. and Bruun, H.H. 2000. Gradient Analysis of dry grassland vegetation in Denmark. Journal of Vegetation Science 11, 573-584.
- Elkington, T., Dayton, N., Jackson, D.L. and Strachan, I.M. 2002. National Vegetation Classification: Field Guide to Mires and Heaths. Joint Nature Conservation Committee, Peterborough.
- Elzinga, C.L., Salzer, D.W., Willoughby, J.W. and Gibbs, J. 2001. Monitoring Plant and Animal Populations. Blackwell Science, New York.
- EN 2005. Target 2010: The condition of England's Sites of Special Scientific Interest in 2005. English Nature, Peterborough.
- Enøe, C., Georgiadis, M.P. and Johnson, W.O. 2001. Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. Preventive Veterinary Medicine 45, 61-81.
- ERFF 2009. UK Environmental Observation Framework. Environment Research Funders' Forum, Swindon.
- ESRI 2002. ArcView Version 3.3. Environmental Systems Research Inc, Redlands California.
- EuMon 2010. EU-wide monitoring methods and systems of surveillance for species and habitats of community interest. Available via http://eumon.ckff.si/summary.php. Accessed 24/01/2010. European Commission.
- Everett, S. 2004. The quality of nature in England. British Wildlife 15, 168-173.
- Fagan, W.F., Fortin, M.-J. and Soykan, C. 2003. Integrating edge detection and dynamic modelling in quantitative analyses of ecological boundaries. BioScience 53, 730-738.
- Farmer, C. 2008. Species descriptions for *Sedum acre* and *Arenaria serpyllifolia*. Available via http://www.plant-identification.co.uk. Accessed 28/03/2009. Skye Flora.
- Faul, F., Erdfelder, E., Lang, A.-G. and Buchner, A. 2007. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods 39, 175-191.
- FC 2009. National inventory of woodland and trees. Available via: http://www.forestry.gov.uk/inventory. Accessed 12/10/2009.

- Ferretti, M. 2009. Quality Assurance in ecological monitoring- towards a unifying perspective. Journal of Environmental Monitoring 11, 726-729.
- Fielding 1993. Qualitative Interviewing. In N. Gilbert, ed. Researching social life. Sage Publications, London.
- Fielding, A.H. and Bell, J.F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 24, 38-49.
- Firbank, L.G., Barr, C.J., Bunce, R.G.H., Furse, M.T., Haines-Young, R., Hornung, M., Howard, D.C., Sheail, J.A., Sier, A., Smart and S.M. 2003. Assessing stock and change in land cover and biodiversity in GB: an introduction to Countryside Survey 2000. Journal of Environmental Management 67, 207-218.
- Firbank, L., Asteraki, L., Chamberlain, D., Gough, S., Haysom, K., Petit, S., Pywell, R., Warman, L. and Vickery, J. 2001. Development and validation of a methodology for the condition assessment of extensively managed habitats. Final Project Report OC9801. Ministry of Agriculture, Fisheries and Food, London.
- Fitter, A.H. and Peat, H.J. 2008. The Ecological Flora Database.
- Fortin, M.-J. 1997. Effects of data types on vegetation boundary delineation. Canadian Journal of Forestry Research 27, 1851-1858.
- Fortin, M.-J. 1999. The effects of quadrat size and data measurement on the detection of boundaries. Journal of Vegetation Science 10, 43-50.
- Fortin, M.-J. and Dale, M. 2005. Spatial Analysis: A Guide for Ecologists. Cambridge University Press, Cambridge.
- Fortin, M.-J., Drapeau, P. and Legendre, P. 1989. Spatial autocorrelation and sampling design in plant ecology. Vegetatio 83, 209-222.
- Fortin, M.-J., Olson, R.J., Ferson, S., Iverson, L., Hunsaker, C., Edwards, G., Levine, D., Butera, K. and Klemas, V. 2000. Issues related to the detection of boundaries. Landscape Ecology 15, 453-466.
- Foster, J.R. 2001. Statistical power in forest monitoring. Forest Ecology and Management 151, 211-222.
- Fowler, J., Cohen, L. and Jarvis, P. 2000. Practical Statistics for Field Biology. Wiley and Sons, Chichester.
- Fox, R., Warren, M., Brereton and T. 2007. A new Red List of British Butterflies. Contract report to the Joint Nature Conservation Committee No. F90-01-915.
- Francini, A., Pellegrini, E., Lorenzini, G. and Nali, C. 2009. Non-sampling error in ozone biomonitoring: the role of operator training. Journal of Environmental Monitoring 11, 736-744.
- Franklin, J. 1995. Predictive vegetation mapping: Geographic modelling of biospatial patterns in relation to environmental gradients. Progress in Physical Geography 19, 474- 499.
- Frogbrook, Z.L., Oliver, M.A., Salahi, M., Ellis and R.H. 2002. Exploring the spatial relations between cereal yield and soil chemical properties and the implications for sampling. Soil Use and Management 18, 1-9.
- Frogbrook, Z. 2000. Geostatistics as an aid to soil management for precision agriculture. PhD thesis. Department of Soil Science, University of Reading, Reading.
- Gaston, K.J., Charman, K., Jackson, S.F., Armsworth, P.R., Bonn, A., Briers, R.A., Callaghan, C.S.Q., Catchpole, R., Hopkins, J., Kunin, .E., Latham, L., Opdam, P., Stoneman, R., Stroud, D.A., Tratt and R. 2006. The ecological

effectiveness of protected areas: The United Kingdom. Biological Conservation 132, 76-87.

- Gaston, K.J. and Spicer, J.I. 2004. Biodiversity: An Introduction. Blackwell Publishing, Oxford.
- Gaucherand, S. and Lavorel, S. 2007. New method for rapid assessment of the functional composition of herbaceous plant communities. Austral Ecology 32, 927-936.

Genstat 2005. Genstat Eighth Edition. Lawes Agricultual Trust.

- Gibbons, P. and Freudenberger, D. 2006. An overview of methods used to assess vegetation condition at the scale of the site. Ecological Management and Restoration 7, 10-17.
- Gibbons, P., Zerger, A., Jones, S. and Ryan, P. 2006. Mapping vegetation condition in the context of biodiversity conservation. Ecological Management and Restoration 7, 1-2.
- Gillison, A.N. and Carpenter, G. 1997. A generic plant functional attribute set and grammar for dynamic vegetation description and analysis. Functional Ecology 11, 775-783.
- Giordani, P., Brunialti, G., Benesperi, R., Rizzi, G., Frati, L. and Modenesi, P. 2009. Rapid biodiversity assessment in lichen diversity surveys: implications for quality assurance. Journal of Environmental Monitoring 11, 730–735.
- Glass, G.V. 1977. Integrating findings: the meta-analysis of research. Review of Research in Education 5, 351-379.
- Gleason, H.A. 1926. The individualistic concept of the plant association. Bulletin of the Torrey Botanical Club 53, 1-20.
- Goldsmith, F.B. 1991. Vegetation Monitoring. Chapman and Hall, London.
- Goslee, S.C. 2006. Behavior of vegetation sampling methods in the presence of spatial autocorrelation. Plant Ecology 187, 203-212.
- Gosz, J.R. 1993. Ecotone hierarchies. Ecological Applications 3, 369-376.
- Gotfryd, A. and Hansell, R.I.C. 1985. The impact of observer bias on multivariate analysis of vegetation structure. Oikos 45, 223-234.
- Gottardini, E., Cristofolini, F., Cristofori, A., Vannini, A. and Ferretti, M. 2009. Sampling bias and sampling errors in pollen counting in aerobiological monitoring in Italy. Journal of Environmental Monitoring 11, 751–755.
- Gray, A.N. and Azuma, D.L. 2005. Repeatability and implementation of a forest vegetation indicator. Ecological Indicators 5, 57-71.
- Green, R.E., Balmford, A., Crane, P.R., Mace, G.M., Reynolds, J.D. and Turner, R.K. 2005. A framework for improved monitoring of biodiversity: responses to the World Summit of Sustainable Development. Conservation Biology 19, 56-65.
- Green, R.H. 1979. Sampling Design and Statistical Methods for Environmental Biologists. Wiley and Sons, Chichester.
- Green, R.H. 1989. Power analysis and practical strategies for environmental monitoring. Environmental Research 50, 195-205.
- Green, R.H. and Young, R.C. 1993. Sampling to detect rare species. Ecological Applications 3, 351-356.
- Gregory, R.D., Strien, A. van, Vorisek, P., Meyling, A.W.G., Noble, D.G., Foppen, R.P.B. and Gibbons, D.W. 2005. Developing indicators for European birds. Phil Trans R Soc B 360, 269-288.
- Gregory, R.D., Wilkinson, N.I., Noble, D.G., Robinson, J.A., Brown, A.F., Hughes, J., Proctor, D.A., Gibbons, D.W. and Galbraith, C.A. 2002. The population status

of birds in the United Kingdom, Channel Islands and the Isle of Man: an analysis of conversation concern 2002-2007. British Birds 95, 410-450.

- Grieg-Smith, P. 1957. Quantitative Plant Ecology. Butterworths Scientific Publications, London.
- Grieg-Smith, P. 1986. Chaos or order- organisation. In: J. Kikkawa and D. J. Anderson, eds. Community Ecology- Pattern and Process. Blackwell Scientific, Oxford; pp. 19-29.
- Grubb, P.J. 1982. Control of relative abundance in roadside *Arrhenatheretum*: results of a long term garden experiment. Journal of Ecology 70, 845-861.
- Grubb, P.J., Lloyd, J.R., Pennington, T.D. and Whitmore, T.C. 1963. A comparison of montane and lowland rainforests in Ecuador. Ecology 51, 567-599.
- Gustafon, E.J. 1998. Quantifying landscape spatial pattern: what is the state of the art? Ecosystems 1, 143-156.
- Haines-Young, R.H., Barr, C.J., Black, H.I.J., Briggs, D.J., Bunce, R.G.H., Clarke, R.T., Cooper, A., Dawson, F.H., Firbank, L.G., Fuller, R.M., Furse, M.T., Gillespie, M.K., Hill, R., Hornung, M., Howard, D.C., McCann, T., Morecroft, M.D., Petit, S., Sier, A.R.J., Smart, S.M., Smith, G.M., Stott, A.P., Stuart, R.C. and Watkins, J.W. 2002. Countryside 2000 Report. Accounting for Nature: Assessing Habitats in the UK Countryside. Department of the Environment, Transport and the Regions, London.
- Hall, D. and Hall, I. 1996. Practical Social Research: Project Work in the Community. Macmillan Press, London.
- Hall, J.B. and Okali, D.U.U. 1978. Observer-bias in a floristic survey of complex tropical vegetation. Journal of Ecology 66, 241-249.
- Hall, J.E. 1997. An analysis of National Vegetation Classification survey data. JNCC Report No. 272. Joint Nature Conservation Committee, Peterborough.
- Hall, J.E., Kirby, K.J. and Whitbread, A.M. 2002. National vegetation classification: field guide to woodland and scrub. Joint Nature Conservation Committee, Peterborough.
- Halley, J.M., Hartley, S., Kallimanis, S., Kunin, W.E., Lennon, J.J. and Sgardelis, S.P. 2004. Uses and abuses of fractal methodology in ecology. Ecology Letters 7, 254-271.
- Hare, T., Boyle, D. and Jeeves, M. 2007. Measuring up to nature. ECOS: A Review of Conservation 28, 20- 27.
- Harte, J., McCarthy, S., Taylor, K., Kinzig, A. and Fischer, M.L. 1999. Estimating species-area relationship from plot to landscape scale using species spatial-turnover data. Oikos 86, 45-54.
- Hartley, S., Kunin, W.E., Lennon, J.J. and Pocock, M.J.O. 2004. Coherence and discontinuity in the scaling of species' distribution patterns. Proceedings of the Royal Society of London B, 81-88.
- Hartley, S.E. and Mitchell, R.J. 2005. Manipulation of nutrients and grazing levels on heather moorland: changes in *Calluna* dominance and consequences for community composition. Journal of Ecology 93, 990-1004.
- Hastings, H.M. and Sugihara, G. 1993. Fractals: A User's Guide for the Natural Sciences. Oxford University Press, New York.
- Haydock, K.P. and Shaw, N.H. 1975. The comparative yield method for estimating dry matter yield of pasture. Australian Journal of Experimental Agriculture and Animal Husbandry 15, 663-667.
- Hellawell, J.B. 1978. Biological Surveillance of Rivers. A Biological Monitoring Handbook. Water Research Centre, Medmenham and Stevenage.

- Hill, M.O. 1979. TWINSPAN- a FORTRAN program for arranging multivariate data in an ordered two-way table by classification of the individuals and attributes. Cornell University, New York.
- Hill, M.O. 1996. TABLEFIT version 1.0, for identifcation of vegetation types. Institute of Terrestrial Ecology, Huntingdon.
- HM Treasury 2004. SR 2004:Public service agreements. Available via http://www.hm-treasury.gov.uk/ spending-review/spend_sr04. Accessed 31/06/2009. HM Treasury, London.
- Hobbs, N.T. 2003. Challenges and opportunities in integrating ecological knowledge across scales. Forest Ecology and Management 181, 223-238.
- Hobbs, R.J. and Gimingham, C.H. 1987. Vegetation, fire and herbivore interactions in heathland. Advances in Ecological Research 16, 87-173.
- Hockings, H. 2003. Systems for assessing the effectiveness of management in protected areas. BioScience 53, 823-832.
- Hockings, M., Stolton, S. and Dudley, N. 2000. Evaluating Effectiveness: A Framework for Assessing the Management of Protected Areas. International Union for the Conservation of Nature: World Commission for Protected Areas, Gland, Switzerland.
- Hockley, N., Farrar, J., Gibbons, J.M., Hearn, S.M., Jones, J.P.G. and Kaiser, M. In draft. Biodiversity Indicators for Wales: A Report to the Welsh Assembly Government. Bangor University, Bangor.
- Hoef, J.M.v. and Glenn-Lewin, D.C. 1989. Multiscale ordination: a method for detecting pattern at several scales. Vegatatio 82, 59-67.
- Hope-Simpson, J.F. 1940. On the errors in the ordinary use of subjective frequency estimations in grassland. Journal of Ecology 28, 193-209.
- Hopley, L. and Schalkwyk, J.v. 2001. The Magnificent ROC (Receiver Operating Characteristic curve). Available via http://www.anaesthetist.com/mnm /stats/roc/Findex.htm. Accessed 24/06/2008.
- Hulme, P.D., Pakeman, R.J., Torvell, L., Fisher, J.M. and Gordon, I.J. 1999. The effects of controlled sheep grazing on the dynamics of upland Agrostis-Festuca grassland. Journal of Applied Ecology 36, 886-900.
- Hurford, C. 2006. Identifying site-specific condition indicators for habitats. In: C. Hurford and M. Schneider, eds. Monitoring Nature Conservation in Cultural Habitats: A Practical Guide and Case Studies. Springer, Dordrecht, The Netherlands; pp. 93-104.
- Hurford, C., Jones, M.R. and Brown, A. 2001. Habitat Monitoring for Conservation Management and Reporting. 2: Field Methods. LIFE- Nature project No. LIFE95 NAT/UK/000821 Integrating monitoring with management planning: a demonstration of good practice on Natura 2000 sites in Wales. Countryside Council for Wales, Bangor.
- Hurford, C. and Perry, K. 2001. Habitat Monitoring for Conservation Management and Reporting. 1: Case Studies. LIFE- Nature project No. LIFE95 NAT/UK/000821 Integrating monitoring with management planning: a demonstration of good practice on Natura 2000 sites in Wales. Countryside Council for Wales, Bangor.
- Hurford, C. and Schneider, M. 2006. Monitoring Nature Conservation in Cultural Habitats: A Practical Guide and Case Studies. Springer, Dordrecht, The Netherlands.
- Hurlbert, S.H. 1984. Pseudoreplication and the design of ecological field experiments. Ecological Monographs 54, 187-211.

- Isaac, N.J.B., Mallet, J. and Mace, G.M. 2004. Taxonomic inflation: its influence on macroecology and conservation. Trends in Ecology and Evolution 19, 464-469.
- Isaaks, E.H. and Srivastava, R.M. 1989. An Introduction to Applied Geostatistics. Oxford University Press, New York.
- IUCN 2009. IUCN Red List of Threatened Species. Version 2009.1. Available via http://www.iucn-uk.org. Accessed 22/05/2009. International Union for the Conservation of Nature, Gland, Switzerland.
- Jackson, D.L. 2000. Guidance on the Interpretation of the Biodiversity Broad Habitat Classification (Terrestrial and Freshwater Types): Definitions and the Relationship with Other Habitat Classifications, JNCC Report No. 307, Joint Nature Conservancy Committee, Peterborough.
- Jackson, S.F. and Gaston, K.J. 2008. The unpredictability of favourability: condition assessment and protected areas in England. Biodiversity Conservation 17, 749-764.
- Jeeves, M. 2007. Targets in nature conservation- sending the wrong signals? ECOS: A Review of Conservation 28, 7-10.
- JNCC 1998. Statement on Common Standards Monitoring (CSM). Available via http://www.jncc.gov.uk/. Accessed 17/04/2006. Joint Nature Conservation Committee, Peterborough.
- JNCC 2004a. Common Standards Monitoring Guidance for Sand Dune Habitats. Available via http://www.jncc.gov.uk/. Accessed 15/02/2008. Joint Nature Conservation Committee, Peterborough.
- JNCC 2004b. Common Standards Monitoring Guidance for Woodland Habitats. Available via http://www.jncc.gov.uk/. Accessed 10/06/2006. In. Joint Nature Conservation Committee, Peterborough.
- JNCC 2004c. Common Standards Monitoring: Introduction to the Guidance Manual. Available via http://www.jncc.gov.uk/. Accessed 25/06/2006. Joint Nature Conservation Committee, Peterborough.
- JNCC 2004d. Handbook for Phase 1 habitat survey a technique for environmental audit. Joint Nature Conservation Committee, Peterborough.
- JNCC 2005a. Common Standards Monitoring Guidance for Upland Habitats. Available via http://www.jncc.gov.uk/. Accessed 21/03/2006. Joint Nature Conservation Committee, Peterborough.
- JNCC 2005b. Guidance for Common Standards Monitoring (CSM). Available via http://www.jncc.gov.uk/. Accessed 16/03/2006. Joint Nature Conservation Committee, Peterborough.
- JNCC 2006. Common Standards Monitoring (CSM). Available via http://www.jncc.gov.uk/. Accessed 15/07/2007. Joint Nature Conservation Committee, Peterborough.
- JNCC 2008a. Key Aspects of Common Standards Monitoring. Available via http://www.jncc.gov.uk/. Accessed 20/04/2009. Joint Nature Conservation Committee, Peterborough.
- JNCC 2008b. The National Vegetation Classification. Available via http://www.jncc.gov.uk/. Accessed 20/10/2008. Joint Nature Conservation Committee, Peterborough.
- JNCC 2008c. Report on Vegetation Sampling Workshop January 2008. Available via http://www.jncc.gov.uk/. Accessed 20/08/2008. Joint Nature Conservation Committee, Peterborough.

- JNCC 2009a. 1.2. Habitats and species of European Interest in the UK. Available via http://www.jncc.gov.uk/. Accessed 23/10/2009. Joint Nature Conservation Committee, Peterborough.
- JNCC 2009b. Surveillance and monitoring. Available via http://www.jncc.gov.uk/. Accessed 03/05/2009. Joint Nature Conservation Committee, Peterborough.
- JNCC 2009c. Tracking Mammals Partnership: Single Species Schemes. Available via http://www.jncc.gov.uk/. Accessed 22/20/2008. Joint Nature Conservation Committee, Peterborough.
- Jones, J.P.G., Andriamarovololona, M.M., Hockley, N., Gibbons, J.M. and Milner-Gulland, E.J. 2008. Testing the use of interviews as a tool for monitoring trends in the harvesting of wild species. Journal of Applied Ecology 45, 1205-1212.
- Jones, L.L. 1967. Studies on hill land in Wales. Technical Bulletin. Welsh Plant Breeding Station, Aberystwyth.
- Jongman, R.H.G., ter Braak, C.J.F. and van Tongeren, O.F.R. 1995. Data Analysis in Community and Landscape Ecology. Cambridge University Press, Cambridge.
- Joseph, L.N. and Possingham, H.P. 2008. Grid-based monitoring methods for detecting population declines: Sensitivity to spatial scale and consequences of scale correction. Biological Conservation 141, 1868-1875.
- Journel, A.G. and Huijbregts, C.J. 1978. Mining Geostatistics. Academic Press, London.
- Keitt, T.H., Urban, D.L. and Milne, B.T. 1997. Detecting critical scales in fragmented landscapes. Conservation Ecology 1, 4.
- Kenkel, N.C. and Podani, J. 1991. Plot size and estimation efficiency in plant community studies. Journal of Vegetation Science, 539-544.
- Kennedy, K.A. and Addison, P.A. 1987. Some considerations for the use of visual estimates of plant cover in biomonitoring. Journal of Ecology 75, 151-157.
- Kent, M. 2006. Numerical classification and ordination methods in biogeography. Progress in Physical Geography 30, 399-408.
- Kent, M. 2007. Biogeography and landscape ecology. Progress in Physical Geography 31, 1-11.
- Kent, M. and Coker, P. 1992. Vegetation, Description and Analysis: A Practical Approach. John Wiley and Sons, Chichester.
- Kent, M., Gill, W.J., Weaver, R.E. and Armitage, R.P. 1997. Landscape and plant community boundaries in biogeography. Progress in Physical Geography 21, 315-353.
- Kent, M., Moyeed, R.A., Reid, C.L., Pakeman, R. and Weaver, R. 2006. Geostatistics, spatial rate of change analysis and boundary detection in plant ecology and biogeography. Progress in Physical Geography 30, 201-231.
- Kercher, M.S., Frieswyk, B.C., Zedler B.J. 2003. Effects of sampling teams and estimation methods on the assessment of plant cover. Journal of Vegetation Science 14, 899-906.
- Kirby, K.J. 2002. A trial of consistency of condition assessment procedures for woodland (unpublished).
- Kirby, K.J. 2003. NVC in Semi-Natural Woodland. In: E. Goldberg, ed. National Vegetation Classification- Ten Years' Experience Using the Woodland Section. Joint Nature Conservation Committee, Peterborough; pp. 7-14.
- Kirk, R.E. 2007. Effect magnitude: a different focus. Journal of Statistical Planning and Inference 137, 1634-1646.

- Kirkham, F.W. 2001 Nitrogen uptake and nutrient limitation in six hill moorland species in relation to atmospheric nitrogen deposition in England and Wales. Journal of Ecology 89, 1041-1053
- Klimes, L. 2003. Scale-dependent variation in visual estimates of grassland plant cover. Journal of Vegetation Science 14, 815-821.
- Klimeš, L., Dančák, M., Hájek, M., Jongepierová, I. and Kučera, T. 2001. Scaledependent biases in species counts in a grassland. Journal of Vegetation Science 12, 699-704.
- Koenig, W.D. and Knops, J.M.H. 1998. Testing for spatial autocorrelation in ecological studies. Ecography 21, 423-429.
- Kuchler, A.W. 1967. Vegetation Mapping. Ronald Press, New York.
- Kuchler, A.W. 1973. Problems in classifying and mapping vegetation for ecological regionalization. Ecology 54, 512-523.
- Kunin, W.E. 1998. Extrapolating species abundance across spatial scales. Science 281, 321-324.
- Kunin, W.E., Hartley, S. and Lennon, J.J. 2000. Scaling down: On the challenge of estimating abundance from occurrence patterns. American Naturalist 156, 560-566.
- Lavorel, S., Grigulis, K., McIntyre, S., Williams, N.S.G., Garden, D., Dorrough, J., Berman, S., Quétier, F., Thébault, A. and Bonis, A. 2008. Assessing functional diversity in the field - methodology matters! Functional Ecology 22, 134-147.
- Legendre, P. 1993. Spatial autocorrelation: trouble or new paradigm? Ecology 74, 1659-1673
- Legendre, P. and Anderson, M.J. 1999. Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. Ecological Monographs 69, 1-24.
- Legendre, P., Dale, M.R.T., Fortin, M.-F., Gurevitch, J., Hohn, M. and Myers, D. 2002. The consequences of spatial structure for the design and analysis of ecological field surveys. Ecography 25, 601-615.
- Legendre, P., Dale, M.R.T. and Fortin, M.-J. 2004. Effects of spatial structures on the results of field experiments. Ecology 85, 3202-3214.
- Legendre, P. and Fortin, M.-J. 1989. Spatial pattern and ecological analysis. Vegatatio 80, 107-138.
- Legendre, P. and Gallagher, E.D. 2001. Ecologically meaningful transformations for ordination of species data. Oecologia 129, 271-280.
- Legg, C.J. and Nagy, L. 2006. Why most conservation monitoring is, but need not be, a waste of time. Journal of Environmental Management 78, 194-199.
- Lengyel, S., Deri, E., Varga, Z., Horvath, R., Tothmeresz, B., Henry, P.-Y., Kobler, A., Kutnar, L., Babij, V., Seliskar, A., Christa, C., Papastergiadou, E., Gruber, B. and Henle, K. 2008a. Habitat monitoring in Europe: a description of current practices. Biodiversity Conservation 17, 3327-3339.
- Lengyel, S., Kobler, A., Kutnar, L., Framstad, E., Henry, P.-Y., Babij, V., Gruber, B., Schmeller, D. and Henle, K. 2008b. A review and a framework for the integration of biodiversity monitoring at the habitat level. Biodiversity Conservation 17, 3341-3356.
- Lennon, J.J. 2000. Red-shifts and red herrings in geographical ecology. Ecography 23, 101-113.
- Lepš, J. and Hadincová, V. 1992. How reliable are our vegetation analyses? Journal of Vegetation Science 3, 119-124.

- Light, S.S. 2001. Adaptive ecosystem assessment and management: the path of last resort? In: M. E. Jensen and P. S. Bourgeron, eds. A Guidebook for Integrated Ecological Assessments. Springer, New York; pp. 55-68.
- Likens, G.E. (Ed.), 1991. Long-term Studies in Ecology: Approaches and Alternatives. Springer- Verlag, New York.
- Lindenmayer, D.B. 1999. Future directions for biodiversity conservation in managed forests: indicator species, impact studies and monitoring programs. Forest Ecology and Management 115, 277-287.
- Linke, S. and Norris, R. 1993. Biodiversity: bridging the gap between condition and conservation. Hydrobiologia 500, 203-211.
- Lister, A.J., Mou, P.P., Jones, R.H. and Mitchell, R.J. 2000. Spatial patterns of soil and vegetation in a 40-year-old slash pine (*Pinus elliottii*) forest in the Coastal Plain of South Carolina, U.S.A. Canadian Journal of Forest Research 30, 145-155.
- Loh, J., Green, R.E., Ricketts, T., Lamoreux, J., Jenkins, M., Kapos, V. and Randers, J. 2005. The Living Planet Index: using species population time series to track trends in biodiversity. Philosophical Transactions of the Royal Society B, 289-295.
- Lundstrom, S. 2000. Plant cover estimates in long-term monitoring- how to assess its precision. MSc dissertation. Department of Environmental Assessment, University of Agricultural Sciences, Uppsala.
- Malloch, A.J.C. 1990. MATCH, a computer program to aid the assignment of vegetation data to the communities and sub-communities of the National Vegetation Classification. University of Lancaster, Lancaster.
- Mandelbrot, B.B. 1983. The Fractal Geometry of Nature. Freeman, New York.
- MapInfo 2004. MapInfo Professional Version 7.8.
- Marchetto, A., Mosello, R., Tartari, G., Tornimbeni, O., Derome, J., Derome, K., Sorsa, P., Konig, N., Clarke, N., Ulrich, E. and Kowalsk, A. 2009. Influence of QA/QC procedures on non-sampling error in deposition monitoring in forests. Journal of Environmental Monitoring 11, 745–750.
- Mark, D.M. and Csillag, F. 1989. The nature of boundaries on 'area-class' maps. Cartographica: The International Journal for Geographic Information and Geovisualization 26, 65-78.
- Marriott, C.A., Fothergill, M., Jeangros, B., Scotton, M. and Louault, F. 2004. Longterm impacts of extensification of grassland management on biodiversity and productivity in upland areas. A review. Agronomie 24, 447-462.
- Maskell, L.C., Norton, L.R., Smart, S.M., Carey, P.D., Murphy, J., Chamberlain, P.M., Wood, C.M., Bunce, R.G.H. and Barr, C.J. 2008. CS Technical Report No.1/07 Field Mapping Handbook. NERC/Centre for Ecology and Hydrology, 143 pp. (CEH Project Number: C03259).
- McCune, B., Grace, J.B. and Urban, D.L. 2002. PCPOrd. Analysis of ecological communities. Mjm Software Design, Gleneden Beach, Oregon.
- McGaw, B. and Glass, G.V. 1980. Choice of the metric for effect size in metaanalysis. American Educational Research Journal 17, 325-337.
- Milberg, P., Bergstedt, J., Fridman, J., Odell, G. and Westerberg, L. 2008. Observer bias and random variation in vegetation monitoring data. Journal of Vegetation Science 19, 633-644.
- Millington, A.C. and Alexander, R.W. 2000. Vegetation Mapping in the Last Three Decades of the Twentieth Century. In: R. W. Alexander and A. C. Millington, eds. Vegetation Mapping. John Wiley and Sons, Chichester; pp. 321-331.

- Milne, J.A. and Hartley, S.E. 2001. Upland plant communities sensitivity to change. Catena 42, 333-343.
- Minium, E.W., Clarke, R.C. and Coladarci, T. 1999. Elements of Statistical Reasoning. John Wiley and sons, New York.
- Mittelbach, C.G., Steiner, C.F. and Scheiner, S.M. 2001. What is the observed relationship between species richness and productivity? Ecology 82, 2381-2396.
- Montgomery, D.R. and Buffington, J.M. 1993. Channel classification, prediction of channel response, and assessment of channel condition. Report TFW-SI-110-93-002, prepared for the SHAMW committee of the Washington State Timber/Fish/Wildlife Agreement. Department of Geological Sciences and Quaternary Research Center, University of Washington, Seattle.
- Mueller-Dombois, D. and Ellenberg, H. 2002. Aims and Methods of Vegetation Ecology. Blackburn Press, Caldwell, USA.
- Nakagawa, S. and Cuthill, I.C. 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. Biological Reviews 82, 591-605.
- Network, T.E.B.E. 2008. Effect Size Calulator. The Curriculum, Evaluation and Management Centre, Durham University, Durham, England.
- Nichols, J.D. and Williams, B.K. 2006. Monitoring for conservation. Trends in Ecology and Evolution 21, 668- 673.
- NIEA 2008. The Condition of Northern Ireland's Areas of Special Scientific Interest: the Results of the First Condition Assessment Monitoring Cycle 2002-2008. A report by the Northern Ireland Environment Agency (2008), Research and Development Series No. 08/10. Available via http://www.nienvironment.gov.uk. Accessed 10/10/2008. Northern Ireland Environment Agency, Belfast.
- Nilsson, I.N. and Nilsson, S.G. 1985. Experimental estimates of census efficiency and pseudoturnover on islands: error trend and between-observer variation when recording vascular plants. Journal of Ecology 73, 65-70.
- Nimis, P.L. 2001. A tale from bioutopia. Nature 413, 21-21.
- Norton, L., Scholefield, P., Maskell, L., Smart, S., Murphy, J., Williams, P., Scarlett, P. and Clarke, R. 2008. Quality Assurance Report: Mapping Quality Assurance Exercise. NERC/Centre for Ecology and Hydrology, 25 pp.
- Noss, R.F. 1990. Indicators for monitoring biodiversity- a hierarchical approach. Conservation Biology 4, 355-364.
- Noss, R.F. 1999. Assessing and monitoring forest biodiversity: A suggested framework and indicators. Forest Ecology and Management 115, 135-146.
- NT 2000. Nature Conservation Evaluation: Hafod y Llan, Gwynedd. The National Trust, Estate Department.
- NT 2004. Hafod y Llan, Nature Conservation Monitoring Report 2004. The National Trust, Craflwyn.
- Oldham, M.J., Bakowsky, W.D. and Sutherland, D.A. 1995. Floristic quality assessment system for southern Ontario. Natural Heritage Information Centre, Ontario Ministry of Natural Resources, Peterborough, Ontario.
- Oliver, I. 2002. An expert panel-based approach to the assessment of vegetation condition within the context of biodiversity conservation: Stage 1: the identification of condition indicators. Ecological Indicators 2, 223-237.
- Oliver, M.A. and Frogbrook, Z.L. 1998. Sampling to estimate soil nutrients for precision agriculture. The International Fertiliser Society conference Proceedings no. 417, Cambridge.

- Packwood, A. 1991. A critical assessment of the National Vegetation Classification (grasslands), MSc dissertation. School of Biological Sciences, University of Wales, Bangor.
- Pakeman, R.J., Hulme, P.D., Torvell, L. and Fisher, J.M. 2003. Rehabilitation of degraded dry heather (*Calluna vulgaris* (L) Hull) moorland by controlled sheep grazing. Biological Conservation 114.
- Pakeman, R.J. and Torvell, L. 2008. Identifying suitable restoration sites for a scarce subarctic willow (*Salix arbuscula*) using different information sources and methods. Plant Ecology and Diversity 1, 105-114.
- Palmer, M. 1992. Trial of MATCH and TABLEFIT computer programs for placing survey data within the National Vegetation Classification. JNCC report No. 20. Joint Nature Conservation Committee, Peterborough.
- Palmer, M.W. 1993. Putting things in even better order: the advantages of canonical correspondence analysis. Ecology 74, 2215-2230.
- Palmer, M.W. 2009. Ordination Methods an overview. Available via http://ordination.okstate.edu/. Accessed 2/02/2009. Oklahoma State University, Oklahoma.
- Palmer, M.W. and White, P.S. 1994. On the existence of communities. Journal of Vegetation Science 5, 279-282.
- Palmer, S.C.F., Gordon, I.J., Hester, A.J. and Pakeman, R.J. 2004. Introducing spatial grazing impacts into the prediction of moorland vegetation dynamics. Landscape Ecology 19, 817-827.
- Parkes, D., Graeme, N. and Cheal, D. 2003. Assessing the quality of native vegetation: the 'habitat hectares' approach. Ecological Management and Restoration 4, 29-38.
- Partnership, U.B. 2009. UK Biodiversity Indicators in Your Pocket 2009: Measuring Progress Towards Halting Biodiversity Loss. Available via http://www.jncc.gov.uk/. Accessed 29/03/2009. Department for Environment, Food and Rural Affairs, London.
- Pearce, J. and Ferrier, S. 2000. Evaluating the predictive performance of habitat models developed using logistic regression. Ecological Modelling 133, 225-245.
- Pereira, H.M. and Cooper, H.D. 2006. Towards the global monitoring of biodiversity change. Trends in Ecology and Evolution 21, 123-129.
- Podani, J., Csontos, P., Tama, J. and Miklo, I. 2005. A new multivariate approach to studying temporal changes of vegetation. Plant Ecology 181, 85-100.
- Porley, R. and Hodgetts, N. 2005. Mosses and Liverworts. Collins, London.
- Preston, C.D., Pearman, D.A. and Dines, T.D. 2002. New Atlas of the British and Irish Flora. Oxford University Press, Oxford.
- Prosser, M. and Wallace, H. 1999. Countryside Survey 2000: A Quality Assurance Exercise. Report to the Institute of Terrestrial Ecology, Merlewood Research Station. Ecological Surveys (Bangor), Herefordshire.
- Pywell, R.P., Bullock, J.M., Roy, D.B., Warman, L., Walker, K. and Rothery, P. 2003. Plant traits as predictors of performance in ecological restoration. Journal of Applied Ecology 40, 65-77.
- Ramsay, P.M., Kent, M., Reid, C.L. and Duckworth, J.C. 2006. Taxonomic, morphological and structural surrogates for the rapid assessment of vegetation. Journal of Vegetation Science 17, 747-754.
- Ringold, P.L., Alegria, J., Czaplewski, R.L., Mulder, B.S., Tolle, T., Burnett, K., Stanford, J.A. and Poole, G.C. 1996. Adaptive monitoring design for

ecosystem management: a protocol for ecosystem management. Ecological Applications 6, 745-747.

- Ringvall, A., Petersson, H., Stahl, G., Lamas and T. 2005. Surveyor consistency in presence/absence sampling for monitoring vegetation in a boreal forest. Forest Ecology and Management 212, 109-117.
- Ritters, K.H., Law, B.E., Kucera, R.C., Galllant, A.L., DeVelice, R.L. and Palmer, C.J. 1992. A selection of forest condition indicators for monitoring. Environmental Monitoring and Assessment 20, 21-23.
- Robertson, G.P. 1987. Geostatistics in ecology: interpolating with known variance. Ecology 68, 744-748.
- Robertson, H.J. and Jefferson, R.G. 2000. Monitoring the Condition of Lowland Grassland SSSIs I: English Nature's Rapid Assessment Methods. English Nature Research Report, Peterborough.
- Rodwell, J.S. (Ed.), 1991a. British Plant Communities Volume 1: Woodlands and Scrub. Cambridge University Press, Cambridge.
- Rodwell, J.S. (Ed.), 1991b. British Plant Communities Volume 2: Mires and Heaths. Cambridge University Press, Cambridge.
- Rodwell, J.S. (Ed.), 1992. British Plant Communities Volume 3: Grasslands and Montane Communities. Cambridge University Press, Cambridge.
- Rodwell, J.S. (Ed.), 1995. British Plant Communities Volume 4: Aquatic Communities, Swamp and Tall-Herb Fens. Cambridge University Press, Cambridge.
- Rodwell, J.S. 1997. The NVC and monitoring. CCW Science Report No. 200. Countryside Council for Wales, Bangor.
- Rodwell, J.S. (Ed.), 2000. British Plant Communities Volume 5: Maritime communities and vegetation of open habitats. Cambridge University Press, Cambridge.
- Rodwell, J.S. 2006. National Vegetation Classification: User's handbook. Joint Nature Conservation Committee, Peterborough.
- Rose, F. 2006. The Wild Flower Key. How to identify wild flowers, trees and shrubs in Britain and Ireland. Revised and updated by C.O'Reilly. Frederick Warne, London.
- Rosenzweig, M.L. 1995. Species Diversity in Space and Time. Cambridge University Press, Cambridge.
- Ross, S. and Bealey, C.E. 2006. Validation Network Project. Lowland heathland monitoring: covering dry and wet heaths. English Nature Research Report No. 669. English Nature, Peterborough.
- Ross, S., Bealey, C.E. and Cox, J. 2004. Validation Network Project. Lowland calcareous grasslands: *Bromus erectus* (CG3) grasslands. English Nature Research Report No. 609. English Nature, Peterborough.
- Rowe, J.S. 1996. Land classification and ecosystem classification. Environmental Monitoring and Assessment 39, 11-20.
- Rowell, T. 1996. Development of monitoring in the UK. Paper at conference 'Monitoring as an integral component of nature conservation planning and management' held at Plas Tan y Bwlch, Snowdonia National Park.
- Rowell, T.A. 1993a. Common Standards for Monitoring SSSIs: Executive Summary. Joint Nature Conservation Committee, Peterborough.
- Rowell, T.A. 1993b. Common Standards for Monitoring SSSIs: Report Volume I. Joint Nature Conservation Committee, Peterborough.

- RS 2003. Measuring Biodiversity for Conservation Policy document 11/03. The Royal Society, London.
- Ruxton, G.D. and Colegrave, N. 2006. Experimental Design for the Life Sciences. Oxford University Press, Oxford.
- Ryti, R.T. 1992. Effect of the focal taxon on the selection of nature reserves. Ecological Applications 2, 404-410.
- Sabine, E., Schreiber, G., Bearlin, A.R., Nicol, S.J. and Todd, C.R. 2004. Adaptive management: a synthesis of current understanding and effective application. Ecological Management and Restoration 5, 177-182.
- Sala, O.E. 2006. Biodiversity Across Scenarios. In: S. Carpenter and P. Pingali, eds. Ecosystems and Human Well-Being: Volume 2 Scenarios: Findings of the Scenarios Working Group (Millennium Ecosystem Assessment Series). Island Press and Millennium Ecosystem Assessment.
- Salafsky, N., Margoluis, R., Redford, H., K., Robinson and G., J. 2002. Improving the practice of conservation: a conceptual framework and research agenda for conservation science. Conservation Biology 16, 1469-1479.
- Schiller, A., Hunsaker, C.T., Kane, M.A., Wolfe, A.K., Dale, V.H., Suter, G.W., Russell, C.S., Pion, G., Jensen, M.H. and Konar, V.C. 2001. Communicating ecological indicators to decision makers and the public. Conservation Ecology 5, 19-28.
- Schlup, B.M. and Wagner, H.H. 2008. Effects of study design and analysis on the spatial community structure detected by multiscale ordination. Journal of Vegetation Science 19, 621-632.
- Schreuder, H.T., Ernst, R. and Ramirez-Maldonado, H. 2004. Statistical Techniques for Sampling and Monitoring Natural Resources. General Technical Report RMRS-GTR-126. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, Colorado.
- Schumaker, N.H. 1996. Using landscape indices to predict habitat connectivity. Ecology 77, 1210-1225.
- Scott, A. and Smart, S. 2006. Analysis of vegetation data from different quadrat sizes. English Nature Research Report (unpublished). English Nature, Peterborough.
- Scott, C.T. 1998. Sampling methods for estimating change in forest resources. Ecological Applications 8, 228-233.
- Scott, W.A. and Hallam, C.J. 2003. Assessing species misidentification rates through quality assurance of vegetation monitoring. Plant Ecology 165, 101-115.
- Seavy, N.E. and Reynolds, M.H. 2007. Is statistical power to detect trends a good assessment of population monitoring? Biological Conservation 140, 187-191.
- Shaw, P. and Wind, P. 1997. Monitoring the condition and biodiversity status of European conservation sites. Paper presented to the European Environment Agency on behalf of the European Topic Centre on Nature Conservation, Paris.
- Sheil, D. 1995a. A critique of permanent plot methods and analysis with examples from Budongo Forest, Uganda. Forest Ecology and Management 77, 11-34.
- Sheil, D. 1995b. A note on sampling and monitoring vegetation for biodiversity: questions, possibilities, and practicalities. Paper presented at: Experts Meeting on Identifying and Monitoring Biodiversity and its Utilisation in Commonwealth Small Island Developing States.

- Sheil, D., Jennings, S. and Savill, P. 2000. Long-term permanent plot observations of vegetation dynamics in Budongo, a Ugandan rain forest. Journal of Tropical Ecology 16, 765-800.
- Simberloff, D. 1998. Flagships, umbrellas, and keystones: Is single-species management passe in the landscape era? Biological Conservation 83, 247-257.
- Slater, J.A., Mole, A.C. and Warring, R. 2006. Strategic Analysis of UK Environmental Monitoring Activity: Report to the Environmental Research Funders' Forum. ADAS, London.
- Smartt, P.F.M. 1978. Sampling for vegetation survey: a flexible systematic model for sample location. Journal of Biogeography 5, 43-56.
- Smit, A.M. 2003. Adaptive Monitoring: An Overview. DOC Science Internal Series. Department of Conservation, Wellington, New Zealand.
- SNH 2008. Scottish Government Condition Targets: Progress Towards the 2010 Target. SNH/08/4/3. Scottish Natural Heritage, Edinburgh.
- Spellerberg, I. 1992. Evaluation and Assessment for Conservation. Chapman and Hall, London.
- Spellerberg, I. 2005. Monitoring Ecological Change. Cambridge University Press, Cambridge.
- SPSS 2003. Statistical Package for Social Scientists Version 12.0.1 for Windows. The Apache Software Foundation, USA.
- Stefano, J.D. 2003. How much power is enough? Against the development of an arbitrary convention for statistical power calculations. Functional Ecology 17, 707-709.
- Stevens, C.J., Dise, N.B., Mountford, J.O. and Gowing, D.J. 2004a. Impact of nitrogen deposition on the species richness of grasslands Science 303, 1876 1879.
- Stevens, J.P., Blackstock, T.H., Howe, E.A. and Stevens, D.P. 2004b. Repeatability of Phase 1 habitat survey. Journal of Environmental Management 73, 53-59.
- Stewart, K.E.J., Bourn, N.A.D., Thomas and J.A. 2001. An evaluation of three quick methods commonly used to assess sward height in ecology. Journal of Applied Ecology 38, 1148-1154.
- Stohlgren, T.J. 2007. Measuring Plant Diversity: Lessons from the Field. Oxford University Press, New York.
- Stohlgren, T.J., Binkley, D., Veblen, T.T. and Baker, W.L. 1995. Attributes of reliable long-term landscape-scale studies: malpractice insurance for landscape ecologists. Environmental Monitoring and Assessment 36, 1-25.
- Stohlgren, T.J., Bull, K.A. and Otsuki, Y. 1998. Comparison of rangeland vegetation sampling techniques in the Central Grasslands. Journal of Range Management 51, 164-172.
- Stohlgren, T.J., Coughenour, M.B., Chong, G.W., Binkley, D., Kalkhan, M.A., Schell, L.D., Buckley, D.J. and Berry, J.K. 1997. Landscape analysis of plant diversity. Landscape Ecology 12.
- Storch, D. and Gaston, K.J. 2004. Untangling ecological complexity on different scales of space and time. Basic and Applied Ecology 5, 389-400.
- Strachan, I. and Jackson, D. 2003. Review and development of the National Vegetation Classification: stability and change. In: E. Goldberg, ed. National Vegetation Classification - Ten years' experience using the woodland section. Joint Nature Conservation Committee, Peterborough; pp. 87-92. JNCC report.

Sutherland, W.J. 2000. The Conservation Handbook: Research, Management and Policy. Blackwell Science, Oxford.

- Svátek, M. and Buček, A. 2007. Methodology of evaluation of state and management of small-scale protected areas. English version of the research report prepared for the Ministry of Environment of the Czech Republic. 2nd Edition. Mendel University of agriculture and forestry, Czech Republic, Brno.
- Sykes, J.M., Horrill, A.D. and Mountford, M.D. 1983. Use of visual cover assessments as quantitative estimators of some British woodland taxa. Journal of Ecology 71, 437-450.
- Sykes, J.M. and Lane, A.M.J. (Eds.), 1996. The United Kingdom Environmental Change Network: Protocols for Standard Measurements at Terrestrial Sites. Stationery Office, London.
- Szaro, R.C., Angelstam, P. and Sheil, D. 2005. Information Needs for Ecosystem Forestry. In: J. Sayer and S. Maginnis, eds. Forests in Landscapes: Ecosystem Approaches to Sustainability. Earthscan, London; pp. 158-190.
- Taverna, K., Peet, R. and Phillips, L. 2005. Long-term change in ground-layer vegetation of deciduous forests of the North Carolina Piedmont. USA Journal of Ecology 93, 202-213.
- Teder, T., Moora, M., Roosaluste, E., Zobel, K., Partel, M., Oljalg, K. and Zobel, M. 2007. Monitoring of biological diversity: a common-ground approach. Conservation Biology 21, 313-317.
- Ter Braak, C.J.F. 1995. Ordination. In: R. H. G. Jongman, C. J. F. ter Braak and O. F. R. van Tongeren, eds. Data Analysis in Community and Landscape Ecology. Cambridge University Press, Cambridge; pp. 19-31.
- Thomas, L. and Meyer, J. 1996. Monitoring long-term population change: why are there so many analysis methods? Ecology 77, 49-58.
- Thompson, M. 1999. Sampling: the uncertainty that dare not speak its name. Journal of Environmental Monitoring 1, 19N-21N.
- Tobin, P.C. 2004. Estimation of the spatial autocorrelation function: consequences of sampling dynamic populations in space and time. Ecography 27, 767-775.
- Tobler, W. 1970. A computer movie simulating urban growth in the Detroit regions. Economic Geography 46, 234- 240.
- Treweek, J.R. 1996. Ecology and Environmental Impact Assessment. Journal of Applied Ecology 33, 191-199
- Trost, J.E. 1986. Statistically nonrepresentative stratified sampling: A sampling technique for qualitative studies. Qualitative Sociology 9.
- Tucker, G., Bubb, P., deHeer, M., Miles, L., Lawrence, A., Bajracharya, S.B., Nepal, R.C., Sherchan, R. and Chapagain, N.R. 2005. Guidelines for Biodiversity Assessment and Monitoring for Protected Areas. King Mahendra Trust for Nature Conservation, Kathmandu, Nepal.
- Turner, A. 2008. Conversations about the National Vegetation Classification: field methodology and assignment of type. Hearn, S.M., Bangor.
- Turner, M.G., Gardner, R.H. and O'Neill, R.V. 2001. Landcape Ecology in Theory and Practice: Pattern and Process. Springer Verlag, New York.
- Turnhout, E., Hisschemoller, M. and Eijsackers, H. 2007. Ecological indicators: Between the two fires of science and policy. Ecological Indicators 7, 215-228.
- UKBAP 2008. UK Biodiversity Action Plan. Available via http://ukbap.org.uk/. Accessed 20/02/2008. UK Biodiversity Partnership.

- UKBG 1998. Tranche two action plans. Volume II: Terrestrial and freshwater habitats. Peterborough, published by English Nature on behalf of the UK Biodiversity Group.
- UKBMS 2009. UK Butterfly Monitoring Scheme. Available via: http://www.ukbms.org/. Accessed 12/10/2009.
- UKCHM 2009. UK Clearing House Mechanism for Biodiversity. Available via: http://uk.chm-cbd.net/. Accessed 12/10/2009.
- Underwood, A.J. 1995. Ecological research and (research into) environmental management. Ecological Applications 5, 232-247.
- UNEP 2002. COP 6, Decision VI/26 Strategic Plan for the Convention on Biological Diversity. Available via http://www.cbd.int/. Accessed 11/04/2006. United Nations Environment Programme, The Hague.
- UNEP 2009. About UNEP-WCMC. Available via http://www.unep-wcmw.org. Accessed 28/04/2009. United Nations Environment Programme, World Conservation Monitoring Centre, Cambridge.
- Urlich, S. and Brady, P. 2003. Monitoring Terrestrial Habitats in Wellington Conservancy: A Strategy for 2003-2012. Department of Conservation, Wellington, New Zealand.
- Van Groenewoud, H. 1992. Robustness of Correspondence, Detrended Correspondence, and TWINSPAN Analysis. Journal of Vegetation Science 3, 239-246.
- Van Strien, A.J., van de Pavert, R., Moss, D., Yates, T.J., van Swaay, C.A.M. and Vos, P. 1997. The statistical power of two butterfly monitoring schemes to detect trends. Journal of Applied Ecology 34, 817-828.
- Vaughan, H., Brydges, T., Fenech, A. and Lumb, A. 2001. Monitoring long-term ecological changes through the ecological monitoring and assessment network: science-based and policy relevant. Environmental Monitoring and Assessment 67, 3-28.
- Vittoz, P. and Guisan, A. 2007. How reliable is the monitoring of permanent vegetation plots? A test with multiple observers. Journal of Vegetation Science 18, 413-422.
- Vos, P., Meelis, E. and Keurs, J.W.T. 2000. A framework for the design of ecological monitoring programs as a tool for environmental and nature management. Environmental Monitoring and Assessment 61, 317-344.
- WAG 2006. Environment Strategy for Wales. Welsh Assembly Government, Cardiff.
- WAG 2008. State of the Environment: Biodiversity Indicators. Percentage of Habitats in Favourable, Recovering or Unfavourable Condition in Wales, 2000-2006. Available via http://www.statswales.wales.gov.uk. Accessed 12/02/2009. Welsh Assembly Government, Cardiff.
- Wagner, H.H. 2003. Spatial covariance in plant communities: integrating ordination, geostatistics, and variance testing. Ecology 84, 1045–1057.
- Wagner, H.H. 2004. Direct multi-scale ordination with canonical correspondence analysis. Ecology, 342-351.
- Wagner, H.H. and Fortin, M.-J. 2005. Spatial analysis of landscapes: concepts and statistics. Ecology 86, 1975-1987.
- Wallace, A., C.S., Watts, M., J., Yool and R., S. 2000. Characterizing the spatial structure of vegetation communities in the Mojave Desert using geostatistical techniques. Computers & Geosciences 26, 397-410.

- Warren, R.D. and Witter, M.S. 2002. Monitoring trends in bat populations through roost surveys: methods and data from *Rhinolophus hipposideros*. Biological Conservation 105, 255-261.
- Webster, R., Oliver and M.A. 1992. Sample adequately to estimate variograms of soil properties. Journal of Soil Science 43, 177-192.
- Webster, R. and Oliver, M. 2001. Geostatistics for Environmental Scientists. John Wiley and Sons, Chichester.
- Welch, D. 1986. Studies in the grazing of heather moorland in North-East Scotland V. Trends in *Nardus stricta* and other unpalatable graminoids. Journal of Applied Ecology 23, 1047-1058.
- Wiens, J.A. 1989. Spatial scaling in Ecology. Functional Ecology 3, 385-397.
- Williams, B.K. 1996. Assessment of accuracy in the mapping of vertebrate biodiversity. Journal of Environmental Management 47, 269-282.
- Wilson, S.M. 2003. Use of the National Vegetation Classification (woodland section) for site assessment in British plantation forests. In: E. Goldberg, ed. National Vegetation Classification- Ten Years' Experience Using the Woodland Section. Joint Nature Conservation Committee, Peterborough; pp. 31-37.
- Withers, M.A. and Meentemeyer, V. 1999. Concepts of Scale in Landscape Ecology. In: J. M. Klopatek and R. H. Gardner, eds. Landscape Ecological Analysis: Issues and applications. Springer, New York.
- Wong, J., Payne, N. and Boden, R. 2005. Non-Wood Forest Products: Resource Assessment Guidelines. UK Department for International Development, London.
- Wyn, G., Brazier, P., Birch, K., Bunker, A., Cooke, A., Jones, M., Lough, N., McMath, A. and Roberts, S. 2006. Handbook for Marine Intertidal Phase 1 Biotope Mapping Survey. Countryside Council for Wales, Bangor.
- Yoccoz, N.G., Nichols, J.D. and Boulinier, T. 2001. Monitoring of biological diversity in space and time. Trends in Ecology and Evolution 16, 446-453.
- Zamora, J., Abraira, V., Muriel, A., Khan, K.S. and Coomarasamy, A. 2006. Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Medical Research Methodology 6.
- Zimmerman, D.L. 1993. Another look at anisotropy in geostatistics. Mathematical Geology 25, 453-470.
- Zou, K.H., Hall, W.J. and Shapiro, D.E. 1998. Smooth non-parametric receiver operating characteristic (ROC) curves for continuous diagnostic tests. Statistics in Medicine 16, 2143 - 2156.
- Zweig, M.H. and Campbell, G. 1993. Receiver operating characteristic (ROC) plots: A fundamental evaluation tool in Clinical Medicine. Clinical Chemistry 39, 561-577.

Appendices

Appendix 2.1 Questionnaire used as the basis for semi-structured interviews with conservation practitioners in the UK between 2006 and 2008.

Monitoring: what do you think? Monitoring- methods which show whether species, habitats or environmental variables meet targets defined by some standard, such as objectives of site management (Hellawell, 1991).

Please take a few minutes to answer some questions.

Respondent information

Name	Organisation		
Gender	Position		
Stakeholder group Policy maker/ Land manager/ Consultant/ Volunteer/ Researcher/ Advisor			

Survey auestions

PERSONAL C		
1.	a. What level of conservation management are you involved with?	Local/ Regional/ National
	band what type?	Taxa/ Habitat/ Landscape
2.	Does your work involve any monitoring of terrestrial vegetation?	Please describe this monitoring (using the matrix overleaf)
3.	Why do you use these monitoring methods?	
4.	Which monitoring method is most effective, and why?	Please score from 1-10, where 1=not effective and 10= very effective (again using the matrix overleaf). Please explain your reasons.
5.	What do you think are the most important criteria to consider when designing a monitoring system?	
	Criteria	
6.	Where do you go for advice about monitoring methodology?	8
7.	What is your understanding of the	

	terms monitoring and surveillance?	
8.	Have you heard of Common Standards Monitoring? (If not mentioned previously)	
	Do you use it?	YES/NO
	If not, why not?	YES/NO
9.	If yes, is it effective for your work and why?	
Wou If sc	uld you mind if I contacted you again to as b, please leave a contact e-mail or phone r	k a few more questions? number.

Matrix for use with questions 2 and 4

	national	regional	local
Таха	score	score	score
habitat	score	score	score
landscape	score	score	score

Please return all completed forms to me, Sue Hearn. Thank you for your time.

School of the Environment and Natural Resources, Bangor University Gwynedd LL57 2UW s.m.hearn@bangor.ac.uk

Hellawell, J. M. (1991). Development of a rationale for monitoring. In: B. Goldsmith (Ed.), *Monitoring for Conservation and Ecology* (pp. 1-14). London: Chapman & Hall.



Appendices

Appendix 3.1 Aerial photographs and Ordnance Survey maps as provided to all surveyors involved in the field trial of the National Vegetation Classification approach to vegetation mapping in Hafod y Llan in summer 2008. Aerial photographs and maps not to scale.

55

Hafod y Llan: Survey Area Part A



1:5000

© This orthophotography has been produced by COWI A/S from digital photography captured by them in 2006. Licensed by the Welsh Assembly Government's Department for Environment, Planning and Countryside. © Mae'r orthoffotograffi hon wedi ei chynhyrchu gan COWI A/S o ffotograffiaeth ddigidol a gymerwyd ganddynt yn 2006. Wedi'u trwyddedu gan Adran Amgylchedd, Cynllunio a Chefn Gwlad Llywodraeth Cynulliad Cymru.

223

Hafod y Llan: Survey Area Part B





© This orthophotography has been produced by COWI A/S from digital photography captured by them in 2006. Licensed by the Welsh Assembly Government's Department for Environment, Planning and Countryside. © Mae'r orthoffotograffi hon wedi ei chynhyrchu gan COWI A/S o ffotograffiaeth ddigidol a gymerwyd ganddynt yn 2006. Wedi'u trwyddedu gan Adran Amgylchedd, Cynllunio a Chefn Gwlad Llywodraeth Cynulliad Cymru.





Crown copyright/database right 2008. An Ordnance Survey/EDINA supplied service.

Appendix 3.2 Large scale maps A-G of National Vegetation Classification community. Data collected in a field trial carried out in Snowdonia, 2008.







228





×





Appendix 4.1 Targets, field methods and criteria for JNCC CSM guidance from CSM Guidance for sand dune habitats and the Grid System taken from the CCW Aberffraw to Abermenai Dunes SAC UK0020021 SAC monitoring report. Text in bold indicates aspects used in field trials at Aberffraw to Abermenai Dunes SAC in 2008.

	JNCC CSM GUIDANCE	CCW Aberffraw to Abermenai Dunes SAC monitoring report
TARGET:No net decrease in extent from the establishedEXTENTbaseline.		As JNCC
TARGET: range of zones	Zonation from beach to fixed dune intact over at least 95% of coastal frontage.	As JNCC
TARGET: bare ground	Bare ground present but not >10% of whole area.	Not recorded as excessive amounts of bare ground or sand is not a problem at Aberffraw-Abermenai Dunes SAC.
TARGET: sward height	30-70% of sward to comprise species-rich short turf, 2-10cm tall.	For at least 70% of sample plots, sward height in April (before grazing or 'burnt off' in summer) between 2-10cm.
TARGET: flowering/ fruiting	Flowering and fruiting of the dune grassland at least frequent* level.	Not recorded as the level and timing of grazing at Aberffraw-Abermenai Dune SAC is deemed to be sufficient to allow adequate seed production.
TARGET: typical species	At least 8 typical species present at more than occasional* level.	At least 8 positive indicator species present in a 50cm radius of each sampling point.
TARGET: negative indicator species	 Non-native species no more than rare*. Any other negative indicators no more than frequent* throughout the sward, or singly or together the cover of negative indicator species no more than 5%. 	Negative indicator species and non-native species are absent within a 2m radius of each sampling point.
TARGET: scrub/trees	Scrub/ trees no more than occasional* or <5% cover.	Scrub/ trees absent within a 2m radius of each sampling point and overall trees/ scrub <5% cover.
TARGET: other negative indicators	Human/ vehicle damage should be absent or rare.	As JNCC
TARGET: indicators of local distinctiveness	Maintain distinctive elements at current extent/levels and/or in current locations.	Not recorded
FIELD METHOD	Structured walk (e.g a W shaped walk) with at least 10 4m ² stops within each assessment unit. The number of stops should be enough to allow the assessor to have an overview and judge the condition of the feature (area)	Grid placed over area with sample points placed at intersections and radii of 50 m and 2 m assessed at each point.
CRITERIA FOR FAVOURABLE CONDITION	The feature (area) must meet all of the targets to be deemed in favourable condition.	A sample point must meet all of the targets to be in favourable condition and each feature (area) has a target proportion of points to be favourable for the whole feature (area) to be favourable condition.

*Based on a version of the DAFOR scale which has been adapted to the particular characteristics of sand dune:

DOMINANT: species appears at most (>60%) stops and it covers no more than 50% of each sampling unit

ABUNDANT: species occurs regularly throughout a stand at most (>60%) stops and its cover is less than 50% of each sampling unit

FREQUENT: species recorded from 41-60% of stops

OCCASIONAL: species recorded from 21-40% of stops

RARE: species recorded from 1-20% of stops

Appendix 5.1 Protocol used in grid sampling at Hafod y Llan, Snowdonia National Park 2006-7

1. Sample point establishment

- Create a grid of 400 m across the whole of Hafod y Llan using a GIS mapping facility; assign a random starting point in the south west corner of the site. Choose the nearest exact 6 figure grid reference to this starting point and align the grid with eastings and northings of the national grid. The grid is chosen as an objective sampling strategy and for ease of navigation in the field. Use each intersection as a sample point (66 locations) and enter 6 figure grid references into a GPS.
- Locate two areas of 100 ha within the 400 m grid which are matched in their dominant vegetation types (acid grassland) and similar in altitude (approximately 300 m), but contrasting in spatial variation, with one relatively homogenous and the other relatively heterogeneous. Make this choice through visual inspection of the site and data from the 400 m plots and comparing the variance and means of ordination sample point scores for axis 1 and 2 from the 400 m plots from the two areas using an independent t-test in SPSS (SPSS, 2003). Create a 200m grid in each 100 ha area, with each intersection being a sample point and enter 6 figure grid references into GPS. Create grids of 100 m, 50 m and 25 m spacings, again with each intersection being a sample point in the central 25 ha, 6.5 ha and 1.56 ha of the 200 m grid successively. This will result in 50 points at 200 m, 100 m, 50 m and 25 m spacings.
- Print out copies of OS map with sampling points marked on them and use these in the field to get as near as possible to each location, abandoning points which are inaccessible due to dangerously steep rock exposures and old quarry workings. If entire areas are inaccessible, change the location of the relevant grid; in this case the 100 m grid in the homogenous area had to be off set. Finally use the GPS unit to locate each point and note the 'level of accuracy'.
- Establish plots of 5 m x 5 m over the 400 m grid and 2 m x 2 m over the other grid spacings, taking the sample point as the south west corner and measuring 5 m (or 2 m) northwards and then 5 m (or 2 m) eastwards from the marked point, marking each corner of the plot with a bamboo cane. If impossible to go 5 m (2 m) northwards due to cliffs etc then go 5 m (2 m) southwards and make a note.
- Divide the plot into 25 x 1 m² (or 16 x 0.5 m²) cells using tent pegs and brightly coloured thin rope (I used 4 mm climbing rope).

2. Whole plot measurements

- Sample point code and grid reference- assign each sample point a unique code and when the location has been found (as near as possible to the 6-figure grid reference), enter it into the GPS as a waymark, noting the grid reference and the accuracy.
- Altitude- use a 10 m resolution Digital Elevation Model to obtain altitude measurements for the origin of each sample point (I used the Landform Profile DTM 1: 1000 supplied by EDINA).
- **Slope** use canes placed across the steepest slope in the quadrat to sight along to measure the angle of slope using a clinometer.
- Aspect- measure the orientation of the broadest part of the quadrat using a compass.
- 3. Establishment of sub-plots (numbered moving north and then back south):
- In each 5 m x 5 m plot, establish 25 sub-plots by dividing the plot into 25 x 1 m² (or 16 x 0.5 m² in the 2 m x 2 m plots), using tent pegs and brightly coloured thin rope (I used 4 mm climbing rope). Number the sub-plots as follows:

N ↑	5	6	15	16	25
Sample point	4	7	14	17	24
	3	8	13	18	23
	2	9	12	19	22
	1	10	11	20	21

4. Measurements in each sub-plot Note presence/ absence:

- Litter- areas of litter where there is no vegetation present
- Dung- indicate species i.e. sheep (S), goats (G) cattle (C), rabbit (R), vole (V)
- Grazing/browsing- note type of browsed vegetation, grasses (G), heather (H), sedges (S), rushes (R), vaccinium (V) and seedlings (SE)
- Other evidence of grazing- note any wool (S- sheep, G- goat) molehills (MH), anthills (AH) tunnels (T) etc

Height of vegetation:

- Trees- note the species present and use the following system:
 - 2 m = wrist with arms stretched
 - 1 m 50 cm = chin with head straight ahead
 - 1 m = belly button
 - 50 cm = top of knee
 - 20 cm = top of boot socks
- Shrubs, grasses, sedges, ferns and bryophytes- use ruler to measure height of each taxon once in the middle of each sub-plot, or of the individual/patch nearest to the centre. Include any species inside the boundary of the sub-plot.

Equipment needed:

- GPS, OS map and site maps with plots marked, 50 m tape, compass, tent pegs and coloured rope cut into appropriate lengths.
- Weather writer, recording sheets, pencil and waterproof paper, ID books, clinometer and slope conversion charts and 1 m plastic ruler.

Appendix 5.2 Targets for sample points using Common Standards Monitoring guidance for wet heath habitats, amended by Countryside Council for Wales for use across the Eryri Special Area of Conservation (The National Trust, 2004; JNCC, 2005a).

At least 70% of the vegetation will be 'good condition wet heath', i.e. meet the following targets:

- 1. Erica tetralix present
- 2. >25% of the vegetation should be made up of species from Group i
- 3. >25% of the vegetation should be made up of species from Group ii
- 4. <1% of the vegetation should be made up of *Agrostis* spp., *Holcus lanatus and Ranunculus repens*
- 5. <10% of the vegetation should consist of Juncus effusus
- 6. <75% of the vegetation should consist of either dwarf shrubs and/or graminoids
- <33% mature shrub shoots, and <66% pioneer shoots should show signs of browsing
- 8. <10% disturbed Sphagnum

<u>Group i</u>

Small-medium sedges Drosera spp. (sundew) Non-crustose lichens Rhyncospora alba (white beak sedge) Sphagnum spp. Trichophorum cespitosum (deer grass)

Group ii

Calluna vulgaris (ling heath) Empetrum nigrum (crowberry) Erica tetralix (cross-leaved heath) Erica cinerea (bell heath) Vaccinium spp. (bilberry) Myrica gale (bog myrtle)