SYSTEMATIC APPROACHES TO SETTING CONSERVATION PRIORITIES USING SPECIES' DISTRIBUTION DATA.

Melanie Kershaw

Thesis submitted for the degree of Doctor of Philosophy,
University of London.
September 1996.
ABSTRACT.

The need to select priority areas for conservation action stems from the fact that throughout the world, human activity is increasingly resulting in the loss or irreversible alteration of natural ecosystems and their component species. Since there are not the resources to protect all species and communities and because diversity is not distributed equally, it has become necessary to select areas for conservation based on an assessment of the worthiness of a site compared with other sites or the countryside as a whole.

The setting of such priorities for conservation is a difficult task and in the past has been highly subjective with no clear objective or targets. Recently there have been moves towards a more systematic approach to the selection of conservation areas, with the aim of representing the widest possible range of the biotic diversity as efficiently as possible. However, while this aim may be universal, its practical implementation and its implications are unclear.

At a fundamental level there are three problems underlying attempts to make systematic selections of priority areas for conservation: (i) What is biodiversity and what are we trying to conserve? (ii) What surrogate measure can be used for overall diversity? and (iii) how should priority sites be selected using this information?

These three problems are evaluated and possible solutions investigated for two extensive data sets of distribution data. The first is for Afrotropical antelopes and includes information on species richness, taxonomic diversity, richness of threatened species and degree of rarity, for 249 areas in Africa. The second data set consists of distribution data for birds, carnivores and ungulates and plants in KwaZulu Natal, South Africa.

Using these data I discuss the different possible definitions of what constitutes biodiversity and investigate the consequences of using different components of biodiversity for selecting priority areas for conservation. These components include simple species richness, taxonomic diversity, threatened species, endemic species and 'rare' species.
The second problem associated with setting priorities for conservation is that since we do not know how many species exist or how they are distributed it is necessary to have some sort of surrogate measure that can be used to evaluate the importance of areas for a wide range of biotic diversity. Different potential surrogate measures, including indicator groups, higher taxon richness, endemic species and habitat representativeness are evaluated in terms of their ability to represent a range of diversity across a variety of taxa.

Finally, once the diversity value of different sites has been determined it is necessary to select a network of priority areas to represent the diversity. This requires some sort of site selection method. Simple scoring approaches where the top ranking sites for particular diversity criteria are compared to more sophisticated techniques using iterative algorithms that take account of complementarity.
ACKNOWLEDGEMENTS.

Many people have provided data and computer software for use in this work. I am very grateful to Paul Williams for making his WORLDMAP software available and for many useful discussions concerning area selection algorithms. I should also like to thank my supervisor, Georgina Mace, as well as Dick Vane-Wright, Chris Humphries, Josh Ginsberg, Åke Berg, John Prendergast, Rachel Quinn and Blair Csuti for helpful ideas over the past few years. John Prendergast, Paul Williams, Andrew Balmford, Tim Coulson and Steve Albon kindly commented on earlier drafts of this thesis. Rod East provided helpful comments concerning the Afrotropical antelopes analysis.

On the computing front I should like to thank Adrian Fisher for helping with the area selection algorithms and for teaching me to program. Steve Wilkinson provided advice on the use of CANOCO and how to understand ordination biplots.

The director (research) of the National Botanical Institute, Pretoria is thanked for the use of data produced by the Pretoria National Herbarium Computerised Information System (PRECIS). These data consisted of all quarter degree distribution records for plant taxa collected in KwaZulu Natal, South Africa. Temperature data for KwaZulu Natal was kindly supplied by Prof. R.E. Schulze, University of Natal. Ingrid Figenschou, Computing Centre for Water Research, University of Natal provided the precipitation data for KwaZulu Natal. The 15 x 15 minute precipitation data were extracted from a data set generated by Steve Lynch, Dept. of Agricultural Engineering, University of Natal. The steering committee responsible for this project consisted of the following persons: Mr D.W.H. Cousens (Water Research Commission chairman), Mr P.W. Weidman (WRC secretary), Dr P.J.T. Roberts (WRC), Prof. P. Meiring (University of Natal), Mr F.J.Kruger (Dept of Environment Affairs), Mr J.F. Erasmus (Dept of Agriculture & Water Supply), Dr P.T. Adamson (Dept. of Water Affairs), Mr F. Viljoen (Dept. of Environment Affairs). The financing of the project by the WRC and the contributions of members of the steering committee is acknowledged gratefully. This project of five years duration was only possible with the cooperation of many individuals and institutions. We therefore wish to record our sincere thanks to the following; Dept

I should particularly like to thank Adrian for all his support over the past three years (and persuading me not to give up on the thesis), my parents for their encouragement (and for daring to proof read some of the chapters), Marcus for being very tolerant of all the mess in the flat and everyone at WWT for distracting me for the last year so I didn’t get a chance to finish the thesis.
Declaration.

I declare that the work contained in this thesis is my own. Some of the work contained in Chapter 2 has been published jointly with Paul Williams and Georgina Mace (Kershaw, M., Williams, P.H & Mace, G.M. 1994. Conservation of Afrotropical antelopes: consequences and efficiency of using different site selection methods and diversity criteria. *Biodiversity and Conservation* 3, 354-372.). This paper was written by me and the analysis in the paper was also carried out by myself. Results presented in Chapter 5 are also published (Kershaw, M., Mace, G.M. & Williams, P.H. 1995. Threatened status, rarity and diversity as alternative selection measures for protected areas: a test using Afrotropical antelopes. *Conservation Biology* 9, 324-334.). Again the analysis was done by myself and I wrote the paper. Some of the data from KwaZulu Natal have been published with Georgina Mace (Mace, G.M. & Kershaw, M. In press. Extinction Risk and Rarity on an Ecological Timescale. In *The Biology of Rarity*, (eds. W.E. Kunin & K.J. Gaston), pp 130-149, 1997.). The results of the analyses contained in this publication are not included in this thesis. Some of the work in Chapter 6 is also published jointly (Csuti, B., Polasky, S., Williams, P.H., Pressey, R.L, Camm, J.D., Kershaw, M., Kiester, A.R. Downs, B., Hamilton, R., Huso, M & Sahr, K. In press. *Biological Conservation*. A comparison of reserve selection algorithms using data on terrestrial vertebrates in Oregon.). Chapter 6 only includes the analysis which I did for this publication.
Contents.

Title page.................................................................1
Abstract.................................................................2
Acknowledgements................................................4
Declaration............................................................6
Contents...............................................................7
List of tables........................................................9
List of figures........................................................12
Chapter 1.
General Introduction..............................................16
Chapter 2.
A comparison of the efficiency of alternative area selection criteria at representing different biodiversity attributes........................................44
  Introduction.........................................................44
  Methods............................................................46
  Results.............................................................58
  Discussion.........................................................75
  Conclusions.......................................................87
Chapter 3.
Quantifying patterns of richness and spatial turnover: what makes a good surrogate measure for wholesale biodiversity?..................88
  Introduction.........................................................88
  Methods............................................................94
  Results.............................................................112
  Discussion.........................................................143
  Conclusions.......................................................156
Chapter 4.
A comparison of the consequences of using two alternative site selection methods to represent the diversity of a region: scoring versus iterative selection methods........................................157
Chapter 5.

A comparison of the efficiency of different iterative reserve selection algorithms.

Introduction. ..............................................................185
Methods. .................................................................187
Results. .................................................................192
Discussion. .........................................................203
Conclusions. ......................................................210

Chapter 6.

General Discussion. .................................................211

Literature cited. ................................................218
List of Tables.

Table 2.1. Numbers of species, threatened species, endemic species and restricted-range species in each of the different taxonomic groups in the data set for KwaZulu Natal.

Table 2.2. Percentage of different diversity attributes represented in the top six and twelve sites selected using alternative diversity criteria for Afrotropical antelopes.

Table 2.3. Top ranking sites according to the alternative diversity criteria for Afrotropical antelopes.

Table 3.1. List of the environmental variables included in the CCA analysis for birds, mammals and grasses in KwaZulu Natal.

Table 3.2. Correlation coefficients (r) between species richness of quarter degree grid cells (n=166) for different groups in KwaZulu Natal.

Table 3.3. Percentage overlap of top sites for bird, mammal and plant species richness with sites selected using alternative diversity criteria.

Table 3.4. Eigenvalues & cumulative percentage variance explained by the first four axes extracted by correspondence analysis compared to canonical correspondence analysis of sites in KwaZulu Natal, southern Africa based on (a) passerine birds, (b) nonpasserine birds, (c) mammals (carnivores & ungulates), (d) grasses (Poa1) and (e) grasses (Poa2).

Table 3.5. Eigenvalues for the first four axes extracted by canonical correspondence analysis for each group (passerines, nonpasserines, mammals and grasses (Poa1 & Poa2)), with their significance levels calculated from 99 Monte Carlo permutations.

Table 3.6a. Interset correlations of the first CCA axis with environmental variables, arranged in rank order for each taxon.

Table 3.6b. Interset correlations of the second CCA axis with environmental variables, arranged in rank order for each taxon.

Table 3.7. Species turnover values for birds, mammals & plants along fourteen east-west transects in KwaZulu Natal, measured using two alternative methods (Harrison e. al. 1992 & Whittaker 1960 - see text for explanation).

Table 3.8. Species turnover values for birds, mammals and plants measured along twelve north-south transects in KwaZulu Natal using the beta diversity measure of Harrison et al. (1992).

Table 3.9. Mean turnover (with S.E.) for fourteen east-west transects in KwaZulu Natal
calculated for habitat specific and habitat generalist passerine species separately using the Harrison et al. (1992) measure of species turnover.

Table 3.10. Mean pairwise similarity (CC_j) (x100) for birds and plants, between 168 pairs of sites from the same dominant habitat type, and 168 pairs of sites from different dominant habitat types.

Table 3.11. Percentage of species represented in the top eight sites (hotspots) selected using alternative diversity criteria.

Table 4.1. Number of sites needed to represent all species using two alternative sites selection methods (selection of top sites (scoring) and an iterative richness algorithm) & efficiency of iterative site selection methods compared to the selection of top scoring sites for birds, mammals and plants in KwaZulu Natal and Afrotropical antelopes.

Table 4.2. Numbers (%) of species represented in the first eight sites (for KwaZulu Natal taxa) or first twelve sites (Afrotropical antelopes) selected by the two alternative selection methods.

Table 4.3. First twelve sites selected to represent Afrotropical antelopes using two alternative selection methods (scoring versus an iterative richness algorithm). The cumulative number of species that would be represented as sites are added to the network is also given.

Table 4.4. First eight sites selected to represent bird species in KwaZulu Natal using two alternative selection methods (scoring versus an iterative richness algorithm). The cumulative number of species that would be represented as sites are added to the network is also given. Sites are coded as letters corresponding to the west-east and north-south position of the quarter degree grid cell.

Table 4.5. First eight sites selected to represent mammal species in KwaZulu Natal using two alternative selection methods (top sites versus an iterative richness algorithm). The cumulative number of species that would be represented as sites are added to the network is also given.

Table 4.6. First eight sites selected to represent plant species in KwaZulu Natal using two alternative selection methods (top sites versus an iterative richness algorithm). The cumulative number of species that would be represented as sites are added to the network is also given.

Table 4.7. Effect of having some sites already preassigned to the site network on the efficiency of an iterative selection of sites for Afrotropical antelopes.

Table 4.8. t-tests for the differences between the mean number of sites needed to represent all Afrotropical antelope species when high richness compared to low richness sites are preassigned to the site network for different numbers of
preassigned sites (2, 5, 10 & 15). P values give the two-tailed significance level.

Table 4.9. t-tests for the difference between the mean number of sites needed to represent all Afrotropical antelope species with no sites preassigned to the network compared to when different numbers of sites, of varying richness are preassigned. P values give the two-tailed significance level.

Table 4.10. Comparison of the ability (in terms of number and percentage of species) of different site selection methods at representing threatened and endemic species for birds, mammals and plants in KwaZulu Natal.

Table 4.11. Comparison of the ability (in terms of % of species/diversity represented) of different site selection methods at representing threatened species, restricted-range species and taxonomic diversity for Afrotropical antelopes.

Table 5.1. Number of sites needed to represent all species using four different iterative algorithms. Numbers in brackets denote the number of sites that were 'redundant' following the backtracking procedure which was applied to the richness algorithms for the KwaZulu Natal taxa.

Table 5.2. Efficiency of the different iterative algorithms.

Table 5.3. Sites selected by the different iterative algorithms for Afrotropical antelopes.

Table 5.4. Number of species that are represented by each of the alternative iterative algorithms when the number of sites that can be selected is limited.

Table 5.5. Number of sites needed to include all species in the Oregon data set using four different iterative algorithms (richness, rarity I, rarity II & rarity III) compared to the optimal solution as derived from integer programming.

Table 5.6. Number of species represented by the alternative algorithms when the number of sites that can be selected is limited (6 sites & 12 sites) for the Oregon data set.
List of Figures.

Figure 2.1. Classification of Afrotropical antelopes (based on Gentry 1992) used to derive taxonomic diversity measures in WORLDMAP.

Figure 2.2. Example illustrating the formulation of codes used to enter a taxonomic hierarchy into WORLDMAP. The program weights different faunas according to the number of taxa and number of internodes subtended by the included species.

Figure 2.3. Map of southern Africa showing the location of KwaZulu Natal.

Figure 2.4. Ability of alternative site selection criteria (x axis) at representing different biodiversity attributes of Afrotropical antelopes: (a) species richness; (b) restricted-range richness; (c) threatened species richness; (d) dispersion-weighted taxonomic diversity, (e) higher-taxon weighted diversity & (f) spanning-subtree weighted taxonomic diversity.

Figure 2.5. Ability of alternative site selection criteria (x axis) at representing different biodiversity attributes of birds in KwaZulu Natal: (a) species richness; (b) threatened species richness; (c) endemic species richness & (d) restricted-range richness.

Figure 2.6. Ability of alternative site selection criteria (x axis) at representing different biodiversity attributes of mammals in KwaZulu Natal: (a) species richness; (b) threatened species richness & (c) restricted-range species richness.

Figure 2.7. Ability of alternative site selection criteria (x axis) at representing different biodiversity attributes of plants in KwaZulu Natal: (a) species richness; (b) threatened species richness & (c) restricted-range species richness.

Figure 2.8. Ability of alternative site selection criteria (x axis) at representing different biodiversity attributes of grasses in KwaZulu Natal: (a) species richness; (b) endemic species richness & (c) restricted-range species richness.

Figure 2.9. Location of the top eight sites selected by the alternative diversity criteria for birds in KwaZulu Natal.

Figure 2.10. Location of the top eight sites selected by the alternative diversity criteria for mammals in KwaZulu Natal.

Figure 2.11. Location of the top eight sites selected by the alternative diversity criteria for plants (all species) in KwaZulu Natal.

Figure 2.12. Location of the top eight sites selected by the alternative diversity criteria for grasses in KwaZulu Natal.

Figure 2.13a. Cladograms showing where species from three of the antelope sites lie on the overall classification (Keran, Waza & Mont Sangbe). Each site has three
species, but different taxonomic diversity scores as calculated by the three WORLDMAP measures.

Figure 2.13b. Cladograms showing where species from three of the antelope sites lie on the overall classification (Meshra, Harrar & Scio). Each site has three species, but different taxonomic diversity scores as calculated by the three WORLDMAP measures.

Figure 3.1. Five hypothetical groups of sites (a-e) showing different levels of species richness and spatial turnover (complementarity) for two taxa.

Figure 3.2. Map of KwaZulu Natal showing the different habitat types, redrawn from Rowe-Rowe 1992a.

Figure 3.3. Concordance of top sites for bird, mammal & plant species richness in quarter degree squares in KwaZulu Natal, southern Africa. The top eight squares for each taxon are shown.

Figure 3.4. Relationship between bird species richness and bird family richness of quarter degree squares in KwaZulu Natal, southern Africa.

Figure 3.5. Ordination diagram based on canonical correspondence analysis of passerine distribution data in 166 quarter degree squares in KwaZulu Natal, with respect to the environmental variables.

Figure 3.6. Ordination diagram based on canonical correspondence analysis of mammal distribution data in 166 quarter degree squares in KwaZulu Natal, with respect to the environmental variables.

Figure 3.7. Ordination diagram based on canonical correspondence analysis of grass (POA1) distribution data in 166 quarter degree squares in KwaZulu Natal, with respect to the environmental variables.

Figure 3.8. Ordination diagram based on canonical correspondence analysis of passerine distribution data in KwaZulu Natal with respect to the environmental variables. Each of the 195 breeding passerine landbird species is shown as the centroid (▲) representing the species’ optima in 2-dimensional space in the ordination.

Figure 3.9. Ordination diagram based on canonical correspondence analysis of passerine distribution data in KwaZulu Natal with respect to the environmental variables. Each of the endemic, breeding, passerine landbird species is shown as the centroid (▲) representing the species’ optima in 2-dimensional space in the ordination.

Figure 3.10. Ordination diagram based on canonical correspondence analysis of nonpasserine distribution data in KwaZulu Natal with respect to the environmental variables. Each of the endemic, breeding, nonpasserine landbird
species is shown as the centroid (▲) representing the species’ optima in 2-dimensional space in the ordination.

Figure 3.11. Ordination diagram based on canonical correspondence analysis of passerine distribution data in KwaZulu Natal with respect to the environmental variables. Each of the threatened, breeding, passerine landbird species is shown as the centroid representing the species’ optima in 2-dimensional space in the ordination.

Figure 3.12. Ordination diagram based on canonical correspondence analysis of nonpasserine distribution data in KwaZulu Natal with respect to the environmental variables. Each of the threatened, breeding, nonpasserine landbird species is shown as the centroid representing the species’ optima in 2-dimensional space in the ordination.

Figure 3.13. Ordination diagram based on canonical correspondence analysis of passerine distribution data in KwaZulu Natal with respect to the environmental variables. Each of the habitat specific, breeding, passerine landbird species is shown as the centroid (▲) representing the species’ optima in 2-dimensional space in the ordination.

Figure 3.14. Ordination diagram based on canonical correspondence analysis of grass distribution data (POA2) in KwaZulu Natal with respect to the environmental variables (omitted for clarity - see figure 3.7). Each of the endemic species is shown as the centroid (▲) representing the species’ optima in 2-dimensional space in the ordination.

Figure 3.15. Ordination diagram based on canonical correspondence analysis of mammal distribution data in KwaZulu Natal with respect to the environmental variables. Each of the threatened mammal species is shown as the centroid (▲) representing the species’ optima in 2-dimensional space in the ordination.

Figure 3.16. Species accumulation curves as sites are added to a site network starting with the most species rich sites and selecting subsequent sites in decreasing rank order (up to a total of 30 sites). The vertical line indicated the number of species that would be represented in the top eight sites.

Figure 3.17. Percentage of threatened species represented in the top eight sites for species richness in KwaZulu Natal for each of the taxa, plants, birds & mammals.

Figure 4.1. Locations of the first eight sites selected for birds in KwaZulu Natal by the two alternative selection methods (top-ranking sites for species richness & an iterative richness selection).

Figure 4.2. Locations of the first eight sites selected for mammals in KwaZulu Natal by the two alternative selection methods (top-ranking sites for species richness & an iterative richness selection).
Figure 4.3. Locations of the first eight sites selected for plants in KwaZulu Natal by the two alternative selection methods (top-ranking sites for species richness & an iterative richness selection).

Figure 5.1. Locations of the first eight sites selected by the different iterative algorithms (Richness, Rarity I, Rarity II & Rarity III) for birds in KwaZulu Natal.

Figure 5.2. Locations of the first eight sites selected by the different iterative algorithms (Richness, Rarity I, Rarity II & Rarity III) for mammals in KwaZulu Natal.

Figure 5.3. Locations of the first eight sites selected by the different iterative algorithms (Richness, Rarity I, Rarity II & Rarity III) for plants in KwaZulu Natal.

Figure 5.4. Species accumulation curves for the different iterative algorithms (Richness, Rarity I, Rarity II & Rarity III) for the Oregon data set.
CHAPTER ONE.

GENERAL INTRODUCTION.

The problem.
The need to select priority areas for conservation action stems from the fact that throughout the world, human activity is increasingly resulting in the loss or irreversible alteration of natural ecosystems and their component species (Ehrlich 1988, Myers 1988, Wilson 1988, Vane-Wright et al. 1991, Groombridge 1992, Grehan 1993). Since there are not the resources to protect all species and communities and because diversity is not distributed equally, it has become necessary to prioritise areas for conservation based on an assessment of the worthiness of a site compared to other sites or the countryside as a whole.

In the past area selection has been highly subjective, often on the basis of individual species or groups of species with a high public profile, or for scenic or recreational reasons, with no clear objective or targets for species, habitat or ecosystem conservation (Terborgh & Winter 1983, Purdie 1987, Vane-Wright 1993). Recently there have been moves towards a more systematic approach to the selection of conservation areas, with the aim of representing the widest possible range of biotic diversity as efficiently as possible (Austin & Margules 1986, Margules et al. 1988, Pressey & Nicholls 1989b, May 1990, Pressey et al. 1993, Vane-Wright et al. 1994, Williams et al. 1996b). However, while this aim may be universal, its practical implementation and its implications in terms of the areas that are important are unclear. Even in cases where more systematic approaches have been used, the areas selected will depend on the attributes used to evaluate sites, the selection method, the motive for protection, and at what level of organisation efforts are concentrated, and aims are often distorted by economics and the availability of land (Green 1985, Pressey 1994).

A major problem in the selection of priority areas is determining what it is that we are trying to conserve; or more generally what is biodiversity? This will be dependent on the values and goals of individuals and these are not universal (Gaston 1994, Williams et al. 1996c).
Priority areas could be based on a variety of different biodiversity attributes. For example, species richness, endemic species, taxonomic richness, genetic diversity, threatened species, ecological type diversity and community diversity are all attributes of biodiversity which could be used to select areas for conservation, depending on the overall conservation goal. Additionally, since we do not fully understand how certain attributes of biodiversity contribute to potential goals like the maintenance of evolutionary potential or ecosystem processes and function, we cannot easily determine how to achieve these goals. A further problem is that there is not the time nor the resources to evaluate all aspects of biodiversity. Society has to decide what the goals for conservation are, but this needs to be based on information about the distribution of biodiversity attributes. The aim of the work presented in this thesis is to provide the sort of information that can guide strategic decisions concerning the best methods to achieve alternative biodiversity goals. If we choose to base area selection on a single goal (for example, maximising the number of species that are represented) then we need to know which attributes of biodiversity will and will not be incorporated. This will be particularly important in cases where a different approach or the use of different attributes would result in very different areas being designated as high priority (Kershaw et al. 1995). Alternatively, it may be possible to use a technique that incorporates many different attributes and goals explicitly.

Aims of priority areas for conservation.
Conserving a representative selection of the full range of biotic diversity (from genes and species to ecosystems) is a widely cited goal (Margules & Usher 1981, Bolton & Specht 1983, Van der Ploeg 1986, Pressey & Nicholls 1989b, Bedward et al. 1992, Nilsson & Götmark 1992, Saetersdal et al. 1993). Representativeness was defined by Austin & Margules (1986) as a criterion for assessing how adequately a reserve or system of reserves represents the range of biological variation in a region. Representativeness was seen as a priority for conservation in Britain when the government published its white papers "Conservation of nature in England and Wales" and "National Parks and the conservation of nature in Scotland" in 1947 following the recommendation by the Huxley committee that nature reserves should represent reference points within the countrywide
field of variation in ecosystem types (Ratcliffe 1977, Usher 1986). This lead to the key area concept - nature reserves in Britain should centre around the safeguarding of a fairly large number of key areas which were representative of all major types of natural and semi-natural vegetation, their fauna and habitat conditions. The key area approach was subsequently considered largely inappropriate and in 1966, in the Nature Conservancy Review, which was launched to reappraise candidate sites for expansion of the National Nature Reserve network, factors like ecosystem fragility, vulnerability to future loss due to fragility and predicted future human pressures and irreplaceability of the ecosystem were seen as priorities.

At an international level Dasmann (1972) recognised the need to establish a worldwide network of reserves on a biogeographical basis. With this in mind the World Conservation Strategy (IUCN 1980) was developed with three fundamental aims: maintenance of essential ecological processes, preservation of genetic diversity and sustainable utilisation of species and ecosystems (IUCN 1980). In 1992 over 150 countries signed The Convention on Biological Diversity at the United Nations Conference on Environment and Development. Part of this convention requires contracting parties to develop national strategies for the conservation of biodiversity. In response to this the UK Action Plan was published, the overall goal being to ‘conserve and enhance biological diversity within the UK and to contribute to the conservation of global biodiversity through all appropriate mechanisms’ (DoE 1994). The Action Plan recognises the multifaceted nature of biodiversity and objectives such as representing the range of species, wildlife habitats and ecosystems and conserving threatened species, ecosystems and elements characteristic of local areas are cited. As well as representing UK biodiversity the need to conserve elements of international and global importance is also an objective (DoE 1994).

While a widely agreed goal of priority areas might be to conserve a representative range of the existing biotic diversity, it is not obvious how this translates practically. At a fundamental level there are three major problems underlying attempts to make systematic selections of priority areas for in situ conservation and these will form the basis of the
results presented in this thesis:

1. What to conserve (and can different attributes of biodiversity be conserved in the same sites)?
2. What surrogate measures can be used to reflect overall biodiversity patterns?
3. How should biodiversity value information be used to select priority areas?

Definitions.
In order to fully review the different goals and methods for the selection of priority conservation areas it is necessary to briefly outline some of the terminology used:

_Biodiversity_ describes the variety and variability among living organisms and the ecological complexes in which they occur (McNeely et al. 1990). Although biodiversity can be considered at a number of different levels (genes, species and ecosystems), species are the most commonly used units for evaluation and areas are usually the most practical way of conserving biodiversity.

_Attributes_ are those properties of a site or area which can be used to reflect its conservation interest. They represent the currency of conservation value. For example a site attribute might be the species that are present, or the habitat types in the site.

_Criteria_ provide a means of expressing the attributes in such a way that they can be used to evaluate the conservation interest of a site. For instance the list of species in a site could be used to calculate a diversity or richness index criterion, or a rarity value criterion.

The total area from which priority sites could be drawn is termed the _region_. _Regional diversity_ is then all the elements of biodiversity that are found in the total area.

Within a region the individual sites that are selected as priority areas together form a _site network_.

19
species richness is used simply to describe the number of species present in a site or region.

Threatened species are those species that have been included on some sort of ‘red list’, for example, Red Data Book (IUCN 1994), indicating that they are considered susceptible to extinction in the near future. Different degrees of threat are usually recognised, and species may be included on red lists for a number of different reasons. New methods for assessing global threatened species lists have been adopted (IUCN 1994) although these new criteria have yet to be applied to the majority of taxa (although see Collar et al. 1994).

Rarity has a number of definitions and the term is used in many different ways (Magurran 1988, Gaston 1994). Rabinowitz (1981) and Rabinowitz et al. (1986) identified seven ways to be rare based on three rarity ‘traits’ - restricted-range size, habitat specificity and small population size.

Endemic species are those that are confined to a particular geographic region. If this region is small then the endemic species could be classified as restricted-range and therefore represent a form of rarity.

Complementarity is the degree to which individual areas or discrete sites contain elements of biodiversity that are represented by other sites. Two sites containing completely different biodiversity elements have high complementarity whilst sites containing the same or similar elements have low complementarity.

A surrogate measure is defined as an attribute which accurately reflects the patterns of distribution across a wide range of additional taxa or other biodiversity attributes. For example plant species richness patterns might mirror richness patterns in a range of other taxa so selecting areas rich in plants would also select areas of high richness for other taxa. Plants could therefore be used as a surrogate for richness in other taxa without having any information on the distribution of these other groups.
Higher taxon is used to refer to any taxonomic rank that is higher than the species level, for example family or order.

When areas have been evaluated for biodiversity value a selection method can be applied. The sites that are selected are priority areas. These will vary according to the attributes, criteria and selection method used.

When sites are selected for inclusion in a network by means of taking the sites ranked most highly on the basis of some criterion (for example, species richness score) this is described as a scoring or ranking selection method.

A selection method that proceeds in a stepwise fashion, at each step taking account of the attributes represented by previous site selections is termed an iterative selection method or iterative heuristic.

The ability of a set of reserves to represent the biotic diversity of an area can be termed the efficiency of the network (Pressey & Nicholls 1989a). This can be measured in terms of the proportion of a total region that is required to represent all the desired attributes. Efficiency must undoubtedly be a primary objective given the pressure on land and the cost involved.

What to conserve (and can different attributes of biodiversity be conserved in the same sites)?

It is generally agreed that the role of protected areas is to maintain a representative selection of the existing biotic diversity, but biodiversity does not have a simple definition. Part of the problem may be that the word biodiversity is treated as if it were a ‘thing’ whereas in reality it is an abstract idea. Biodiversity describes the range and variability of all living organisms (McNeely et al. 1990). If only a subset of this variability can be conserved or protected there is no clearly objective way of determining which group of organisms represents the ‘most’ biodiversity. Species richness, taxonomic diversity, genetic diversity, community/ecosystem diversity, evolutionary
potential and threatened or vulnerable species are all components of biodiversity which could be used to assess the conservation value of an area. Problems will arise when a set of areas selected to maximise one aspect of biodiversity, for example, species richness are not the best sites for other aspects of diversity, like rare or threatened species (Prendergast et al. 1993, Kershaw et al. 1995, Williams et al. 1996a).

Threatened species.
The ultimate goal of conservation is to prevent the extinction of species (Diamond 1976, ICBP 1992), thus species that are threatened could be an appropriate focus for the selection of conservation areas. Species vary in their vulnerability to extinction, so if the aim of a reserve network is to maintain a representative selection of species diversity then it is necessary to take account of the fact that some species are less likely to persist in the future than others. Optimisation of the amount of biodiversity represented in reserves could jeopardise the long term representation of biodiversity if limited resources are used to represent common species leaving vulnerable species unprotected (Witting & Loeschcke 1995).

Species which are not threatened or potentially threatened do not need protection (Ratcliffe 1986). Giving these species equal weighting with species that are threatened in the short term could affect the areas that are valued as conservation priorities (Saetersdal 1994). However, if a large number of species are ultimately threatened with extinction then concentrating on species currently threatened or vulnerable may lead to biases in the representation of biotic diversity in the future, especially if threatened species tend to occur in certain taxonomic, ecological or life history groups. For example, if threatened species are more likely to be classically rare species (restricted-range, small population size and habitat specific) then concentration on single threatened species may lead to habitats being preserved on the basis of rare species and for these species to therefore be over-represented in protected areas.

Also, it is not possible to identify every species which is threatened or potentially threatened with extinction, and then set aside areas to conserve these priority species.
Firstly not all species are known, secondly it would be far too time consuming and practically unrealistic to evaluate the status of all species, and thirdly, there are not the resources to set aside areas to conserve all the species that are or are likely to become threatened (Noss 1991). Additionally, we cannot accurately predict which species will be subject to threatening processes in the future.

In the past some conservation efforts have revolved around last ditch attempts to save threatened species - the so-called ‘Noah Principle’ (Heywood 1988) - for example the Black footed ferret (Groves & Clark 1986), Red-cockaded woodpecker (Reed et al. 1988), Grizzly bear (Schaffer 1983) and Spotted owl (Barrowclough & Coats 1985). Such an approach could result in valuable resources being channelled into ‘hopeless cases’ rather than attempting to ensure the maintenance of species and ecosystems that have a real chance of persistence.

Nonetheless, richness of threatened species should highlight those areas in imminent danger of losing species and therefore most in need of protection (Rebelo & Tansley 1993). Areas could be selected that maximise the number of threatened or potentially threatened species that would be represented. Rapoport et al. (1986) proposed a method for the designation of optimum reserve areas based on the conservation value of each component species, weighted towards threatened species, which in turn gave a conservation value per unit area. Species’ values depended on the local area occupied, geographical range, density or abundance and tendency to become extinct.

However, approaches based around maximising the number of threatened species in protected areas have also been criticised. An analysis of the plant species listed as threatened in New South Wales, Australia revealed that 55% were confined to one of two distinctive ecological types (McIntyre 1992). Herb species were significantly under represented in the threatened species list compared to their contribution to the overall flora suggesting that being less conspicuous than trees and shrubs has lead to their under documentation (McIntyre 1992). Eighty-four percent of the species listed as threatened were either restricted-range and habitat specific or had a wide range and were habitat
specific. Only 11% of the species were widely distributed habitat generalists. This may indicate that widespread, but uncommon species have been overlooked in assessments of conservation status (McIntyre 1992). These species may however be the most demanding to conserve as they will need large areas to be set aside. In contrast endemic species may be successfully conserved in relatively small areas (McIntyre 1992).

So while there is no doubt that threatened species are a priority for conservation, concentrating purely on threatened species could lead to biases in the future representation of biodiversity.

Rare species.
An alternative to focusing on species categorised as threatened is to prioritise areas on the basis of rare species, since there is a general assumption that the rare species of today are the extinct species of tomorrow (McIntyre 1992, Gaston 1994, Mace & Kershaw in press). But this is not necessarily the case, especially since species can be rare for a number of different reasons and in a number of different ways, which in turn are highly dependent on the scale used (Rabinowitz 1981, Magurran 1988, McIntyre 1992, Gaston 1994, Saetersdal 1994). Mayr (1963) recognised that species could be rare for different reasons - either because they are highly specialised or because they are very localised. Rabinowitz (1981; Rabinowitz et al. 1986) classified seven different ways in which rarity could manifest itself - a species can be rare if it is geographically restricted, if it only ever exists in small populations or if it is habitat specific. These then give rise to seven different combinations in which a species could be rare according to one or more of the categories, the extreme case being a species that is geographically restricted, habitat specific and only occurs in small populations. Such species are those which would be considered to be classic endemics, but a species can also be rare if it is widely distributed, a habitat generalist but only ever exists in small populations and these are the sort of species that tend to be overlooked. It is these species that are most likely to be unprotected if areas are selected for conservation based on classically rare, restricted-range species.
A number of studies have looked at the distributions of various forms of rarity in different taxonomic groups and at different scales and how this relates to the future persistence of populations (Thomas & Mallorie 1985, Pagel et al. 1991, Kattan 1992, Saetersdal 1994). Cowling & Bond (1991) found that the species most likely to be missing from small islands fragmented from the mainland were endemics and specialists. Similarly Newmark (1991) found that forest-dependent understorey species in Tanzania varied greatly in their vulnerability to fragmentation, but the most susceptible groups were rare and forest interior species (i.e. habitat specialists). Bolger et al. (1991) found a strong positive correlation between the relative persistence ability of a species and its density in habitat fragments and concluded that density is a primary factor in determining a species’ relative vulnerability to extinction. Small population size was found to be the most common reason (more important than range-size) for threatened species listing in a recent conservation review of birds using the new IUCN threatened species criteria (Collar et al. 1994, IUCN 1994, Mace & Kershaw in press). Given & Norton (1993) found that the most useful criteria for prioritising threatened plant conservation in New Zealand were population decline rate, size of largest population and habitat specificity. Geographic distribution was a redundant criterion in that it did not account for any of the differences between species in terms of their threat status (Given & Norton 1993).

However, population size or abundance may not always be a good measure of a species’ vulnerability. There are a number of examples where persistence has been related to distribution - that is species with small ranges are more susceptible to extinction (Terborgh & Winter 1983, Thomas 1991, Thomas & Mallorie 1985, ICBP 1992). A recent study by Birdlife International found that three quarters of threatened bird species had restricted-ranges, defined as a total range of less than 50,000km² (ICBP 1992). Thomas & Mallorie (1985) found that while vulnerability of Moroccan butterfly species was related to distribution, it was not necessarily related to abundance. Similarly, Thomas (1991) compared the geographic range of Costa Rican butterflies with their potential ability to use human-modified habitats and found that restricted-range species were less able to utilise such areas, and so were most vulnerable to habitat loss.
Burke & Humphrey (1987) looked at rarity as a criterion for endangerment in Florida's fauna. Generally, initial rarity (defined as limited distribution, low local density and large body size) was a good indicator of a species' subsequent vulnerability (as indicated by a species' inclusion on threatened species lists), however there were some significant discrepancies. Some rare species were more extinction prone than expected and likewise some species which were not rare were threatened for other biological or economic reasons, for example by excessive settlement within their range. For instance the bat species *Myotis grisecens* was very abundant where it occurred but needed caves to breed which were uncommon and very sensitive to disturbance. So rarity may be useful as an initial filter for assigning conservation priorities, but it will also be necessary to consider other factors that could affect species' persistence.

Species that have been rare for long periods may have adaptations that allow them to persist in a rare state (Kunin & Gaston 1993, Gaston 1994, Mace & Kershaw in press). For example, Simberloff & Gotelli (1984) looked at the effects of insularisation on plant species richness in the prairie-forest ecotone and found that rare species were found more often in small sites than a random colonisation model would predict i.e., they were less susceptible to local extinction than expected. Rare species may therefore be in less need of protection than more common species which may lack adaptations to cope with population reduction or fragmentation.

The scale at which rarity is defined is likely to influence those species listed and is an important consideration when setting priorities for conservation. Species that are rare at a local scale may be common at a regional level and *vice versa*. When Saetersdal (1994) classified vascular plants in Norway according to Rabinowitz's (1986) seven forms of rarity, a number of the species changed rarity category depending on whether the scale used was local (deciduous woodlands in Hordaland, western Norway) or regional (the whole of western Norway). Sixty-seven of the 225 deciduous woodland species that were common at the local scale fell into one of the rarity categories at the regional level (Saetersdal 1994). Seven species were rare at the local level but common at the regional level and 45 species changed from one form of rarity to another.
The same effect can be seen with threatened species listings. In North America, most states and provinces have lists of state/provincially endangered species and in Europe individual countries have their own threatened species lists (Hunter & Hutchinson 1994). A number of the species on these lists are there because they are peripheral species, on the edge of their range in the country in question (Hunter & Hutchinson 1994). For example, the South African Red Data Book for birds lists a number of species that are on the edge of their range in southern Africa but are widespread throughout the rest of the Afrotropics and do not feature on the global list of threatened species like the Wattle-eyed Flycatcher (*Platysteira peltata*) and African Broadbill (*Smithornis capensis*) (Brooke 1984, Hurford et al. 1996). Similarly, there are some species that are listed as globally threatened yet do not appear on the South African red list because they are common within the region, for example, the Buffstreaked Chat (*Oenanthe bifasciata*). Globally rare or threatened species are likely to be the greatest priority for conservation but there are arguments for protecting locally rare species as well. To maintain the full range of genetic diversity may require the conservation of populations across the whole of a species’ range and populations which are at the edge of their species’ range may also be more able to adapt to environmental change compared to core populations (Hunter & Hutchinson 1994). There is also the possibility that loss of populations at a local level might affect ecosystem functioning and lead to the loss of other (possibly globally rare) species (Hunter & Hutchinson 1994). Most conservation organisations operate within a political framework and a system globally implemented with respect to biogeographical regions, while ideally being more appropriate would in reality be impractical (Hunter & Hutchison 1994).

*Species richness.*

A commonly used measure of biodiversity value has been simple species richness (for example, Myers 1988, Prendergast et al. 1993, Williams et al. 1996a), but it is debatable whether it is an adequate measure of conservation value or priority. It is attractive because it is relatively easy to measure and is quantifiable (Magurran 1988). However,
species rich communities are often ecotones and not typical habitats (Rapoport et al. 1986), or disturbed habitats where many of the species are opportunistic invaders. From an evolutionary viewpoint, hybrid areas, consisting of regions of overlap between two or more biotas may be more species rich than the ‘parent’ areas, but these hybrid areas may represent marginal habitats (Harrison pers comm, Brooks et al. 1992). Additionally, richness can only be used where sites are comparable in terms of habitat type, for instance, reedbeds and peatbogs tend to be species poor but contain many rare species unique to these habitats (Magurran 1988).

If all species are ultimately threatened then a strategy aimed at protecting areas of high species richness might be valid, but selection of high richness areas does not guarantee representativeness. Alternatively if species richness is a good indicator of other attributes like endemism, rarity or threat, then richness might be a short-cut method for evaluating areas. This is most likely to be the case when species richness is split into its components i.e. when alpha, beta and gamma diversity are measured as separate attributes. Alpha diversity describes diversity within individual, ‘homogenous’ habitats, beta diversity the spatial turnover of species between habitats and gamma diversity the spatial turnover of species in the same habitats across geographical regions (Whittaker 1960, 1972, Magurran 1988). Patterns of alpha diversity could be used to identify areas of high species richness, or to compare the relative diversity of similar habitats. But areas of high species richness may not pick up other attributes of biodiversity like endemic or threatened species. Selecting areas to maximise beta diversity should result in different complements of species being represented so might pick up species that do not occur in the most species rich areas. Likewise, selecting areas to maximise gamma diversity should pick up areas rich in endemic species (Terborgh & Winter 1983).

Another problem with the use of species richness as a measure of biotic diversity is that species are not equivalent units in terms of the genetic, taxonomic and ecological variation they represent.
Taxonomic diversity.

Each species differs in the amount of ‘evolutionary information’ that is unique to it compared to that which it shares with other species (Vane-Wright et al. 1991, Witting & Loeschcke 1995). A group of more distantly related species can be quantified as representing more genetic character diversity than an equal number of more closely related species (Vane-Wright et al. 1991, Williams et al. 1991, Williams 1993, Williams & Humphries 1994). Since biodiversity also involves genetic level diversity, then areas selected to maximise the representation of biodiversity need to take account of the amount of phylogenetic diversity that would be represented and not just the species level diversity. Use of taxic weighting predicts genetic diversity in the absence of direct information, avoiding some of the limitations of diversity measures based solely on taxon richness (Vane-Wright 1996). For example the endangered Australian dragonfly Hemiphlebia mirabilis is the only member of its genus and may be the living sister group of all other extant Odonata. If only species richness is applied then H. mirabilis is just one species among many other dragonfly species; weighted according to its taxonomic rank its significance for the conservation of Australian dragonfly biodiversity would be much more pronounced (Vane-Wright 1993).

However phylogenetic diversity can be measured in different ways, and there is the question of how to value different species relative to one another (May 1994). Topological methods are based solely on the number of nodes between a species and the root of a monophyletic tree. Species with many nodes between them and the root of the tree are given the lowest scores and those closest to the root of the tree the highest scores, but it is not clear how to weight species (Vane-Wright et al. 1991, Faith 1994, Krajewski 1994). For example should the two species of Tuatara, the sister group to all other living reptiles, be given equal weight to the other 6800 species of reptile together (Daugherty 1990, May 1990, Vane-Wright et al. 1991)? If information on branch lengths is available then this, rather than branching points can be used to derive a measure of character divergence between species (Krajewski 1994, Williams et al. 1994b). Branch lengths are proportional to the character differences between species, so that species with longer branches make a disproportionate contribution to group diversity (Krajewski 1994).
Additionally there is the problem of how to value taxonomic differences in the context of conserving biodiversity. One goal of conservation might be to maximise future evolutionary potential - that is to protect areas where lineages originate innovations in their evolution (Erwin 1991, Fjeldså 1994). To achieve this, priority areas should be those where current evolutionarily dynamic lineages occur and where future biodiversity will be created (Erwin 1991). Recent lineages are where radiations and dynamic changes in taxa are currently occurring, however the species involved are often widespread, opportunistic ‘weed’ species (Erwin 1991). Areas rich in such species would have low taxonomic diversity scores as measured by the methods described above which give greater weight to more relict groups. Brooks et al. (1992) suggest that areas of endemism are important because they have been the focus of biodiversity production in the past and may be centres of evolutionary potential for the future. However endemic forms may merely indicate where past evolutionary production has occurred, not present potential, and endemic species may be on the way to extinction (Erwin 1991).

Ecological type diversity.
Species are also not equivalent at the intraspecific level. For example, in the Cape Floral Kingdom, southern Africa, there are over 500 species of Erica all of which have a very similar growth form (Bond 1989). In contrast Acacia karroo is a single species which occurs over a very wide geographical region and in a diverse range of habitats. The species is highly variable among populations both morphologically and as measured by protein electrophoresis (Bond 1989). It might therefore be more important to conserve a number of Acacia karroo populations from different locations or ecological niches than to conserve a number of different Erica species. Extinction of a highly polymorphic species should be considered a greater biodiversity loss than extinction of a monomorphic one (Krajewski 1994).

What surrogate measures can be used to reflect overall biodiversity patterns?
The question of what to conserve is compounded by the second problem - that we only know about a fraction of the biotic diversity that exists. The ratio of known to unknown species has been variously estimated to range from 1:1.67 to 1:21 (Hunter 1991,
McNeely et al. 1990, May 1992). For arthropods, with just over a million species described, total estimates for tropical species vary from between 3-5 million to 10-80 million (Erwin 1982, May 1986, Stork 1988, Thomas 1990). The same problems of poor information apply to other biodiversity measures such as genetic or ecosystem diversity. Any area evaluation will necessarily be based on a subset of the biotic variation in a region. This is inevitably biased towards conspicuous, well studied groups like birds and large mammals (Murphy & Wilcox 1986, Sutton & Collins 1991). Systematic evaluation ideally requires a surrogate measure that will accurately reflect the distribution of diversity in a region across a variety of taxa and different aspects of biodiversity.

A number of different surrogate measures have been suggested, including indicator taxa (Pearson & Cassola 1992, Kremen 1992, 1994, Beccaloni & Gaston 1995), higher-taxon richness (Gaston & Williams 1993, Williams et al. 1994, Williams & Gaston 1994), community, habitat or landscape diversity (Richerson & Lum 1980, Thomas & Mallorie 1985, Austin & Margules 1986, Margules & Stein 1989, Austin 1991a, Noss 1987, Scott et al. 1987), keystone, mobile link or umbrella taxa (Gilbert 1980, Pearson & Cassola 1992, Ryti 1992, Devries 1995) and restricted-range (or otherwise rare species) or endemic species (Terborgh & Winter 1983, Game & Peterken 1984, Thomas & Mallorie 1985, ICBP 1992, Kremen 1994, Balmford & Long 1995). However, the validity of most surrogate measures is rarely established and more often species or groups of species are used opportunistically without consideration of their ability to reflect patterns of diversity in other groups.

Indicator taxa.

These are species or other taxonomic groupings whose distribution accurately reflects the patterns of distribution and richness of a wide variety of other species, or which can be used to monitor changes to ecosystem structure, function and composition in response to natural factors or anthropogenic influences (Kremen et al. 1993). Only the former use will be discussed here - the use of indicator taxa to identify biogeographic zones, areas of endemism, evolutionary centres of radiation and patterns of species richness and complementarity (Erwin 1991, Ryti 1992, Kremen et al. 1993). Indicator groups are
most commonly used as a surrogate for species richness in other taxa but they could be
used to indicate genetic, community or ecosystem level structure or diversity (Kremen
1992). Suitable indicator groups are likely to possess traits such as habitat specificity of
individual species, wide distribution in a region for the taxon as a whole and susceptibility
to disturbance (Kremen 1992, Pearson & Cassola 1992, Kremen 1994). For example,
Pearson & Cassola (1992) propose that tiger beetles (Coleoptera; Cicindelidae) are good
biodiversity indicators. Patterns of species richness in tiger beetles are highly correlated
with those of other invertebrate and vertebrate taxa (butterflies and birds). But this was
at a scale of 275 x 275 km², which is much larger than the size of a reserve could be.
Features of tiger beetles which make them good indicator species include that they are
a well known taxon, easily observed, have a broad geographic range but individual
species are specialised and habitat specific, a stable taxonomy and patterns of species
richness that are highly significantly correlated with other taxa.

Butterflies may also be good indicator species, particularly of plant species richness due
to the interactions of both larvae and adults with specific host plant species (Murphy &
identify particular, notably highly endemic species assemblages, sensitive to ecological
boundaries, and which correlate strongly with the diversity of other organisms in differing
local ecosystems on Madagascar. Potential target taxa included a species rich genus
containing a large proportion of endemic species which appeared to have undergone an
extensive evolutionary radiation in the region (Kremen 1992, 1994). The genus was
characterised by a high beta diversity, with many of the species restricted to a particular
forest microhabitat (Kremen 1992). However while butterfly assemblages were found
to be good indicators of environmental heterogeneity they were poor indicators of plant
species richness and diversity (Kremen 1992).

Beccaloni & Gaston (1995) evaluated the indicator properties of a number of Central and
South American butterfly families in terms of their ability to reflect overall butterfly
species richness at local and regional levels. They found that, for a number of families,
the proportion of species occurring in that family was relatively constant across areas and
not dependent on the total number of species in an area or the size of the area. This means that it is possible to use the average proportion of an indicator species across all areas to predict the total species richness of an area with a known number of indicator species. Beccaloni & Gaston (1995) took the subfamily Ithomiinae, consisting of 309 species, as a potential indicator group and demonstrated that there is a strong positive correlation between ithomiine species richness and overall species richness for all other butterflies across sites (0.044-<314 km\(^2\)), countries (4,828-1,285,215 km\(^2\)) and the Neotropics as a whole (17,500,000 km\(^2\)).

More generally, arthropods may be good indicators for both inventory and monitoring purposes due to their high diversity, rapid growth and evolutionary rates and the large number of functional niches and microhabitats they occupy (Sutton and Collins 1991, Kremen et al 1993, Launer & Murphy 1994). However the consequences of selecting areas on the basis of such indicator groups have rarely been tested to determine whether arthropods are good indicators of general patterns of biodiversity at scales suitable for conservation planning (Kremen 1992, Launer & Murphy 1994).

Launer & Murphy (1994) looked at the effect of a conservation strategy based on the Bay checkerspot butterfly (\textit{Euphydryas editha bayensis}) on the conservation of other grassland community elements in central California. The diversity of native, spring-flowering forbs was high on patches occupied by the butterfly, and rare plant species were found on at least one of the patches where the butterfly occurred. Therefore if all the habitat patches occupied by the butterfly were protected in their entirety then most of the grassland species would also be protected. However the distribution of the plant and butterfly species appears to be influenced by different factors and the overlap in plant-butterfly distributions may in fact be temporary thus reducing the effectiveness of the butterfly as an indicator of plant richness patterns.

\textit{Restricted-range species.}

Restricted-range species are predicted to be among the species most vulnerable to extinction and so should be targeted for conservation action. If restricted-range species
also tend to be concentrated in areas of high species richness they could also be good indicators of other elements of biodiversity. A number of studies, at different spatial scales, have illustrated a positive correlation between the number of restricted-range species and absolute species richness. For example, richness of restricted-range vascular plant species was positively correlated to overall species richness in the southern Appalachian mountains (White et al. 1984). ICBP (1992) has advocated the use of restricted-range bird species as indicators of priority conservation areas. Restricted-range bird species were defined as those with a global breeding range of less than 50,000km² (ICBP 1992). Priority areas for conservation were identified on the basis of the distribution of such species and it was found that the areas with concentrations of restricted-range bird species were generally also important for a range of other taxa (ICBP 1992). Additionally, the number of these restricted-range bird species at a country level was found to be a good predictor of national levels of endemism, and to a lesser extent, species richness and threat in other animal and plant groups in the tropics (Balmford & Long 1995). However, when area effects were controlled for there was no longer a significant correlation between restricted-range bird species richness and bird, amphibian, swallowtail or threatened mammal species richness (Balmford & Long 1995).

However, use of restricted-range as an area selection criterion may be biased against taxa which are widespread, but exist at low densities and/or are habitat specific. Several families of landbirds which do not contain any restricted-range species, include species that fall into this category - notably cranes, storks and bustards (Gaston 1994). Another potential problem in the use of restricted-range species as a surrogate for wholesale diversity is that it requires information on the distribution of species, both at the level at which the restricted-range criterion is being assessed (usually at a biogeographical scale), and also at the level at which area prioritisation is taking place.

*Ecosystem representativeness.*

Evaluation of areas based on factors like the geographic range of the component species will only be possible when the biota in question is well known. Such a method is likely to be biased towards well studied and conspicuous groups of organisms like birds, and
the areas selected will similarly be those which have the highest values for these species. A network of sites that collectively represent the range of habitat or environmental types within a region might also represent the range of species and other biodiversity components found in the region, without the need for detailed species distribution data.

According to Noss (1983) the long term maintenance of biodiversity requires a strategy that considers regional biogeography and landscape patterns with the goal of perpetuating natural ecosystems. Maintaining the landscape mosaic is more appropriate than single sites, therefore reserve networks should aim to maintain high quality examples of pre-settlement ecosystems in approximate proportion to their former abundance, rather than maximising habitat or species diversity. A community approach to the selection of sites for conservation involves selecting priority areas with the aim of representing the range of community or landscape types in a region (e.g. Bolton and Specht 1983, Purdie 1987, Carrier & Smith 1989, Walker 1992, Faith & Walker 1996). For example, Purdie et. al (1986) analysed resource maps of land system and vegetation in Australia and selected key areas on the basis of ecosystem diversity and representativeness as well as rare ecosystems. Sites were narrowed down after field analysis and 3.1% of the region selected in this manner could incorporate 92% of the major vegetation types and most of the rare species.

Environmental representativeness can be assessed by gradient analysis of vegetation-environmental variation (DeVelice et al. 1988). For example, Taggart (1994) used ordination techniques to delimit different savanna communities in Carolina, North America on the basis of vegetational and environmental variation. This enabled an assessment of the extent to which the current protected areas represented the range of community variation in the region. In this way additional sites could be identified that would fill in gaps in the representation of community diversity. Similarly, Saetersdal & Birks (1993) used multivariate analyses to delimit representative types of woodland in western Norway and to identify suites of species associated with these woodland types. It was then possible to determine whether any major groups of species were under represented by the existing reserve network. Whitehead et al. (1992) looked at
environmental correlates of avian distributions and from this classified geographically coherent groups of species thus allowing reserve selection based on maximising the range of climatic options available to species. This could be important, since if climatic conditions change, the areas currently important for biodiversity may shift. Thomas & Mallorie (1985) found that when knowledge of the species in question is incomplete, the best approach is to preserve a complete sample of ecosystems. For Moroccan butterflies this proved to be only slightly less effective than an approach based around conserving as many restricted-range species as possible.

Margules & Stein (1989) analysed species distribution patterns in New South Wales, Australia and found that the best approach to represent the species in the region would be to select sites that encompassed the full range of environmental variation along a steep rainfall-temperature gradient across the region. Plant communities are predicted to be one of the most efficient surrogates for overall species diversity because vegetation consists of the principal autotrophs upon which most other organisms depend and it integrates many important environmental gradients (Lesica 1993).

Problems with this approach may arise if different taxa respond to different environmental variables or at different levels of resolution. Linder (1991) found that patterns of species richness in the south-western Cape Province, southern Africa were similar for a variety of plant taxa and correlated to a similar range of environmental variables, notably precipitation. Similarly, distributions and species richness of frogs, snakes, lizards, large mammals and birds show similar patterns in southern Africa which are related to the rainfall regime (and hence vegetation) (Crowe 1990). However, Daniels (1992) found that while angiosperm and amphibian species richness were broadly related to levels of precipitation, in the Western Ghats, India, local patterns of amphibian species richness and endemism are different from those of angiosperms and birds and this may be due to differences in dispersal ability and in responses to environmental conditions.

However concentration on communities as a way of maintaining high levels of species diversity may only be useful as a coarse filter approach in contrast to the fine filter
approach of single species protection (Hunter et al. 1988). The method might not be a fine scale enough to pick up threatened species or areas of high species richness for example. Hunter et al. (1988) also point out that since most North American communities are mostly less than eight thousand years old and are only transient assemblages, rather than highly organised units reflecting long-term co-evolution, site selection should focus on the distribution of physical environments rather than communities. In a similar way it has been suggested that since patterns of diversity are to some degree attributable to climatic and topographical patterns then one approach to preserving diversity would be to focus on landscape heterogeneity (eg. Richerson & Lum 1980). However this is an even coarser scale method than the use of community or ecosystem level diversity to select areas. Clarke & Bell (1986) analysed biotic community representation in Malawi's system of reserves and concluded that quantitative representation of biotic communities alone was an inadequate guide to conservation priority because the classification of communities is subjective as is any decision over what constitutes an adequate level of representation. Secondly the method is insufficiently selective in terms of the protection of rare or threatened species. The approach is useful for revealing under-representation of communities at a gross scale, but is weak at a finer scale especially where rare species of restricted geographic range are concerned, and is also biased towards plants, since ecosystem classification is often based on vegetation associations or even structure.

**Higher-taxon richness.**

Given the paucity of information about the distributions of many species, one potential surrogate for species information might be to use distribution data at a higher taxonomic level than that of species. Positive correlations between species richness and higher taxon richness (typically families) have been documented for a variety of regions and taxonomic groups (e.g. Gaston & Williams 1993, Williams & Gaston 1994, Williams et al. 1994a). If this relationship holds at the scale at which sites are evaluated then it might be a shorthand method of evaluating the relative species diversity of sites. Sites will have fewer higher taxa than species and as a group, the higher taxa will be better known.
Gentry (1988) found that plant species richness in 0.1 hectare plots increased with an increase in family richness and this relationship held for islands, on most continents and worldwide. But this relationship may not always hold. Plant species richness is much higher in the Neotropics compared to the Palaeotropics (Prance 1994). However family level richness is not greater in the Neotropics and generic level richness is only slightly higher (Prance 1994). The ratio of species to families is approximately 308:1 in the Neotropics, 157:1 in Africa and 135:1 in Malesia (Prance 1994). Additionally, centres of endemism and centres of diversity identified at the species level are not evident when only higher taxa are used in the Neotropics (Prance 1994). At a finer scale, the relationship between species and genus richness was strong for Australian ant faunas within regions, but varied substantially across regions. In addition the relationship was confounded by biogeographical factors, sampling intensity and area effects (Andersen 1995). Andersen (1995) concludes that higher taxon richness is not a good surrogate for ant species richness in Australia and more generally for taxa where a relatively small number of genera contribute a large proportion of the total species; however this is exactly the sort of relationship that is needed for a good surrogate measure. While the relationship may hold for taxa where there is a low species - genera ratio this will not reduce the level of sampling required by much (Andersen 1995).

_Umbrella and keystone species._

Species with large geographical ranges or those which utilise a large range of habitats within a region might act as an umbrella for the majority of other species in the region. Measures designed to conserve the umbrella species may coincidentally also benefit a large number of other species. For example, Devries (1995) proposes the use of large herbivores to determine the design of large-scale reserves in western Europe. Large herbivores require large areas of land and also have a major influence on the vegetation composition of an area. They could therefore act as an umbrella for the preservation of other plants and animals and would also counter the increase in habitat fragmentation occurring in many areas. A similar idea is the use of keystone or mobile link species (Gilbert 1980). Keystone species are those that provide an essential function in a community (e.g. provide nectar or leaves) while mobile link species are for example
pollinators or species that disperse the seeds of a large number of plants (Gilbert 1980). Protection of these sorts of species may also ensure the persistence of all the other species dependent on them.

*Combined surrogates.*

Use of just one surrogate measure may not be sufficient to pick up all aspects of biodiversity. In particular, the use of community or environmental diversity may be too coarse an approach. To avoid these sorts of problems Clarke & Bell (1986) suggest a combined approach: within the framework of broad representation, sites should be selected on the basis of rare, endemic and threatened species. Similarly Hunter (1991) established representative samples of communities and then used a fine filter approach to capture rare and threatened species. Klopatek et al. (1981) used four criteria for evaluating areas: vegetation, avian communities, mammal communities and endangered or threatened species. All these criteria were evaluated at three levels: national, regional and ecoregional and combined into one index. Many areas were high ranking for individual parameters, but not overall for the combined index.

One method in which community and species level diversity could be combined is to use surrogate taxa to identify areas of high diversity within different biogeographical regions. Rebelo & Siegfried (1992) looked at the efficiency of rarity and richness algorithms at selecting reserve areas for flora in the Cape Floristic Region, South Africa. They concluded that the best strategy is to identify areas of high richness and endemism within biogeographic zones and to protect these. Crowe & Crowe (1982) and Crowe (1990) used cluster analysis to identify biogeographic zones characterised by coherent groups of species. This enabled centres of unexpectedly high species richness or endemism within uniform biogeographic areas to be identified.

**How should biodiversity value information be used to select priority areas?**

In selecting priority areas for conservation there is also the question of how values should be assigned to sites and how sites should then be selected.
**Scoring approaches.**

In the past the most commonly employed method has been to assign values to criteria and combine them into some sort of index (where multiple criteria are used) (Usher 1986). For example Gehlback (1975) scored sites for heritage value, educational utility, species significance, community representation and human impact and added these values together to give a site score.

However there are a number of problems associated with the use of scoring indices. They are affected by the distribution of attributes within sites, inclusion of less valuable sites within the mainland habitat and the fact that different taxa are likely to have different values, so some kind of species or other taxonomic weighting is necessary and this will tend to be arbitrary (Götmark et al. 1986). There is also the problem of what values to give to criteria, and the relative importance of each attribute within the index. The efficiency of criteria (i.e. the number or area of highest scoring sites required to represent all natural features) can be influenced by whether scores or score classes (i.e. ordinal values) are used (Bedward et al. 1991). The precise effect differs with the criterion and with the data set and is difficult to predict *a priori* (Bedward et al. 1991). Additionally two sites could have the same score but for very different reasons because the values given to the different attributes are hidden within the agglomerative site score. While indices may not be sufficient for complete evaluation they might provide a means of arranging similar sites in some sort of priority order.

The main problem with using a scoring selection method is that sites are selected without reference to what has already been represented in other sites. In an extreme case it is possible that all the top scoring sites have exactly the same species complement, and so species only occurring in low scoring sites would not be represented at all. This problem can be partially solved by identifying different species assemblages and selecting sites to represent each assemblage. For example McKenzie et al. (1989) identified species assemblages, then used isolines to plot species richness and selected areas which contained a range of assemblage richness nodes plus areas with unique assemblages.
However if you want a network that maximises the representation of biodiversity as efficiently as possible then you need to consider complementarity between sites in the network. This may mean that the lowest scoring site can be selected if it contains unique aspects of diversity.

Iterative selection procedures.

Iterative site selection methods aim to achieve 'maximally efficient' solutions for fully representative reserve networks, either as the minimum number of areas, or smallest configuration of sites, capable of providing sustainable protection (living space) for all species found in a given region (Vane-Wright et al. 1994). Sites are selected in a stepwise fashion, at each step maximising the number of previously unrepresented species or other attributes that are added to the network.

The application of systematic methods to conservation requires an assessment of the complementarity and irreplaceability of attributes across sites (Vane-Wright et al. 1994). Iterative selection methods are based on maximising the complementarity between the sites in a reserve network and therefore achieving representation of all the desired attributes in the 'minimum' area or number of sites (Pressey et al. 1993). Kirkpatrick (1983) first applied an iterative algorithm to the assessment of priorities for threatened species in the central east coast of Tasmania in an attempt to get over the problem that when weighted attributes and applied formulae are used in a single step process then different areas may contain the same species while other species may go unrepresented. Ackery & Vane-Wright (1984) also used an iterative selection process in their critical faunas analysis. They identified the minimum set of areas or faunas which would contain at least one population of each of 157 species of milkweed butterfly in 350 areas. The process involved picking the areas containing most endemics, then adding areas to complete list for all other species - this required just 31 sites. A number of different iterative algorithms have now been developed eg. (Margules et al. 1988, Pressey & Nicholls 1989b, Csuti et al. in press).

An iterative approach is invariably more efficient than scoring sites - that is it requires a
smaller number of sites to achieve representation of all attributes (Pressey & Nicholls 1989a, Bedward et al. 1991, Kershaw et al. 1994). Although iterative algorithms are highly efficient - that is they can find close to the minimum number of sites that would be needed to represent all the species in a region - this does not necessarily imply that this number of sites is sufficient to ensure the persistence of all the species. However iterative selection algorithms can be modified to incorporate viability factors like multiple representation, contiguity of sites and availability of land (Bedward et al. 1992, Williams et al. 1996a). When insufficient sites to represent all species can be included in a reserve network then iterative algorithms can be used to ensure close to the maximum possible attributes are included within the number of sites that can be selected.

A problem with such algorithms is that the process of iteration by definition comes up with sequences of sites, so that selection of a particular sites is constrained by the sites selected previously. If one of these selected sites cannot be reserved then the other sites in that network may no longer be the best sites. Iterative algorithms have also been criticised on the grounds that they do not necessarily come up with the optimal solution for a reservation problem and in some cases may be 'grossly suboptimal' (Underhill 1994). Instead of selecting sites iteratively it is possible to select sets of sites that are the best solution for a predefined problem (for example represent all the species in the region in the smallest number of sites) using integer programming techniques (Cocks & Baird 1989, Saetersdal et al. 1993, Underhill 1994, Camm et al. in press, Church et al. in press, Csuti et al. in press). Various constraints can then be built into the program, for example, that sites are as contiguous as possible or that species are represented in more than one site. However, while such techniques can guarantee to find the optimum solution to a problem, they are very computer intensive and may take many weeks to find a solution when a simple heuristic can find close to the optimum for the same size data set in a matter of seconds (Pressey et al. 1996).

The three major problems that face priority area selection (what to conserve, what surrogate to use, and how to use biodiversity value information to prioritise areas) are discussed in the following chapters of this thesis. In chapter two the effect of using
different biodiversity attributes to select priority areas is investigated for a number of
different taxonomic groups in KwaZulu Natal, southern Africa and for Afrotropical
antelopes. Attributes considered include simple species richness, endemic and restricted-
range species, threatened species and taxonomic diversity. In chapter three different
potential biodiversity surrogates are evaluated including, species richness of indicator
taxa, environmental representativeness, restricted-range species, higher taxon richness
and habitat specialist species, with reference to the patterns of species richness and
complementarity for different taxa in KwaZulu Natal. Chapters four and five are
concerned with the selection of priority area networks once sites have been evaluated for
biodiversity value. In chapter four the efficiency of scoring versus iterative approaches
is compared. In chapter five alternative iterative algorithms are evaluated in terms of
their consequences for the representation of biodiversity in a region and in relation to
optimal reserve configurations derived from integer programming solutions.
CHAPTER TWO.

A COMPARISON OF THE EFFICIENCY OF ALTERNATIVE AREA SELECTION CRITERIA AT REPRESENTING DIFFERENT BIODIVERSITY ATTRIBUTES.

INTRODUCTION.

In practical terms, conservation is directed at areas, but areas are most easily evaluated in terms of the species that they contain. It is generally agreed that the rôle of protected areas is to maintain a representative selection of biotic diversity; however, this can have a number of different meanings (Vane-Wright et al. 1991). For example, species richness, taxonomic diversity, genetic diversity, community/ecosystem diversity, evolutionary potential and threatened or vulnerable species are all components of biodiversity that could be used to assess the conservation value of an area (Vane-Wright et al. 1991, Fjeldså 1994, Erwin 1991, Chapin et al. 1992). However, while a universal aim of priority areas might be to support a representative selection of biodiversity, it is not obvious how to translate this aim into practice. Use of alternative biodiversity criteria to evaluate potential protected areas could have very different consequences in terms of the areas selected and the attributes that will be represented by these areas. Problems will arise if a set of areas selected to maximise one aspect of biodiversity, e.g. species richness, is not the best for other aspects of diversity, e.g. threatened species (Kershaw et al. 1995).

Three different approaches are currently at the heart of evaluations of area-based priorities. First, much documentation exists at the species level, in national and international lists of threatened species (e.g. IUCN 1992) and areas containing one or more threatened species might be an appropriate focus for conservation action (e.g. Collar & Stuart 1985). Secondly restricted-range or endemism is considered a good indicator of vulnerability to extinction (Terborgh & Winter 1983, Arita et al. 1990, Thomas 1991) and ICBP (1992) have recommended that conservation action for birds be focused on areas that contain a large number of restricted-range species. Thirdly, a
number of recent studies have advocated area-based priorities developed from the overall species richness found at each site (e.g. Pressey & Nicholls 1989b, Lombard et al.1995, Csuti et al. in press), which can include species weighting for some kind of taxonomic distinctiveness (Vane-Wright et al. 1991, Williams & Humphries 1994).

These three approaches are all focused on preserving biodiversity, but their consequences are not necessarily the same. A focus on threatened status may be seen as a short-term measure to limit extinctions, and could detract from a longer term aim of preserving a taxonomically representative sample of taxa. This is especially likely to occur if threatened status is common among particular ecological, taxonomic or trophic categories. Use of restricted-area species can be a very practical approach where distribution data are available but could be biased against widely distributed, but threatened forms, and have the longer term consequence of losing diversity if restricted-range species are common in particular taxonomic groups, or simply occur in areas that are low diversity. Conversely, a focus on species richness might lead to greater diversity in the long-term, but could fail in its aim through not recognising that some elements of the species complement are at high risk of imminent extinction.

In this chapter I investigate the consequences of basing site selection on different biodiversity criteria for two, contrasting data sets. First, at a biogeographical scale I use species diversity (simple species richness and taxonomic diversity measures), richness of restricted-area forms, and numbers of threatened species to prioritise areas for Afrotropical antelopes in sub-Saharan Africa. In this data set the area selection units are discrete areas within the Afrotropical region, and sites are selected solely on the basis of one taxonomic group, i.e. antelopes. Secondly, across a smaller political unit, site networks were selected for birds, mammals and plants in KwaZulu Natal, southern Africa using species richness, richness of endemics, richness of species that are sparsely distributed within the political unit, and numbers of threatened species for each taxon separately. Here, a range of different taxa is considered, with information on their distributions across the entire political region. For both examples the species complement represented by each of the selected site networks is evaluated. In particular I focus on
the following questions:

1. How effective are networks of sites selected using alternative biodiversity criteria at selecting sets of species that score highly for the other biodiversity attributes? Are there any generalisations that can be made across different geographical scales and for different taxa?
2. Are the same areas selected by each of the different criteria?

METHODS.

Afrotropical antelopes.

Data.

The analysis focused on 99 species and subspecies of Afrotropical antelope distributed across 249 sites in sub-Saharan Africa, compiled from the IUCN's Global survey and regional action plan (East 1988, 1989, 1990). The sites were all existing or proposed conservation areas in sub-Saharan Africa. All the antelope subspecies included were recognised as distinct by wildlife managers in the field and also by the IUCN/Species Survival Commission/Conservation Assessment Management Plan (CBSG 1992).

Only species occurring in 'substantial' populations in each site were included in the data set. East (1988, 1989, 1990) defined substantial populations according to how rare he perceived species to be; so that for endangered species a substantial population was defined as greater or equal to 100 individuals; for rare/threatened species, greater or equal to 500 individuals and for species considered satisfactory, greater or equal to 1000 individuals. In cases where population numbers were unknown, substantial populations were assumed to occur in areas where the species/subspecies was known to be widespread and relatively common, and where there was an extensive area of suitable habitat (East 1988, 1989, 1990).

Following a Conservation Assessment and Management Plan (CAMP) workshop in September 1992 all the species and subspecies included in the antelope data set had been assigned to one of five conservation status categories, critical (5 species), endangered
(8 species), \textit{vulnerable 1} (11 species), \textit{vulnerable 2} (17 species) and \textit{not of concern} (58 species), according to the perceived degree of threat (CBSG 1992). These categories were based on the then proposed system by Mace & Lande (1991) which defined three categories of threat - \textit{critical} (50% probability of extinction within five years or two generations, whichever is longer), \textit{endangered} (20% probability of extinction within twenty years or ten generations whichever is longer) and \textit{vulnerable} (10% probability of extinction within 100 years). In addition the CAMP workshop recognised a fourth category of threat - \textit{vulnerable 2}. Species that did not qualify for any of the threatened categories were placed in a \textit{not of concern} category.

\textbf{Analysis.}

\textit{WORLDMAP} (version 2.5) computer software was used for the analysis. This is a graphical software package designed to allow rapid assessment of priority areas under various measures of biodiversity (Williams 1992). Species distributions are stored by area, and areas can be scored to give species richness values. Species can also be weighted by a measure of their taxonomic distinctiveness or by how restricted their ‘range’ is (derived from the number of sites a species is recorded in), and areas then scored by these values. \textit{WORLDMAP} calculates the cumulative percentage diversity (according to the specified weighting) represented as sites are selected and can therefore be used to calculate the percentage of the total diversity represented in a set of areas.

The 249 sites were scored according to the following criteria:

\textit{Unweighted species richness.}

This was simply measured as the number of species and subspecies occurring in each site.

\textit{Taxonomic diversity measures.}

Taxonomic weightings for each species and subspecies of antelope were derived from a classification based on a cladogram of fifteen species of antelope by Gentry (1992) (Williams pers. comm., Kershaw et al. 1994) (Figure 2.1). Figure 2.2 gives an example of how the taxonomic weightings were calculated. \textit{WORLDMAP} can be used to calculate several different measures of taxonomic diversity, including the three described
Dispersion-weighted species richness.
This is a taxonomic weighting which aims to maximise the range of subgroups in the classification and to ensure an even representation of the classification that reflects the shape of the tree (see Williams 1993, Williams & Humphries 1994). A highly dispersed fauna will be characterised by a high mean divergence among the species (as calculated by the number of nodes separating the species, i.e. a large number of different subgroups) and a narrow standard deviation of divergences around this mean (i.e. the different subgroups are evenly spaced across the classification) and this can be used to weight the faunas of each site (Williams et al. 1991, Faith 1992, Williams & Humphries 1994). (More recently a set of p-median measures have been used to quantify regularity which give more consistent results (Faith 1994), but these were not available for this analysis).

Higher-taxon weighted species richness.
This measures diversity as species richness with higher weights for faunas containing species from early diverging, higher taxa, irrespective of the number of extant species in these higher taxa. For example the two extant species of tuatara represent the sister group to all other reptile species. A site containing a tuatara species would therefore have a large higher-taxon score relative to a site with the same number of reptile species, but no tuatara species.

Spanning-subtree weighted species richness.
This measure gives higher weighting to faunas with species that represent the greatest proportion of subgroups (branching points) at all levels of the classification. This is similar to the dispersion-weighted diversity measure, but does not attempt to ensure evenness of species representation across the whole tree.
Figure 2.1. Classification of Afrotropical antelopes (based on Gentry 1992) used to derive taxonomic diversity measures in WORLDMAP.
Figure 2.2. Example illustrating the formulation of codes used to enter a taxonomic hierarchy into WORLDMAP. The program weights different faunas according to the number of taxa and the number of internodes subtended by the included species.
Restricted-range weighted diversity.

This measure weights the fauna of each site according to how unique it is relative to the faunas of other sites. WORLDMAP calculates restricted-range diversity for each fauna as the sum of the inverse number of sites each component species is found in, expressed as a percentage of the sum of the scores for every species in the data set. The restricted-range score therefore takes account of both the absolute number of species in a site and how geographically restricted (in terms of the total number of sites occupied) its fauna is. Areas with many species that occur in one or only a few sites will have the highest restricted-range scores. This score is not strictly a measure of restricted-range since it is based solely on the number of sites a species is found in and does not take account of the location or contiguity of these sites. For example a species occurring only in two sites that are widely separated would have a higher restricted-range score than a species found in three contiguous sites.

\[ \text{RANGE SIZE SCORE} = \frac{\sum_{i=1}^{n} \frac{100}{X}}{\sum_{i=1}^{n_t} \frac{1}{X}} \]  

(1)

Where \( X \) is the number of sites occupied by each species; \( n \) is the number of species in the site; and \( n_t \) is the number of species in the complete data set.

Richness of threatened species.

This was simply the number of species in any of the four categories of threat (critical, endangered, vulnerable 1 and vulnerable 2) that occurred in each site.

Once sites had been assigned values based on each of the above criteria, priority sets of sites were selected from the total 249 sites in the data set. These consisted of the top twelve sites (approximately 5 percent of the total number of sites) ranked for each of the different diversity criteria. Selecting five percent of sites for each network was an arbitrary number of sites but is close to the percentage land area that is currently protected at a global scale (Groombridge 1992). Where there were ties between sites
there was the potential for different site networks to be selected. Ties between sites were broken using one of the alternative diversity criteria (progressing through the list of criteria until the tie had been resolved). Use of one of the other diversity criteria to break ties will tend to make the different site networks more similar than if a site were selected at random, for example, from the tied sites. For instance, for the threat selection, if two sites had the same number of threatened species, but one had more species overall, then this site would be selected. This means that the complete site network may contain more species than if a site had been selected at random and the lower richness site had been selected. Sets of selected priority areas are referred to as site networks. The selected site networks were evaluated in terms of their ability to represent different aspects of diversity, i.e. for each of the priority site networks, the following evaluations were made for the first six and twelve sites selected:

1. Percentage of total species represented.
2. Percentage of total taxonomic diversity (dispersion, higher taxon and spanning subtree weighted, calculated using the WORLDMAP software).
3. Percentage of total restricted-range diversity represented.
4. Percentage of all those species listed as threatened.

This enabled an evaluation of the consequences of using different diversity measures to select sites in terms of the representation of other aspects of biodiversity.

**Birds, mammals and plants in KwaZulu Natal, southern Africa.**

**Data.**

Distribution data for birds, mammals (carnivores and ungulates) and plants in KwaZulu Natal were used in the analysis (Table 2.1, Figure 2.3). Bird distribution data for 570 species were obtained from a bird atlas, compiled from ten years of fieldwork (Cyrus & Robson 1980). Only 23 species that had been previously recorded from KwaZulu Natal were not recorded during the atlassing period. All vagrants, introduced species and pelagic species that were only seen from the coast and did not use the province to breed were excluded from the analyses. Mammal distribution data for 65 species came from
an atlas of ungulates and an atlas of carnivores (Rowe-Rowe 1992a, 1992b, 1994). These contained all available species records obtained from a variety of sources (field records, literature, museum specimens etc.). Plant distribution data for 6111 species, subspecies and varieties were collated from the computerised herbarium database PRECIS (National Herbarium, Pretoria (PRE) Computerised Information System (CIS)), maintained by the National Botanical Institute. This contains locality data for all indigenous and naturalised plants recorded from southern Africa, and Natal is one of the most comprehensively covered provinces in the subregion (Gibbs-Russell et al. 1984, Gibbs-Russell 1985).

Distribution data for all the taxa were mapped at a quarter degree grid square resolution (24 x 27.5 km in South Africa, Osborne & Tigar 1992) which gives squares of 660 km². KwaZulu Natal is covered by 166 of these squares. This represents a smaller area than the mean size of protected areas globally which is 911 km² (Groombridge 1992), although the median size is much smaller than this, and in practice it is unlikely that blocks of land 660 km² in area could be set aside exclusively for conservation purposes.

Threatened species data for birds, mammals and plants came from southern African red data books and international red data books (Hall et al. 1980, Brooke 1984, Collar & Stuart 1985, Smithers 1986, Branch 1988, Collar & Andrew 1988, Collar et al. 1994). All species listed in either the international or the southern African red data books were treated as threatened for the purposes of this analysis, including species listed as near threatened, indeterminate and uncertain (Table 2.1). Threatened plant data were undoubtedly incomplete.

Data on which species were endemic to the southern African subregion were only available for birds and grasses (Poaceae). Information on endemism for birds was obtained from field guides (Cyrus & Robson 1980, Sinclair et al. 1993) and for grasses from ‘Grasses of southern Africa’ (Gibbs-Russell 1990). Sixty-eight bird species that occurred in KwaZulu Natal were endemic to the southern African subregion. Bird species classed as endemic to southern Africa were defined as those species confined to the Republic of South Africa, Botswana, Namibia, south and central Mozambique,
Zimbabwe and south west Angola (Sinclair et al. 1993). Endemic Poaceae (66 species) were species that occurred only within the Republic of South Africa, Botswana, Namibia, Mozambique, Zimbabwe, Angola and Zambia (Gibbs-Russell 1990).

Species that occurred in only a few sites within KwaZulu Natal (irrespective of their distribution elsewhere) were classified as restricted-range species. Whereas the true endemic species could be viewed as a global priority for conservation, species with restricted-ranges in KwaZulu Natal might be considered a national/local priority (within KwaZulu Natal) and if these species also had sparse/restricted distributions outside KwaZulu Natal they might also be a global priority, although this will not necessarily be the case. Restricted-range species were defined as the bottom 10 percent of species ranked by the number of sites in which they occurred. Where the bottom 10 percent of species did not fall on a natural cut off point, the number of species defined as restricted was determined by the nearest natural break in the data. Using this definition, restricted-range bird species were those recorded from less than five of the 166 sites, restricted-range mammals were those recorded from less than seven sites, and restricted-range plants (and grasses) were those species recorded from one site only (Table 2.1). Using the bottom 10 percent of species to define restricted-range was an arbitrary criterion but resulted in a similar number of species compared to the number of endemic species in each taxon. Definitions of restricted-range species were not intended to be comparable across the different taxa.

Table 2.1. Numbers of species, threatened species, endemic species and restricted-range species in each of the different taxonomic groups in the data set for KwaZulu Natal.

<table>
<thead>
<tr>
<th></th>
<th>BIRDS</th>
<th>MAMMALS</th>
<th>PLANTS (GRASSES)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPECIES (TOTAL)</td>
<td>570</td>
<td>65</td>
<td>6111 (411)</td>
</tr>
<tr>
<td>ENDEMIC SPECIES</td>
<td>68</td>
<td>-</td>
<td>- (66)</td>
</tr>
<tr>
<td>THREATENED SPECIES</td>
<td>81</td>
<td>21</td>
<td>124 (-)</td>
</tr>
<tr>
<td>RESTRICTED-RANGE SPECIES</td>
<td>56</td>
<td>7</td>
<td>1407 (66)</td>
</tr>
</tbody>
</table>

Definitions of the different categories are given in the text.
Figure 2.3. Map of southern Africa showing the location of KwaZulu Natal.
Analysis.

For each of the taxa, networks of eight sites (five percent of KwaZulu Natal) were selected according to the alternative diversity criteria. Five percent is an arbitrary level but is close to the percentage area that is currently protected at a global scale. As with the antelope data set, site networks consisted of the top ranking sites for each diversity component, and networks were evaluated in terms of their ability to represent the other components of diversity for the taxon. The amount of diversity represented was calculated for the first four sites and the first eight. To give an indication of how well the different networks performed at representing the different aspects of diversity, 100 sets of random site networks (consisting of eight sites) were generated for each taxon and evaluated in the same way as the real site networks. Random site networks were generated using the random number generator in FORTRAN, seeded with the time for each run of sites. For each eight site selection, sites were selected without replacement so that a site could only be selected once in each network. Ideally, more than 100 random draws of sites would have been made, however inspection of the means and standard errors as n was increased indicated that 100 samples provided a reasonable estimate of the true values.

Summary of the site networks selected for each taxon:

**Birds**

1. Top sites for species richness (570 species total).
2. Top sites for richness of species endemic to southern Africa (68 species total).
3. Top sites for richness of range size-rarities (56 species total).
4. Top sites for threatened species (81 species total).

These different networks were then evaluated in terms of how well they represented (1) species richness, (2) endemic species, (3) species found in less than five sites (range-size rarities) and (4) threatened species.

**Mammals**

1. Top sites for species richness (65 species total).
2. Top sites for richness of range size rarities (7 species total).
3. Top sites for threatened species (21 species total)

These different networks were evaluated in terms of how well they represented (1) species richness, (2) species found in less than seven sites (range-size rarities) and (3) threatened species.

*Plants (all).*
1. Top sites for species richness (6111 species/subspecies total).
2. Top sites for range size rarities (1407 species/subspecies total).
3. Top sites for threatened species (124 species/subspecies total).

These different networks were evaluated in terms of how well they represented (1) species richness, (2) species found in only one site (range-size rarities) and (3) threatened species.

*Grasses (Poaceae).*
1. Top sites for species richness (411 species total).
2. Top sites for range size rarities (66 species total).
3. Top sites for endemic species richness (66 species total).

These different networks were evaluated in terms of how well they represented (1) species richness, (2) species found in only one site (range-size rarities) and (3) endemics. Grasses were also included in the full plant analysis, but the analysis was repeated separately using just grasses since they were the only plant group for which information on endemism was available. This does however mean that any relationships for grasses cannot be treated as independent from results derived for all plants.

The differences in the criteria used to select and evaluate sites for the different taxa merely reflect the information available. As with the antelope analysis, overlap between the sites selected using the alternative criteria was also measured. Data manipulation
and analysis were carried out using FORTRAN programs and DBASE IV PLUS (Borland 1993). The restricted-range criterion used in the KwaZulu Natal analysis was based solely on the number of restricted-range species that were present in each site and did not incorporate any kind of weighting for the absolute number of species present in the site as used for the Afrotropical antelopes.

RESULTS.

Afrotropical antelopes.

Consequences of using different selection criteria for the representation of biodiversity. When the number of sites in each network was limited to six, there were some noticeable compromises in the representation of those aspects of biodiversity that the selections were not explicitly designed to represent (Table 2.2, Figure 2.4). The top sites for species richness and taxonomic diversity were very poor at representing range size rarities. Simple species richness incorporated only 16% of the total restricted-range diversity and the different taxonomic measures represented less than 15% of restricted-range diversity (Figure 2.4). In comparison, the top six sites for restricted-range species resulted in the representation of 36% of these species (Table 2.2). Sites selected on the basis of species richness were also poor at representing threatened species. Only 12% of threatened species were in the top six sites selected using species richness, half the number represented by the threatened species network (Figure 2.4).

In contrast sites ranked highly for restricted-range species performed almost as well as the species richness and taxonomic diversity criteria in terms of representing species richness and most measures of taxonomic diversity (Figure 2.4). The restricted-range criterion incorporated 37% of species richness (compared to 39% in the species richness network) and 36% of dispersion-weighted taxonomic diversity (42% represented by dispersion criterion). Restricted-range diversity performed least well at representing higher taxon diversity (29% compared to 45% represented by the higher taxon selection), suggesting that restricted-range species tended to be from the most recently diverging groups in the classification. Top sites for restricted-range diversity were also effective at incorporating threatened species - in fact the top sites for restricted-range species
represented more threatened species between them than the top six sites for threatened species richness.

Table 2.2. Percentage of different diversity attributes represented in the top six and twelve sites selected using alternative diversity criteria for Afrotropical antelopes.

<table>
<thead>
<tr>
<th>SITE SELECTION CRITERION</th>
<th>TAXONOMIC DIVERSITY MEASURE</th>
<th>SPECIES RICHNESS</th>
<th>RESTRICT-RANGE</th>
<th>THREATENED SPECIES</th>
<th>DISPERSION-WEIGHTED</th>
<th>HIGHER-TAXON</th>
<th>SPAN-SUBTREE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>PERCENTAGE OF ATTRIBUTE REPRESENTED IN THE TOP SIX SITES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPECIES RICHNESS</td>
<td></td>
<td>39.4</td>
<td>37.4</td>
<td>14.1</td>
<td>41.4</td>
<td>40.4</td>
<td>40.4</td>
</tr>
<tr>
<td>RESTRICT-RANGE</td>
<td></td>
<td>16.3</td>
<td>36.3</td>
<td>9.49</td>
<td>14.3</td>
<td>14.9</td>
<td>14.1</td>
</tr>
<tr>
<td>THREATENED SPECIES</td>
<td></td>
<td>12.2</td>
<td>26.8</td>
<td>24.4</td>
<td>17.1</td>
<td>9.76</td>
<td>14.6</td>
</tr>
<tr>
<td>DISPERSION-WEIGHTED</td>
<td></td>
<td>39.2</td>
<td>36.1</td>
<td>11.2</td>
<td>42.0</td>
<td>38.9</td>
<td>40.8</td>
</tr>
<tr>
<td>HIGHER-TAXON</td>
<td></td>
<td>35.1</td>
<td>29.2</td>
<td>14.9</td>
<td>37.2</td>
<td>45.1</td>
<td>35.7</td>
</tr>
<tr>
<td>SPAN-SUBTREE</td>
<td></td>
<td>54.8</td>
<td>51.4</td>
<td>24.7</td>
<td>57.5</td>
<td>55.5</td>
<td>56.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ATTRIBUTE</th>
<th>PERCENTAGE OF ATTRIBUTE REPRESENTED IN THE TOP TWELVE SITES</th>
<th>SPECIES RICHNESS</th>
<th>RESTRICT-RANGE</th>
<th>THREATENED SPECIES</th>
<th>DISPERSION-WEIGHTED</th>
<th>HIGHER-TAXON</th>
<th>SPAN-SUBTREE</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPECIES RICHNESS</td>
<td></td>
<td>53.5</td>
<td>60.6</td>
<td>45.5</td>
<td>47.5</td>
<td>47.5</td>
<td>49.5</td>
</tr>
<tr>
<td>RESTRICT-RANGE</td>
<td></td>
<td>20.7</td>
<td>57.7</td>
<td>27.9</td>
<td>19.4</td>
<td>17.3</td>
<td>19.7</td>
</tr>
<tr>
<td>THREATENED SPECIES</td>
<td></td>
<td>19.5</td>
<td>46.3</td>
<td>48.8</td>
<td>22.0</td>
<td>14.6</td>
<td>24.4</td>
</tr>
<tr>
<td>DISPERSION-WEIGHTED</td>
<td></td>
<td>53.4</td>
<td>60.2</td>
<td>43.7</td>
<td>48.2</td>
<td>46.0</td>
<td>50.1</td>
</tr>
<tr>
<td>HIGHER-TAXON</td>
<td></td>
<td>58.4</td>
<td>58.6</td>
<td>49.5</td>
<td>48.6</td>
<td>54.5</td>
<td>51.5</td>
</tr>
<tr>
<td>SPAN-SUBTREE</td>
<td></td>
<td>66.4</td>
<td>72.6</td>
<td>57.5</td>
<td>61.6</td>
<td>61.6</td>
<td>63.0</td>
</tr>
</tbody>
</table>

As expected, dispersion weighted and spanning subtree selections were very similar in terms of the amount of diversity that they represented (Figure 2.4). The greatest discrepancies were between higher taxon diversity and the two tree spanning measures, but all three taxonomic measures were similar in their ability to represent alternative diversity criteria.
Figure 2.4. Ability of alternative site selection criteria (x axis) at representing different biodiversity attributes: (a) species richness; (b) restricted-range richness; (c) threatened species richness; (d) dispersion-weighted taxonomic diversity; (e) higher-taxon weighted diversity & (f) spanning-subtree weighted taxonomic diversity.

(a) Species richness

(b) Restricted-range species

(c) Threatened species

(d) Dispersion-weighted diversity

(e) Higher-taxon weighted diversity

(f) Spanning-subtree weighted diversity
The top six sites for threatened species were poor at representing all aspects of species diversity (richness and taxonomic diversity) and also at incorporating restricted-range diversity (Figure 2.4). Top sites for threatened species represented less than half the number of species in other site networks (Table 2.2, Figure 2.4).

When the number of sites in the network was increased to twelve, similar patterns in the representation of diversity emerged (Figure 2.4, Table 2.2). Some site networks incorporated as little as a third of the amount of diversity represented by other networks. All species richness and taxonomic measures were poor at representing restricted-range diversity (less than 21% compared to 58% represented by the range size criterion). Likewise both unweighted and taxonomically weighted species richness proved to be ineffective at representing threatened species (less than half the number represented compared to the threat selection) (Table 2.2).

There was little difference between the richness and taxonomic measures in terms of overall species richness represented. The use of unweighted species richness did not significantly reduce the amount of taxonomic diversity that was represented compared to using a taxonomic diversity measure, indicating that for this particular data set the faunas of the most species rich sites were reasonably taxonomically diverse relative to the data set as a whole. In general the amount of taxonomic diversity represented in a network was dependent on the total number of species represented. A network that represents the most species will also represent a large amount of the potential taxonomic diversity in this data set.

Restricted-range richness proved to be good at representing species richness, taxonomic diversity and also threatened species (Figure 2.4). The mean number of sites occupied by threatened species was 3.51 (S.E. 0.492) compared to a mean of 16.6 (S.E. 2.18) sites occupied by non-threatened species. Since all threatened species occurred in less than 14 sites they would have contributed more to the restricted-range diversity score than the average non-threatened species and this accounts for the good performance of the restricted-range network at representing threatened species.
After twelve sites had been selected for threatened species the amount of restricted-range diversity represented improved, with more restricted-range diversity represented compared to the species richness and taxonomic diversity networks (although still less than half the diversity represented by the restricted-range selection). The relative ineffectiveness of the threat selection to represent restricted-range diversity (despite restricted-range being good at representing threatened species) was due to the difference in total species richness represented by the two criteria. The threat algorithm was poor at representing absolute numbers of species (fifteen species fewer than the restricted-range network), and since the restricted-range score also took account of species richness this partly accounts for the surprisingly low amount of restricted-range diversity represented by the threat network (particularly when only six sites were selected).

**Are the same areas selected by the different criteria?**

Agreement in the areas selected was high between unweighted species richness and the taxonomic diversity measures (between four/six and five/six (and eight/twelve - nine/twelve) sites in common) (Table 2.3).

<table>
<thead>
<tr>
<th>Site</th>
<th>Species</th>
<th>Restricted-Rank</th>
<th>Threat</th>
<th>Taxonomic Diversity Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Species Richness</td>
</tr>
<tr>
<td>1</td>
<td>Serengeti</td>
<td>De Hoop</td>
<td>O-Rime-O-Achim</td>
<td>Kafue</td>
</tr>
<tr>
<td>2</td>
<td>Kafue</td>
<td>O-Rime-O-Achim</td>
<td>Haut Dodo</td>
<td>Kafue</td>
</tr>
<tr>
<td>3</td>
<td>Selous</td>
<td>Meru</td>
<td>Cavally-Gouin</td>
<td>Kafue</td>
</tr>
<tr>
<td>4</td>
<td>Boma</td>
<td>West Pretorious</td>
<td>Air &amp; Tenere</td>
<td>Kafue</td>
</tr>
<tr>
<td>5</td>
<td>Gorongosa</td>
<td>Boni Dodori</td>
<td>Krah-Bassa</td>
<td>Kafue</td>
</tr>
<tr>
<td>6</td>
<td>Luangwa</td>
<td>Serengeti</td>
<td>Sapo</td>
<td>Kafue</td>
</tr>
<tr>
<td>7</td>
<td>Tsavo</td>
<td>Yangudi Rassa</td>
<td>Meru</td>
<td>Kafue</td>
</tr>
<tr>
<td>8</td>
<td>Kruger</td>
<td>Southern</td>
<td>Yangudi Rassa</td>
<td>Kafue</td>
</tr>
<tr>
<td>9</td>
<td>Zakouma</td>
<td>Gamba</td>
<td>Grebo</td>
<td>Kafue</td>
</tr>
<tr>
<td>10</td>
<td>Manova</td>
<td>Kangandala</td>
<td>Kafue</td>
<td>Kafue</td>
</tr>
<tr>
<td>11</td>
<td>Ruaha</td>
<td>Luangwa</td>
<td>Bale Mts.</td>
<td>Kafue</td>
</tr>
<tr>
<td>12</td>
<td>Odzala</td>
<td>Bale Mts.</td>
<td>Kangandala</td>
<td>Kafue</td>
</tr>
</tbody>
</table>

The top sites for higher taxon richness contained two unique sites and spanning subtree diversity, one unique site. Neither the species richness nor the dispersion weighted.
diversity selections contained any unique sites. Only one of the first six sites selected using restricted-range richness featured in the top twelve sites for unweighted species richness or any of the taxonomic weightings. Overall there were only two sites common to the species richness and restricted-range selections of twelve sites (Table 2.3). The threat network shared none of its first six sites with the species richness or taxonomic diversity selections and only one site with the restricted-range network. Overall there was only one site common to the threat and richness selections (out of twelve) and five sites common to the threat and restricted-range networks. Serengeti featured in the top six sites for five of the six alternative networks, partly due to it being the most species rich site (equal with Kafue). Kafue was in five of the site selections and Selous, Kruger, Manover and Gorongosa were all selected four times (Table 2.3).

**KwaZulu Natal.**

*Consequences of using different selection criteria for the representation of biodiversity.*

As with the antelope data there were some substantial differences in the amount of diversity represented in the various networks.

**Birds.**

When only four sites were selected a number of the site networks were inferior to the average random selection of sites at representing certain components of diversity (Figure 2.5). The site network selected on the basis of endemic bird richness incorporated fewer species and fewer threatened species compared to the average random selection of sites. Species richness, threatened species and restricted-range networks all represented fewer endemic birds than the average of the random selections. Threat and species richness were particularly poor in this respect, incorporating fewer than half the number of endemic species compared to the random selections. The restricted-range network was overall the best at representing all components of diversity. The top four sites for restricted-range birds incorporated more bird species overall (77 more species than were in the top four most species rich sites) and more threatened species. The restricted-range selection criterion was also an improvement over simple species richness in terms of representing endemic species. Unlike the antelope data, species richness proved a
Figure 2.5. Ability of alternative site selection criteria (x axis) at representing different biodiversity attributes: (a) species richness; (b) threatened species (c) endemic species & (d) restricted-range species, for birds in KwaZulu Natal.
reasonable site selection criterion for representing threatened species and *vice versa*.

The top eight sites for species richness represented on average 73 species more than the mean of the random site selections (Figure 2.5). Threat and endemism selections were less effective at representing simple species richness compared to the top sites for species richness. The network selected using threat incorporated only 36 species more than the mean of the random selections and the endemism selection represented 83 species less than the mean random selection and 156 species fewer than the richness network. Richness of endemic birds was therefore ineffective at representing overall bird species richness. Restricted-range proved to be a good criterion for representing species richness. As with the top 5 percent of sites selected to represent antelope restricted-range diversity, the top eight sites for restricted-range bird species incorporated more species than the top sites for bird species richness.

Species richness, threat and restricted-range selections all represented a large proportion of threatened species (between 77 and 83% of the total threatened species in the top eight sites). In contrast the endemism selection was poor at incorporating threatened species, representing only 38% of threatened species (14 species fewer than the mean random selection) (Figure 2.5).

Both restricted-range and endemism selections incorporated more than 50 of the 68 endemic species (more than 74%) in KwaZulu Natal - over 20 endemic species more than the threat selection. Species richness was intermediate with 40 (59%) endemic species occurring in the top eight sites, although this was less than the number of endemic species represented by the mean random network. The threat network only incorporated 31 of the endemic birds (46%). In comparison a mean of 48 of the endemic species (71%) were represented by the random selections (Figure 2.5).

Restricted-range species were not well represented in any of the networks except the one selected using restricted-range diversity as the criterion. The endemism network was particularly poor at representing restricted-species, incorporating only 11 (20%) of those
species found in less than five sites. However, all selected networks were an improvement on the random selections.

Mammals
There was little difference between the richness, threat and restricted-range selections in their ability to represent species richness or threatened mammals (Figure 2.6). Ninety percent of threatened mammals occurred in top sites for species richness and 76% in top sites for restricted-range diversity. The greatest difference between the alternative selections was in their ability to represent restricted-range species. Both species richness and threatened species networks picked up less than half the restricted-range species compared to the restricted-range network that represented six out of seven of these species in the top eight sites.

Plants.
All site networks were a substantial improvement over the mean random selection at representing species richness (Figure 2.7). The top eight sites for species richness incorporated 63% of total plant species richness, top sites for restricted-range diversity 60%, and top sites for threatened species 53% of all species. Randomly selected sites represented on average less than half the species compared to other selections (and only an average of 25% of all species). Unlike birds and mammals, threatened species were not well represented in top sites for species richness or restricted-range diversity, although both were a substantial improvement on the mean random selection. After eight sites had been selected only 39% of threatened species were represented by the richness network and 41% by the restricted-range network compared to 60% by the threatened species selection. The mean random network only incorporated 17% of the threatened species. The top four sites for richness and the top four sites for threatened species were not good at representing restricted-range species (although were better than the mean random selection), but when eight sites were selected the richness selection improved. The eight site richness network incorporated 22% of restricted-range species compared to 24% in the restricted-range network, but only 12% of restricted-range species were represented by the threat network.
Figure 2.6. Ability of alternative site selection criteria (x axis) at representing different biodiversity attributes: (a) species richness; (b) threatened species & (c) restricted-range species, for mammals in KwaZulu Natal.

(a) Species richness

(b) Threatened species

(c) Restricted-range species

Legend
- Four sites
- Eight sites
Figure 2.7. Ability of alternative site selection criteria (x axis) at representing different biodiversity attributes: (a) species richness; (b) threatened species & (c) restricted-range species, for plants in KwaZulu Natal.
Figure 2.8. Ability of alternative site selection criteria (x axis) at representing different biodiversity attributes: (a) species richness; (b) endemic species & (c) restricted-range species, for grasses in KwaZulu Natal.
Poaceae.
Top sites for endemic grasses were poor surrogates for species richness, but were a slight improvement over the mean random selection of sites (Figure 2.8). All site selections represented a similar number of the endemic species - more than twice the number represented by the mean random selection (eight sites). Endemism and richness networks were not as good as the restricted-range network at representing restricted-range species, but again were considerably better than the mean of the random networks.

Are the same areas selected by the different criteria?

Birds
Overlap was greatest between the species richness and threat selections (six/eight sites in common) (Figure 2.9). Top sites for bird species richness and for threatened bird richness were concentrated in the north east of the region. None of the top sites for endemic species was a top ranking site for threatened species or simple species richness and there was only one site common to the endemic and restricted-range selections. Top sites for endemic bird richness were clustered in the west of the region, in the Drakensberg Mountains. This explains why the restricted-range selection was better than the richness and threat selections at representing endemic birds. Top sites for restricted-range species included three of the top sites for threatened species and four of those for species richness.

Mammals.
Six of the top eight sites ranked for species richness were also in the top sites for threatened mammals (Figure 2.10). Top sites for species richness and threatened species were predominantly in the north east and north bushveld regions. Three sites were common to the richness and restricted-range selections and two sites to the threat and restricted-range top sites. The top ranking sites for restricted-range species were more dispersed, but included several western sites in the Drakensberg Mountains.

Plants.
Only two of the top sites for plant species richness and restricted-range diversity were
Figure 2.9. Location of the top eight sites selected by the alternative diversity criteria for birds in KwaZulu Natal.

Species richness

Threatened species richness

Endemic species richness

Restricted-range species richness
Figure 2.10. Location of the top eight sites selected by the alternative diversity criteria for mammals in KwaZulu Natal.

Species richness

Threatened species richness

Restricted-range species richness
Figure 2.11. Location of the top eight sites selected by the alternative diversity criteria for plants in KwaZulu Natal.
Figure 2.12. Location of the top eight sites selected by the alternative diversity criteria for grasses in KwaZulu Natal.
also in the top eight sites ranked for threatened plants (Figure 2.11). Five of the top eight sites for species richness were also top sites for restricted-range species. Top sites for all aspects of plant diversity tended to be more dispersed than for birds and mammals. The top sites for plant species richness included sites on the south east coast, the Drakensberg Mountains and the north east bushveld. Sites ranked highly for restricted-range species were concentrated in the west, in the Drakensberg, whereas threatened plant richness was generally highest in the north east and along the south east coast of KwaZulu Natal.

Poaceae.

Three of the top sites for species richness were in the top sites for endemic species. These sites were all in the western part of KwaZulu Natal (Figure 2.12). Three sites were common to the richness and restricted-range selections and the endemism and restricted-range selections. Two of the western sites (in the Drakensberg Mountains) were in all three site networks. Sites of high endemic species richness were almost exclusively in the Drakensberg Mountains. Restricted-range species richness was highest in the north east and also in the Drakensberg. Sites with the greatest number of species were not confined to any particular area within KwaZulu Natal - top sites for species richness occurred in the Drakensberg, on the south east coast and in the north east of the region.

DISCUSSION.

One approach to the conservation of biodiversity is to prevent species from becoming extinct. As Diamond (1976) wrote, it is species that would be ‘doomed to extinction in the absence of refuges’ that should be the priority for conservation. Indeed species that are not threatened or potentially threatened do not need protection (Ratcliffe 1986). Short term losses could be reduced by selecting priority areas that represent species currently on threatened species lists. Here the assumption is that species not currently threatened with extinction either co-occur with threatened species (so will be protected), or will persist unaided in the future (Williams 1993).
When only a limited number of sites (six sites) could be chosen to conserve Afrotropical antelopes, selecting areas to incorporate threatened species was not efficient at representing overall numbers of species or taxonomic diversity. So in this example, protecting species at risk of extinction in the short term could compromise future species richness and taxonomic diversity. However when the top twelve sites (5 percent of all sites) for threatened species were included in the site network the percentage of species and taxonomic diversity represented did improve substantially. The top eight sites (5 percent of sites) for threatened birds in KwaZulu Natal represented 82% of the total bird species richness, which, although lower than the amount represented by the richness network, was an improvement over a random selection. Even when only four sites were selected the threat selection was still good at representing overall species richness. However, the top sites for threatened birds were poor at representing endemic and restricted-range bird species. The apparent success of a threatened species approach at representing bird species richness in KwaZulu Natal may have been partly due to the inclusion of species that were threatened at a regional, but not global level. The plant threatened species network (eight sites) incorporated almost 600 species less than the richness network, but again this was an improvement over a random selection of sites.

Lack of concordance between sites with high species richness and those with many threatened species may not be a problem unless more species become threatened. In this case a focus on currently threatened species could restrict the representation of diversity in the future if, for example, the currently threatened species (and therefore protected species) are taxonomically clumped or occur in species poor areas (as with the antelope in this analysis). Another problem with a threatened species approach is that for many taxa there is little or no information on their status. For example only 124 of the 6111 plant species recorded from KwaZulu Natal were listed as threatened. This almost certainly is an underestimation and reflects the lack of information about the status of many plant species in KwaZulu Natal (Hall et al. 1984). A threatened species approach could be valid in cases where few species are expected to become threatened in the future, or if other aspects of diversity are also conserved by such a method. Using threatened species as a focus for the selection of priority areas for conservation also
requires there to be threatened species present in the region under evaluation. This may not be the case for all regions and for all scales at which priority area evaluation is applied.

An alternative approach is to select areas that would collectively represent the full range of existing biodiversity, where biodiversity might be measured as taxonomic diversity, or simply species richness, for example. Here the assumption is that over the long term most species will become threatened so the best approach to conservation is to set aside areas that protect as wide a range of biotic diversity as possible. If species threatened with extinction in the short term occur in the sites selected for species diversity, then short term losses will not be too great.

However, when sites were selected on the basis of species richness for Afrotropical antelope, the sites initially selected contained few threatened species. Only half the amount of threatened species richness incorporated in the threat network was represented by the top six sites for antelope species richness, and by the time twelve sites had been selected less than half the threatened species were represented in the richness network compared to the threat network. Similarly for plant species in KwaZulu Natal, only 39% of threatened species were represented in the top eight sites for species richness compared to 60% in sites selected to incorporate threatened species, although three of the top sites for plant species richness were in the top eight sites ranked for threatened species. A network of sites designed to maximise taxonomic diversity for Afrotropical antelope was also poor at representing threatened species. Thus, in practice protection would not be given to the threatened species, and their likely short term extinctions would compromise the longer term aim of protecting a representative range of biotic diversity.

However, for birds, carnivores and ungulates in KwaZulu Natal, a network selected on the basis of species richness would be good at representing threatened species. For birds, this may have been in part due to the choice of threatened species included in the analysis. All species designated as threatened at an international level and also in the
south African Red Data Book were included. Some bird species included in the south African list, and not present on an international threat list were species that were at the southern edge of their range in South Africa, for example the Wattle-eyed Flycatcher (*Platysteira peltata*). Several species included in the South African Red Data Book are predominantly tropical species whose ranges only just extend into the north east part of South Africa (Hurford et al. 1996). These species tend to occur in the north east of KwaZulu Natal, which is also the region with the highest bird species richness. So the apparent success of the richness criteria at representing threatened birds in KwaZulu Natal may have been due to the inclusion of a certain category of threatened species in the data set (i.e. those species that are rare in KwaZulu Natal due to being a peripheral population). Some of the globally threatened species, which are arguably a higher conservation priority than regionally threatened species, were restricted to the high altitude west of KwaZulu Natal which has low species richness (e.g. Rockjumper, *Chaetops aurantius*, Yellow-breasted Pipit, *Anthus chloris* and Drakensberg Siskin, *Serinus symonsi*). These species would not be picked up if sites were selected on the basis of absolute species richness.

A compromise might be to use biological traits that correlate well with vulnerability to extinction in the future, for example species with large body sizes, high trophic level, small population sizes, restricted-ranges or habitat specialists (Terborgh 1974, Terborgh & Winter 1980, Thomas & Mallorie 1985, Cowling & Bond 1991, International Council for Bird Preservation (ICBP) 1992). Use of these criteria might also pick up areas that are biologically diverse. Birdlife International (ICBP 1992) has used restricted-range in birds as an indicator of vulnerability to extinction and has identified priority regions for conservation on the basis of the distribution of such species. These areas important for restricted-range birds not only contain many of the bird species most threatened with extinction, but are often also important for other taxa including mammals, plants and insects (ICBP 1992). For example, the number of restricted-range bird species in tropical countries is positively correlated with levels of endemism and species richness in several other animal and plant taxa (Balmford & Long 1995).
For the antelope data the sites were not contiguous so the number of sites occupied by a species did not necessarily indicate that species had a small geographic range. Nonetheless threatened antelopes were found in, on average, fewer sites than non-threatened species. The mean number of sites occupied by threatened species was 3.51 (S.E. 0.492) compared to a mean of 16.6 (S.E. 2.18) sites occupied by non-threatened species. Since all threatened species occurred in less than 14 sites they would have contributed more to the restricted-range diversity score than the average non-threatened species and this accounts for the good performance of the restricted-range network at representing threatened species. The relationship between restricted-range and threat would have been reinforced if range had been a criterion used to assign the categories of threat to species.

Use of restricted-range as a selection criterion was generally better than the threat algorithm in terms of the overall species diversity represented. For the antelope analysis this to some extent results from the fact that the restricted-range weighting used in WORLDMAP includes the species richness of each site in the score, although this was not a factor in the KwaZulu Natal results. The success of the restricted-range criteria used in these analyses may well have been because the networks of sites selected were more likely to contain different suites of species (because restricted-range species occur in fewer sites) and therefore be more complementary, rather than any biological reason per se. However, Thomas & Mallorie (1985) found that restricted-range butterfly species in Morocco tended to occur in the most species rich communities so that conservation of restricted-range species would coincidentally protect most other species. Using richness of species that were restricted-range within KwaZulu Natal as a selection criterion proved to be efficient at representing species richness, and in many cases other diversity components. Birds that occurred in only a few sites within KwaZulu Natal were often species that were at the southern limit of their global range (for example Lagonosticta rhodopareia and Vidua purpurascens that are both on the southern edge of their range in northeastern KwaZulu Natal), or were passage migrants or non-breeding visitors. These species tended to occur in areas of high overall species richness, in the north east and along the east coast of KwaZulu Natal. Other restricted-range species
were threatened and/or endemic species. As a result, top sites for restricted-range species included high diversity sites for species richness (four of the eight top sites for species richness), threatened species (three/eight) and endemic species (one/eight). If the restricted-range criterion had been applied at a biogeographical scale (rather than a political level as was done in this analysis), the number of endemic species represented is likely to have been greater. Restricted-range grasses included 16 species that were endemic to southern Africa (24% of all endemics occurring in KwaZulu Natal). Top sites for restricted-range grasses included three of the top sites for endemic species and three of the top sites for species richness. In fact the restricted-range site network incorporated more species than the richness network and more endemic species than the endemic network. There were five sites common to the restricted-range and richness networks for all plants, and restricted-range was good at representing species richness. It did less well at incorporating threatened species, particularly when only four sites were selected.

Species defined as restricted at a political scale (rather than a biological one) may not represent a high priority for conservation. Although the KwaZulu Natal bird data set did exclude vagrant species, many of the species defined as restricted-range were non-breeding visitors or passage migrants which is why some east coast sites had high restricted-range scores. Endemic species (i.e. biogeographically restricted) were recorded from a mean of 49.6 sites (S.E. 5.24, n=68) in KwaZulu Natal, compared to a mean of 59.8 sites (S.E. 2.22, n=502) for non-endemics. However, since restricted-range species were those defined as occurring in less than five sites, many of the endemics would not be classed as restricted within KwaZulu Natal. Only twelve of the 68 endemic species occurred in less than five sites, but not even all these species were found in the highest ranking sites for restricted-range diversity. Using a restricted-range criterion applied at a biogeographical scale for southern Africa would probably do less well at representing species richness than was the case for KwaZulu Natal.

Species endemic to southern Africa were species with globally restricted-ranges. When endemic species richness was used to select sites rather than richness of species only
occurring in a few sites within the region, there were noticeable differences in the amount of diversity represented in the site network. None of the top sites for endemic birds in KwaZulu Natal was high ranking for bird species richness and a network of sites selected using endemic bird richness was very poor at representing overall bird species richness. The majority of the endemic bird species occurring in KwaZulu Natal are found in the Drakensberg Mountains and uplands in the extreme west of KwaZulu Natal. All the top ranking sites for endemics were in this western region. However these areas were overall poorest for species richness, which was greatest in the north east and along the east coast of KwaZulu Natal. No Drakensberg or western areas featured in the top eight sites for species richness or threatened species. In the case of the threatened species selection this is surprising since endemic species are predicted to be vulnerable and therefore are more likely to be considered under threat. However only 18 of the 68 endemic species were listed as threatened and no western areas ranked in the top eight for threatened species.

A few endemic species were confined to coastal evergreen forest which occurs along the east coast and a number of the endemic species occurring in the Drakensberg winter further east, sometimes as far as the south east coast. This accounts for the large number of endemics represented by the restricted-range selection since this includes a site on the south east coast and also a Drakensberg site. If only breeding records were included the coastal site would not contribute so many endemic species. Several endemic species are confined to the north east of KwaZulu Natal (also a centre of plant endemism) and a number of endemics occur throughout KwaZulu Natal which explain why the species richness selection represents 40 of the endemic species. Twenty-eight species are confined to the west of KwaZulu Natal so will not be represented by the richness selections.

Similarly top sites for endemic grass richness were inferior to other selections in their ability to represent overall richness, although they did include three sites ranked in the top eight for species richness. High scoring sites for endemic grasses were all in the west of KwaZulu Natal (all but one in the Drakensberg). This area has high species richness for grasses (unlike birds), but the north east of KwaZulu Natal is also species rich (like birds). The endemic network did not include any northeastern sites so represented fewer
species overall compared to the other networks. Several top sites for species richness were also in the Drakensberg which explains why the richness network was good at representing endemic grass species.

Species richness proved to be a good surrogate for taxonomic measures of diversity in the antelope data set. This indicates that the species in the different sites were not taxonomically clumped, and was expected given that the data set was reasonably large (Williams & Humphries 1994, Williams 1993). Species richness was therefore a reasonable first approximation to taxonomic diversity although the general applicability of this needs to be tested for other data sets (see also Williams & Humphries 1994). While taxonomic diversity scores were dominated by the number of species in the fauna, they did vary between sites with the same number of species. Taxonomic diversity could therefore be useful in resolving ties between sites, or in selecting alternative sites in situations where the best site is not available for conservation. For instance Waza, Keran, Scio, Mont Sangbe, Meshra and Harrar all contain three antelope species, but their taxonomic diversity scores vary substantially (Figures 2.13a & b). Waza and Keran have the highest dispersion and spanning subtree scores because they represent the greatest number of nodes in the classification and also because the pairwise divergence between species is most even. However Keran has twice the higher taxon richness score of Waza, because *Sylvicapra grimmia* which occurs in Keran is closer to the base of the classification than any of Waza's species. Scio contains three species all from the genus *Cephalophus*. This means that it has low dispersion weighted diversity and spanning subtree diversity scores but has a greater higher taxon score than Waza because *Cephalophus* is closer to the root of the tree than any of the taxa occurring in Waza. Similarly, Mont Sangbe, with two species from the genus *Cephalophus* and one *Tragelaphus* species (the two most basal taxa) has the largest higher taxon score, but a very low spanning subtree and dispersion scores. Meshra and Harrar both contain three species that between them do not span many nodes of the classification and also do not lie close to the base of the tree. Both these sites have low dispersion/spanning subtree scores and also low higher taxon scores.
Figure 2.1a. Cladograms showing where species from three of the anole genera fit in the overall classification scheme. Each site has three species, but different taxonomic diversity scores as calculated by the three WMRDAMP measures.
Figure 2.13b. Cladograms showing where species from three of the antelope sites lie on the overall classification (Meshra, Harrar & Scio). Each site has three species, but different taxonomic diversity scores as calculated by the three WORLDMAP measures.
In reality, it is likely that the resources and land will not be available to conserve every species. If the number of sites that can be selected is limited, then conflicts occur and no one criterion will incorporate all the desired attributes. Thus, in practice strategic decisions will need to be made about the goals for particular sets of reserves. Given that we do not know what the situation will be in the future regarding numbers of threatened species, the best approach might be to select areas that minimise short term losses of species but also maximise components of biotic diversity. Sites could be selected using one diversity component, but adjusting the strategy from time to time to incorporate aspects of diversity not selected by that particular criterion. For example threatened status could be used to make the initial selections, switching the selection criterion to species richness for later selections. Alternatively sites could be given different levels of protection enabling a larger number to have some form of protection but saving outright protection for those already subject to threatening processes.

Generally, it is useful to look at the relationship between the distribution of richness, threatened species and restricted-range/endemic species across sites when choosing a policy for reserve selection. If areas containing threatened species also contain large numbers of other species that are not taxonomically clumped, then selecting areas to conserve threatened species may also incorporate aspects of diversity such as species richness and taxonomic diversity. Problems will arise if threatened species occur in sites of low species diversity because there will be little overlap between the sites selected for species richness and those needed to protect threatened species. For African antelopes there was little overlap between the most important sites for threatened species and those for species richness - only one out of the first twelve sites was common to the two selections and consequently areas selected on the basis of threatened species were relatively poor at representing overall species richness. Similarly, top sites for threatened plants included only three of the eight top sites for plant species richness. In contrast overlap between top sites for threatened birds and species richness in KwaZulu Natal was 75%. Coincidence between sites will improve as the number that could be reserved increases; however, in practice the number of areas that could be reserved is likely to be limited. Overlap was similarly low between sites selected for endemic birds and grasses.
and those selected for species richness and for restricted-range antelope and species richness. Likewise for a number of British taxa Prendergast et al. (1993) found that many of the rare or restricted species did not occur in the most species rich sites, and they concluded that a limited number of species rich areas do not guarantee effective conservation for rare and restricted organisms.

In the selected networks a large number of the species were represented in just one or a few sites, and this might not constitute a viable population. This will apply particularly to those species that are threatened or are otherwise susceptible to extinction. It will therefore be desirable in practice to use a method which aims to achieve multiple representations of species wherever possible, especially those species that are threatened. This will inevitably increase the disparity between the site networks selected using different diversity criteria.

The kind of analysis described here is important for determining the outcome of different selection criteria in terms of which aspects of diversity will be represented or underrepresented. If sites that are good for one component of biodiversity (e.g. species richness) also represent a wide range of other diversity components (e.g. threatened species or taxonomic) then the question of which aspects of diversity should be given priority when selecting conservation areas would not be a problem. For a number of the taxa in this example a selection programme based solely on species richness or threatened species leads to tradeoffs in the other aspects of diversity represented. The best approach might be to combine the different selection strategies to determine the amount of different aspects of diversity conserved at each step and adjusting the subsequent selection strategy accordingly. The results presented here illustrate that it is necessary to be explicit about the aim of conservation area networks, since this may have a significant influence on which sites are chosen and which attributes will be represented, especially if the designation of priority areas is seen as a 'carte blanche' for the destruction of other areas (Rylands 1991).
CONCLUSIONS.

- Restricted-range species richness (defined in the context of the scale of evaluation) was the best criterion for including a range of different biodiversity attributes.

- In contrast, areas rich in 'true' endemics were not the most species rich areas and sites rich in endemic species were poor at representing overall species richness and *vice versa*.

- A focus on threatened species (especially globally threatened ones) may compromise the representation of other aspects of diversity (e.g. species richness, taxonomic diversity and endemic species), although this is dependent on the taxon.

- Use of a simple species richness measure did not introduce biases in the representation of taxonomic diversity, but may not be a fine enough filter to pick up threatened, restricted-range or endemic species.
CHAPTER THREE.

QUANTIFYING PATTERNS OF RICHNESS AND SPATIAL TURNOVER: WHAT MAKES A GOOD SURROGATE MEASURE FOR WHOLESALE BIODIVERSITY?

INTRODUCTION.

The question of which attributes of biodiversity to conserve is compounded by another problem - that we only know about a fraction of the biotic diversity that exists. The ratio of known to unknown species has been variously estimated to range between 1:1.67 and 1:21 (Hunter 1991, May 1992). Any area evaluation will necessarily be based on a subset of the biotic variation in a region. This is inevitably biased towards conspicuous, well-studied groups like birds and large mammals (Murphy & Wilcox 1986, Sutton & Collins 1991). Systematic evaluation ideally requires a surrogate measure that will accurately reflect the distribution of diversity in a region across all taxa and different aspects of biodiversity (wholesale biodiversity; Williams & Gaston 1994).

A number of different surrogate measures have been suggested, including higher-taxon richness (Gaston & Williams 1993, Williams et al. 1994, Williams & Gaston 1994), community, habitat or landscape diversity or representativeness (Richerson & Lum 1980, Thomas & Mallorie 1985, Austin & Margules 1986, Noss 1987, Scott et al. 1987, Margules & Stein 1989, Pressey & Nicholls 1989b, Austin 1991a, Faith & Walker 1996), keystone, mobile link or umbrella taxa (Gilbert 1980, Pearson & Cassola 1992, Ryti 1992) and restricted-range (or otherwise rare species) or endemic species (Terborgh & Winter 1983, Game & Peterken 1984, Thomas & Mallorie 1985, ICBP 1992, Kremen 1994, Balmford & Long 1995). However the validity of most surrogate measures is rarely established and more often species or groups of species are used opportunistically without consideration of their ability to reflect patterns of diversity in other groups.

In this chapter I examine whether there are any general rules to determine the identity of taxa that are good correlates of overall diversity and could therefore be used to select
sites that will collectively represent a large proportion of the regional biodiversity. Species distribution data for a variety of taxa (birds, mammals and plants) in KwaZulu Natal, southern Africa were used to investigate patterns of spatial turnover and richness for different biodiversity attributes. This in turn was related to the ability of areas to represent a range of the regional biodiversity and to identify potential surrogate measures for wholesale biodiversity (Williams & Gaston 1994).

Specifically the following questions were addressed:

1. Do the different taxa show similar gradients of species richness in KwaZulu Natal?
2. Are patterns of spatial turnover the same for each taxon, and what are the environmental correlates of the richness and complementarity patterns?
3. Do the different taxa show similar levels of spatial turnover?

The answers to these questions were used to predict:

1. Groups of species that should be good surrogates for diversity across a range of taxa.
2. How well subsets of sites would represent regional diversity.

**Background to describing the distribution of diversity in a region.**

The biotic diversity of any region is not evenly distributed. The success of any surrogate measure applied across a range of taxa will depend on the extent to which different taxa show similar patterns of species richness and spatial turnover in a region and also the amount of spatial turnover in species composition between sites. Turnover can mean a number of different things, but is used here as a measure of spatial complementarity between different sites in terms of the attributes that those sites contain. In this context, spatial turnover refers to the extent to which different sites within a region contain different elements of biodiversity. Demonstration that high diversity areas coincide for different taxa is not sufficient to conclude that a set of sites selected using one taxon will
also represent the diversity of other taxa. The other important factor is the extent to which the species composition of different sites varies - i.e. the degree of spatial turnover in species between sites. A successful surrogate measure needs to reflect patterns of richness and spatial turnover across a wide range of taxa and diversity components. In this chapter I investigate concordance in patterns of richness across taxa, measure the amount of species turnover for different taxa and use these factors to derive methods for developing good surrogate measures.

Richness and complementarity (spatial turnover) are in some ways similar to the concepts of alpha diversity (richness) and beta and gamma diversity (complementarity) described by Whittaker (1960, 1972, 1977) and Cody (1975, 1986) but at a coarser scale. Alpha diversity describes within habitat or point diversity and so is similar to the within site richness used here (although individual sites may contain several habitats and so there may be spatial turnover within the site), beta diversity describes the extent to which communities of species change in composition along environmental or habitat gradients and gamma diversity is turnover across geographic areas of similar habitat. Together beta and gamma diversity describe spatial turnover between communities of species in different habitats or areas.

Consider the following distributions of two taxa, illustrated in Figure 3.1(a-e):

(i) Equal site richness, no spatial turnover for either taxon (Figure 3.1a).
All sites contain 100% of the regional species richness for both taxa.

(ii) Equal site richness, different level of spatial turnover for the two taxa (Figure 3.1b).
Any one site contains 100% of the regional species richness for taxon 1 but only 33% of the species in taxon 2.

(iii) Same richness gradient for both taxa, no spatial turnover (Figure 3.1c).
There is a gradient of species richness from site 1 (high) to site 3 (low) for both taxa. With no spatial turnover for either taxon, selecting the most species rich site for either taxon would result in the representation of all species in the region.
Figure 3.1. Five hypothetical groups of sites (a - e) showing different levels of species turnover and richness for two taxa. See text for explanation.

(a) SITE 1

SITE 2

SITE 3

TAXON 1 = SPECIES A, B & C
TAXON 2 = SPECIES a, b & c

(b) SITE 1

SITE 2

SITE 3

TAXON 1 = SPECIES A, B & C
TAXON 2 = SPECIES a, b, c, d, e, f, g, h & i

(c) SITE 1

SITE 2

SITE 3

TAXON 1 = SPECIES A, B & C
TAXON 2 = SPECIES a, b & c

(d) SITE 1

SITE 2

SITE 3

TAXON 1 = SPECIES A, B, C, D, E & F
TAXON 2 = SPECIES a, b & c

(e) SITE 1

SITE 2

SITE 3

TAXON 1 = SPECIES A, B, C, D, E & F
TAXON 2 = SPECIES a, b, c, d, c & f

MAXIMUM COMPLEMENTARITY FOR TAXON 2

MAXIMUM COMPLEMENTARITY FOR TAXON 1

RICHNESS GRADIENT FOR BOTH TAXA

91
(iv) *Same patterns of richness, different levels of spatial turnover for each taxon* (Figure 3.1d).

For both taxa there is a richness gradient which runs from site 1 (high richness) to site 3 (low richness). The richest site for both taxa is site 1. However if site 1 were selected as a reserve it would incorporate 100% of taxon 1 but only 50% of taxon 2. This is because the two taxa show different levels of spatial turnover in the region.

(v) *Same pattern of richness, same level of spatial turnover, different pattern of spatial turnover for the two taxa* (Figure 3.1e).

Both taxa show a gradient of species richness running from left (high richness) to right (low richness). However the pattern of spatial turnover is opposite - taxon 1 shows maximum turnover (complementarity) running left to right (the same as the richness gradient) whereas taxon 2 shows maximum spatial turnover running up and down the region. The top three sites for the two taxa are the same since they show exactly the same patterns of richness, however these three sites (1, 4 and 7) represent 50% of taxon 1 but 100% of taxon 2. The three richest sites do not pick up all the species of taxon 1 because the sites do not maximise complementarity. In contrast three sites that maximised complementarity (e.g. 1, 2 & 3) would represent 100% of taxon 2 even though the mean site richness is lower than for sites 1, 4 and 7.

It is therefore not sufficient to only consider patterns of species richness when selecting areas to represent the diversity of a region. The level and pattern of spatial turnover of species within the region will also be important in determining the optimal configuration of sites. For taxa showing little spatial turnover in a region, any network that includes just one high scoring site will encompass the majority of the regional diversity for the taxon. High turnover groups, in contrast will require a suite of sites that encompass the range of compositional variation in the region.

This has the following implications for the representation of diversity in a region:

(i) If patterns of richness and complementarity for two taxa are the same then the same set of sites will maximise the amount of diversity conserved for each taxon. This will not
necessarily be the most species rich sites if the complementarity gradient runs against the richness gradient. If the levels of complementarity are different for the two taxa (even though the pattern is the same) then the best sites for representing regional diversity for the two taxa will be the same still, but a lower proportion of the species in the taxon with higher turnover will be represented. Put another way more sites will be needed to represent the same proportion of diversity for the higher turnover taxon.

(ii) If patterns of richness are the same for two taxa, but patterns of complementarity are not the same, then the set of sites that maximises diversity for one taxon will not be the same as the set of sites that maximises representation of the other taxon.

(iii) If complementarity between sites is generally low then the top sites for species richness will probably come close to the optimum number of species that can be represented in a subset of sites. This will particularly be the case when there is a large difference in the number of species that occur in different sites i.e. between species rich and species poor sites. If the difference between species rich and species poor sites is very large and differences in the complementarity of different sets of sites within the region is relatively small then it will again be best to select species rich sites even though these may not maximise complementarity.

Turnover and richness patterns will also affect the ability of a set of sites to represent different components of diversity such as threatened or endemic species. If sites that are good for one component of biodiversity (e.g. species richness) also represent a wide range of other diversity components (e.g. number of threatened species or taxonomic diversity) then the question of which aspects of diversity should be given priority when selecting conservation areas would not be a problem. This will again be determined to some extent by the amount of turnover shown by species in the region. If there is no turnover, then species poor sites will contain an exact subset of the species found in more diverse sites (rather than a unique suite of species) and so will also contain a subset of certain other aspects of the regional biodiversity, for example threatened species. So for taxa with little or no turnover, selecting the most species rich sites may not compromise representation of some of the other aspects of diversity. This may not be the case for high turnover taxa however. Species poor sites will contain species that are not found
in the most diverse sites so that conserving the most species rich sites will not necessarily guarantee representation of other aspects of diversity. For instance threatened species may occur in the most species poor sites so would not be conserved if only the most species rich sites were selected.

Factors affecting species turnover.
Different taxa may show different patterns of spatial turnover between sites for a variety of reasons. For example small-bodied taxa will respond to the environment at a finer scale compared to larger-bodied taxa (due to fractal dimensions). Invertebrate taxa are likely to exhibit species turnover at a finer scale than large vertebrates since an individual plant, for example, can provide a habitat for an invertebrate. Other factors that could influence the level of species turnover in a region include the degree of mobility and dispersal distances of individual species, the presence of geographical barriers to movement, the degree of habitat specificity of species, and the spatial heterogeneity of habitats within the region. Highly mobile taxa (for example birds) are predicted to show lower species turnover compared to less mobile taxa. Taxa that are very habitat specific will exhibit higher turnover compared to habitat generalists, especially if their specific habitat is restricted within the region. Geographical barriers to movement or dispersal, for example mountain ranges or rivers, may also promote high species turnover within a region. If different taxa do show similar patterns in the distribution of diversity components (e.g. if species richness patterns for birds mirrored those of plants, or if endemic birds occurred in areas important for endemic mammals) then it might be possible to use some subset of taxa as a surrogate or indicator of the overall importance of areas for a wide variety of taxa.

METHODS.

Data sources.
Species distribution data.
Distribution data for birds, mammals and plants were the same as that used in the previous chapter. Bird data consisted of atlas records for 570 regularly recorded bird species (Cyrus & Robson 1980). Mammal distribution data came from an atlas of
ungulates and an atlas of carnivores (Rowe-Rowe 1992a, 1992b, 1994) which consisted of all available species records obtained from a variety of sources (e.g. distribution records from Natal Parks Board staff working in protected areas, personal observations from private land and museum material (post 1977)). Plant distribution data for 6111 species, subspecies and varieties were collated from the computerised herbarium database PRECIS (National Herbarium, Pretoria (PRE) Computerised Information System (CIS)). Subspecies and varieties of plants were treated separately as they are recognised as distinct in the Red Data Book for South African Plants (Hall et al. 1980). However to determine the extent to which the use of different taxonomic rankings might influence the results of this study, some of the analyses were also carried out at the species level with all plant subspecies and varieties treated together at the species level.

Threatened species data for birds, mammals and plants came from South African Red Data Books and International Red Data Books (Hall et al. 1980, Brooke 1984, Collar & Stuart 1985, Smithers 1986, Collar and Andrew 1988, Collar et al. 1994). All species ever listed in either the International or the South African Red Data Books were treated as threatened for the purposes of this analysis, including species listed as Near Threatened, Indeterminate and Uncertain. Field guides and the bird atlas for KwaZulu Natal were used to identify breeding passerine landbird species which could be considered habitat specialists. Although this is a rather arbitrary term, specialist species were defined as those only recorded from one habitat type. The following six habitat types were recognised: high altitude/upland grassland, coastal/riverine evergreen forest (lowland evergreen forest), thornveld (arid savanna), broadleaved woodland (moist savanna), montane evergreen forest/mistbelt (high altitude evergreen forest) and mountainous cliffs/ravines. Species were treated as habitat specific if recorded from only one of the above habitat categories and borderline species were classified as generalists. This gave a total of 68 habitat specific, terrestrial passerines, and 127 generalist species. Only passerine species were used for the habitat specificity analysis due to a limit on the number of species that could be processed by some of the computer software used for the analysis.
Data on endemicity were obtained from fields guides where possible. Information was available for all bird species endemic to the southern African subregion (as defined in the previous chapter). There was less information available for plants. Comprehensive data on endemic grass (Poaceae) species was obtained from the Grasses of Southern Africa (Gibbs-Russell 1990). Since grassland is the major biome present in KwaZulu Natal, and given the paucity of information on the distribution of other plant species, it was decided to restrict the endemic plant analyses to just grass species (Rutherford & Westfall 1994).

**Environmental data.**

Mean annual precipitation and mean altitude data at a 15 minute by 15 minute (quarter degree) scale were obtained from the Computing Centre for Water Research (CCWR) at the University of Natal, Pietermaritzburg. The data are derived from estimates of mean annual precipitation at one minute by one minute of a degree intervals. The estimate of mean annual precipitation (MAP) at each grid point was determined using an adjusted regression surface in which MAP (measured at various meteorological stations) was regressed against factors such as altitude, latitude, longitude, continentality and aspect. Temperature data were also provided by the University of Natal. Data consisted of mean monthly maximum, mean monthly minimum temperature estimates and mean monthly mean temperature estimates calculated from the maximum and minimum estimates. Temperature data were available for each minute of latitude and longitude in KwaZulu Natal. This provided 225 temperature readings for each complete quarter degree square for each temperature variable measured. Coastal grid squares were often not complete and therefore had less than 225 one minute data readings. Complete data (225 readings) were used for quarter degree squares that crossed political boundaries even if the square was not entirely in KwaZulu Natal. These data were used to extract a variety of different temperature statistics for each quarter degree grid cell. A number of these statistics were highly correlated with one another and were subsequently reduced to those that were most important in explaining the variation in species distribution patterns in the following analyses, namely:

1. mean annual temperature
2. mean minimum temperature of coldest month (July)
Coarse habitat data for each quarter degree square were taken from a map of the bioclimatic regions of Natal in Rowe-Rowe (1992a). This was derived from the eleven bioclimatic regions of Phillips (1973) which were based on altitude, rainfall, temperature and vegetation characteristics. Four of the original regions are grouped into ‘bushveld’, giving a total of eight bioclimatic regions within KwaZulu Natal (Rowe-Rowe 1992a). These are coast lowlands (evergreen grassland, and tropical forest and thicket); coast hinterland (grassland and semi-deciduous woody vegetation); bushveld regions of the north-east interior, and lower reaches of major river valleys (semi-deciduous thicket and scrub, dominated mainly by *Acacia* spp.); mistbelt of the midlands (grassveld and Afromontane forest); moist upland (tall grassveld and open savanna); drier upland (tall grassveld and open savanna, drier than previous region); highland (grassland with short, dense cover and patches of Afromontane forest); and montane (temperate grassland and fynbos) (Rowe-Rowe 1992a) (Figure 3.2).

For each quarter degree grid square, the bioclimatic type that covered the largest area of the square was designated as the dominant habitat. If more than one habitat type was equally dominant (for example if two habitats each covered 50% of the grid square) then these habitats were all listed as dominant for that particular square. All other habitat types present were also recorded for each grid square.

*Resolution of data.*

All distribution data were analysed at a quarter degree grid square resolution (approximately 24 x 27.5 km in South Africa, Osborne & Tigar 1992) which gives squares of 660 km². KwaZulu Natal is covered by 166 of these squares. These quarter degree squares were used as the units for site selection in the following analyses.
Figure 3.2. Map of KwaZulu Natal showing the different habitat types, redrawn from Rowe-Rowe (1992a).
Analysis.

Correlations between site level richness for different taxa and diversity components.

To test for interdependence between site level richness for the different groups, Pearson's product-moment correlation coefficient was calculated for the following site richness variables (see data matrix below). Where one of the groups was a subset of the other group in the correlation (e.g. endemic birds against all birds) then the subgroup was excluded from the main group (i.e. endemic birds were correlated against richness of non-endemic birds). Additionally the number of sites that were common to the top eight and sixteen sites scored for each of the above was recorded.

<table>
<thead>
<tr>
<th>BIRD SPECIES RICHNESS</th>
<th>MAMMAL SPECIES RICHNESS</th>
<th>PLANT SPECIES RICHNESS</th>
<th>ENDEMIC BIRD SPECIES RICHNESS</th>
<th>ENDEMIC GRASS SPECIES RICHNESS</th>
<th>HABITAT SPECIFIC PASSERINE RICHNESS</th>
<th>BIRD FAMILY RICHNESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIRD SPECIES RICHNESS</td>
<td>-</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>MAMMAL SPECIES RICHNESS</td>
<td>-</td>
<td>-</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>PLANT SPECIES RICHNESS</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Quantifying spatial autocorrelation.

Autocorrelation refers to the phenomenon whereby successive values of residuals along a regression line vary in a systematic way. Spatial autocorrelation occurs when values that are nearby spatially vary in a systematic manner. Spatial autocorrelation can be positive or negative; positive spatial autocorrelation occurs when nearby values of a variable are more similar than more distant values and negative autocorrelation when nearby values are more dissimilar than more distantly distributed values of the variable. Statistically, if there is spatial autocorrelation in data, there will be a degree of redundancy of information about a phenomenon which is related to spatial proximity. In trying to explain patterns in the distribution of a phenomena in terms of causal
processes then it is necessary to be aware of the amount of spatial autocorrelation present in the data and to correct for it when using statistics to test for systematic relationships in the data. For example if two sets of spatial values relating to two phenomena $x$ and $y$ exhibit positive spatial autocorrelation in the data, nearby values of each single variable will be more related to each other than more distant values irrespective of any real correlation between $x$ and $y$. This violates the statistical assumption that observations are pairwise independent (Legendre 1993). Therefore in testing the significance of any correlation between $x$ and $y$ it is necessary to take account of the fact that nearby values cannot be treated as independent, and so the degrees of freedom will be lower than the sample size suggests.

Several tests exist for measuring the amount of spatial autocorrelation between spatially distributed ordinal or interval values e.g. the Geary Ratio (Geary 1954) and the Moran Coefficient (Moran 1950).

The Moran spatial autocorrelation coefficient ($I$).

\[
I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij}(x_i - \bar{x})(x_j - \bar{x})}{J \sum (x-\bar{x})^2}
\]  

(1)

where $n$ is the number of areas in the region; $J$ is the number of joins between adjacent sites; $x$ is the value for an area; $x_i$ and $x_j$ are the values for two contiguous areas ($c_{ij} = 1$ for adjoining areas). This coefficient only measures the similarity of adjacent areas to one other compared to non-adjacent values. Spatial autocorrelation could also exist between non-adjacent squares i.e. it is a continuous phenomenon, so ideally spatial autocorrelation should be quantified with respect to the distance between pairs of sites rather than only on the basis of whether two sites are adjacent or not.

The expected value of $I$ for a random arrangement of data (i.e. no spatial autocorrelation) is small and negative. A more negative number indicates dispersion - that is similar values are not aggregated, while a positive value of $I$ indicates clustering of similar
values.

Testing the significance of the value involves calculating a standard normal deviate from the derived value of $I$, the expected value of $I$ (with no autocorrelation present) and its standard deviation. There are two possible forms of null hypothesis: normality or randomisation. Under an assumption of normality the null hypothesis is that the observed values of $x$ are the result of taking $n$ values at random from a normally distributed population of values. If randomisation is assumed the $n$ values of $x$ are taken as given, but the particular spatial arrangement of these values is considered in terms of all the possible ways in which those values could be arranged within the region. So, given the set of values for $x$, what is the probability that their spatial arrangement could have occurred by chance?

The equation for the expected value of $I$ under the null hypothesis of normality is:

$$E_i = -\frac{1}{n - 1}$$

and the equation for the standard deviation of $E_i$ is:

$$\sigma_i = \frac{n^2J + 3J^2 - n\sum L^2}{J^2(n^2 - 1)}$$

where $n$ is the number of areas; $J$ is the number of joins between adjacent areas; $L$ is the number of other areas to which an area is joined.

Assuming randomisation as the null hypothesis the equation for the expected value of $I$ is the same as under normality. The equation for the standard deviation is different however:

$$\sigma_i = \frac{n[J(n^2 + 3n) + 3J^2 - n\sum L^2] - k[J(n^2 - n) + 6J^2 - 2n\sum L^2]}{J^2(n - 1)(n - 2)(n - 3)}$$
where $k$ is the kurtosis of the variable $x$.

\[
kurtosis = \frac{\sum (x - \bar{x})^4}{n\sigma^4}
\]  

(5)

The observed value of $I$ can then be converted into a standard normal deviate:

\[
z = \frac{I - E_i}{\sigma_i}
\]

(6)

If the direction of departure from randomness has been suggested (e.g. the data are expected to be clustered rather than dispersed) then a one-tailed significance test is used. If no prediction concerning the direction of departure from randomness is made then a two-tailed test is used.

**Determining whether spatial autocorrelation exists in the KwaZulu Natal data.**

Moran’s coefficient of spatial autocorrelation was calculated for the bird, plant and mammal distribution data for KwaZulu Natal. Spatial autocorrelation might exist in these data as a result of the physical forcing of environmental variables, from community processes (e.g. competition) or as a result of the data collection method used (e.g. a species might be more likely to be recorded from an area if it has been recorded from an adjacent area because observers might expect it to be found there and hence spend more time looking for it there than in more distant areas) (Legendre 1993).

A computer program that calculates Moran’s coefficient from Ebden (1985) was used to measure the degree of spatial autocorrelation present in the KwaZulu Natal data. The number of areas was 166 and the number of joins 572. These include joins between sites that are diagonally spaced and therefore only share corners. The null hypothesis for all three taxa was that the observed arrangement of values is random. The expected value of $I$ was the same for all three taxa, since it is dependent only on the number of sites in the data set (166) equation (2).
In an attempt to control for the problem of spatial autocorrelation in the data set, correlation coefficients between bird, mammal and plant species richness were additionally calculated using only non-adjacent grid squares. Selection of non-adjacent squares produced four sets of between 38 and 45 squares which did not share any borders (including diagonal borders) with any of the other squares in the data set. As mentioned previously, spatial autocorrelation could still exist between non-adjacent squares. For example, observers might expend more effort trying to record a species in a square that is as far away as possible from other squares where that species has previously been recorded from. So that while excluding adjacent squares from the analysis might remove or reduce some forms of spatial autocorrelation from the data it cannot guarantee to exclude it altogether. Ideally a method is required that can identify the scale at which spatial autocorrelation occurs in a region so that variables can be measured using the appropriate spatial lag in order to eliminate the effect.

**Patterns of diversity and complementarity and their environmental correlates.**

Multivariate analyses were used to identify major patterns of distribution for the different taxa. The relationships between the species distribution patterns and a number of environmental variables were quantified using canonical correspondence analysis (ter Braak 1987-92). This is a direct gradient analysis technique that attempts to explain species distribution patterns by ordination axes that are constrained to be linear combinations of the supplied environmental variables. Canonical correspondence analysis (CCA) was implemented using the FORTRAN program CANOCO 3.10 (ter Braak 1990). CCA was carried out separately for breeding passerine landbirds (195 species), breeding nonpasserine landbirds (154 species), mammals (65 species) and grasses (400 species). Waterbirds were excluded from the avian analyses since data on the availability of water bodies for the grid cells were not available and this was considered an important determinant of distribution for these species. In this context waterbirds were defined as any species dependent on aquatic habitats for any part of the year. The analysis could not be carried out on all plant species since these exceeded the available computing capacity, analysis was therefore restricted to Poaceae (split into two data sets alphabetically, consisting of 195 and 205 species respectively) since grasses represent one
of the largest plant families in KwaZulu Natal and the grassland biome of southern African covers much of the province (Rutherford & Westfall 1994).

CCA assumes that the relationship between a species’ distribution and the environmental gradients is unimodal, i.e. each species has an optimal location in the environment space. Each species score is estimated by the technique of weighted-averaging. For all taxa, four canonical axes were initially extracted. The eigenvalue for each ordination axis is a measure of the separation of the species distributions along the axis. Eigenvalues can vary from zero to one, the higher the value the more important the axis. The variance of the species data that is explained by each ordination axis is derived by dividing the eigenvalue for the axis by the sum of all unconstrained eigenvalues (i.e. independent of the environmental data). These values are often low for presence-absence data but the ordination axes are still informative (ter Braak 1987-92). The variance of the species-environment relation for each axis is given by the eigenvalue divided by the sum of all constrained eigenvalues.

To test whether the measured environmental variables were adequate to explain the major variation in species composition an indirect gradient analysis technique - Correspondence Analysis (CA) - was applied to the data. CA like CCA assumes a unimodal species response to major gradients but unlike CCA does not attempt to relate species responses to any external variables. It therefore represents major patterns in species distribution without saying anything about the factors associated with these gradients. If the environmental variables used in CCA are not important in predicting the variation in species composition extracted by CA then there will be a large drop in the amount of variation explained by the axes of CCA compared to CA.

CCA was used to identify the major patterns and differences in the distribution of the different taxonomic groups in relation to environmental variables. Environmental variables that were important in explaining the patterns were identified by looking at the interset correlations of environmental variables with the ordination axes. The number of variables used was reduced to just 22 (listed in Table 3.1) after initially running the
analysis and removing variables with variance inflation factors (VIF) of greater than twenty (as recommended in ter Braak 1987-92). Variables with VIFs of more than twenty are almost perfectly correlated with other variables and so make no unique contribution to the ordination.

Table 3.1. List of the environmental variables included in the CCA analysis for birds, mammals and grasses in KwaZulu Natal.

<table>
<thead>
<tr>
<th>ENVIRONMENTAL VARIABLE</th>
<th>TYPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 HIGHLAND</td>
<td>NOMINAL</td>
</tr>
<tr>
<td>2 DOMINANT HIGHLAND</td>
<td>NOMINAL</td>
</tr>
<tr>
<td>3 MOIST UPLAND</td>
<td>NOMINAL</td>
</tr>
<tr>
<td>4 DOMINANT MOIST UPLAND</td>
<td>NOMINAL</td>
</tr>
<tr>
<td>5 DRIER UPLAND</td>
<td>NOMINAL</td>
</tr>
<tr>
<td>6 DOMINANT DRIER UPLAND</td>
<td>NOMINAL</td>
</tr>
<tr>
<td>7 MONTANE</td>
<td>NOMINAL</td>
</tr>
<tr>
<td>8 DOMINANT MONTANE</td>
<td>NOMINAL</td>
</tr>
<tr>
<td>9 COAST HINTERLAND</td>
<td>NOMINAL</td>
</tr>
<tr>
<td>10 DOMINANT COAST HINTERLAND</td>
<td>NOMINAL</td>
</tr>
<tr>
<td>11 MISTBELT</td>
<td>NOMINAL</td>
</tr>
<tr>
<td>12 DOMINANT MISTBELT</td>
<td>NOMINAL</td>
</tr>
<tr>
<td>13 BUSHVELD</td>
<td>NOMINAL</td>
</tr>
<tr>
<td>14 DOMINANT BUSHVELD</td>
<td>NOMINAL</td>
</tr>
<tr>
<td>15 COAST LOWLANDS</td>
<td>NOMINAL</td>
</tr>
<tr>
<td>16 DOMINANT COAST LOWLANDS</td>
<td>NOMINAL</td>
</tr>
<tr>
<td>17 NONCOAST</td>
<td>NOMINAL</td>
</tr>
<tr>
<td>18 MEAN ANNUAL PRECIPITATION</td>
<td>CONTINUOUS</td>
</tr>
<tr>
<td>19 MINIMUM JULY TEMPERATURE</td>
<td>CONTINUOUS</td>
</tr>
<tr>
<td>20 S.E. MINIMUM JULY TEMPERATURES FOR EACH CELL</td>
<td>CONTINUOUS</td>
</tr>
<tr>
<td>21 SPECIES RICHNESS</td>
<td>CONTINUOUS</td>
</tr>
<tr>
<td>22 LATITUDE</td>
<td>CONTINUOUS</td>
</tr>
</tbody>
</table>

To test whether the environmental variables were significantly correlated to the species data, unrestricted Monte Carlo permutation tests of the individual axes were made to determine whether the eigenvalues of the canonical axes were greater than would be expected given no relationship between the species gradients and environmental variables. The eigenvalue for each canonical axis was compared against eigenvalues obtained by relating the species gradients randomly to the environmental variables in 99 unrestricted permutations of the data. Ninety-nine permutations is the default number...
used by CANOCO and was also the maximum number that could be processed by the computer used for the analysis.

Results from the CCA analyses were displayed as ordination diagrams depicting two axes at a time. Continuous environmental variables (e.g. mean annual precipitation) are shown as arrows representing the direction of maximum change in the variable. The longer the arrow the more important the variable. Nominal environmental variables (e.g. habitat type) are best represented in ordination diagrams as a point at the centroid (weighted average) of the sites belonging to that class (ter Braak 1987-92). Species are represented by points corresponding to their optima in two-dimensional space. The relative position of the species optima with respect to continuous environmental variables is obtained by dropping a perpendicular from the species centroid to the arrow. Since the mean of the environmental variables lies at the origin of the plot, species on the same side of the origin as the arrow have above average responses to that particular variable. Species occurring close to the centre of the biplot are therefore widespread, generalist species which do not show a strong response to any particular variable. The ordination diagram gives an indication of patterns of diversity and complementarity for the different taxa. Sites that are clustered in the same part of the diagram will have similar species composition and will show low complementarity, sites at different ends of the axes will be the most complementary. By comparing the ordination diagram for the different taxa it is possible to determine whether patterns of complementarity are similar for different taxa.

Quantifying complementarity using turnover and similarity measures.

In order to predict the ability of a small subset of sites to represent regional diversity for the different taxa, the amount of species turnover (where turnover is defined as a measure of spatial complementarity) along several transects in the region was quantified.

Measurement of species turnover.

There are a number of different measures which can be used to quantify the complementarity or dissimilarity of different sites. Many of the measures have
specifically been used to measure beta diversity along environmental gradients but they can be used more broadly to measure the degree to which sites have different species complements.

The most widely used beta diversity measure is that of Whittaker (1960):

\[ \beta = \frac{S}{\alpha} - 1 \]  

(7)

where \( S \) is the total number of species in the region and \( \alpha \) is the mean species richness of the individual sites.

This therefore measures the degree to which regional species richness exceeds average local richness. If complementarity between sites is high then the total regional species richness will be much greater than the mean local richness so beta will be high. If there is no complementarity between sites i.e. they all have a similar species composition then regional diversity will be equal to local species richness and beta will be zero.

However there are a number of problems associated with the use of dissimilarity measures like the one above. One of the main problems is how to deal with variation in site (alpha) diversity. Wilson & Schmida (1984) evaluated six different measures of turnover or dissimilarity in terms of their ability to quantify the desired characteristics of spatial turnover along a gradient of compositional change. Only one measure - that of Whittaker (1960) as described above - accurately reflected the extremes of community turnover when within site diversity varied. One other measure worked if the mean site diversity at the ends of the transect was equal to the mean site diversity over the whole gradient. Two of the measures required within site diversity to be constant across the whole gradient and two other measures did not reflect community turnover under any condition. However Whittaker's measure does not in fact give a reasonable measure of turnover under all conditions as was observed by Harrison et al. (1992). They pointed out that Whittaker's measure will overestimate turnover in cases where there is a nested
dropout of species along a richness gradient. If sites with not many species contain an exact subset of the species in the most species rich site, then the value for regional diversity will be greater than the mean alpha value even though there is no turnover (in the sense of complementarity) in species composition between sites. For example:

Site 1.

species: a, b, c, d, e

Site 2.

species: d, e

Total regional species richness = 5 species
Mean within site richness = (5+2)/2 = 3.5
Whittaker’s beta diversity = 0.43 (43%) when there is no complementarity between these two sites.

Harrison et al. (1992) suggest that one way around this problem would be to use the maximum local richness instead of mean local richness:


\[
\beta = \left( \frac{S}{a_{\text{max}}} \right) - 1 \times 100
\]

Where \( S \) is the total number of species in the sample; \( a_{\text{max}} \) is the maximum alpha diversity value in the sample; and \( n \) is the number of sites in the sample.

This beta diversity score measures the extent to which regional diversity (that of the whole transect) exceeds the maximum level of within site (alpha) diversity and can range from 0 (no turnover between sites) to 100 (100% turnover between sites, hence each site has a completely unique species complement). If there is no turnover in species composition between sites then \( S \) will equal \( a_{\text{max}} \) so \( \beta \) will equal 0. If turnover is high then \( S \) will be much greater than \( a_{\text{max}} \) and so \( \beta \) will be higher. Using the maximum value of alpha diversity in the score controls for variation in alpha diversity across the gradient (Harrison et al. 1992). In the above example total regional diversity is five species and
the maximum value of alpha diversity is also five species and so the calculated turnover will be zero.

However there is also a problem with the use of this modification, which while solving the problem of nested dropout will underestimate complementarity in situations where there is a large variation in site diversity as in the following example:

**Site 1.**

*species: a, b, c, d, e*

**Site 2.**

*species: f, g*

If Whittaker's measure is used in this case turnover would be 100% which is correct in terms of between site complementarity since there is complete compositional turnover between sites even though the first site contains many more species than the second. However using the Harrison et al. (1992) modified measure turnover would be 40% which is an under-estimate of the complementarity.

Overall the Harrison et al. (1992) measure of turnover will never overestimate complementarity, but may underestimate it when the maximum species richness of the gradient is much greater than the mean value. Whittaker's measure will tend to overestimate complementarity when there is a big difference between the richness of individual sites. For the KwaZulu Natal data this was not a major problem for the bird and mammal data (as there was not a large difference between the richness of diverse sites compared to species poor sites), but was likely to affect the turnover measures for the plant data.

One solution to the problem of large alpha diversity variation is to use a modified version of Sørenson's (1948) coefficient of similarity. Sørenson's coefficient is a pairwise measure calculated from the number of species shared by two sites relative to the number of species in each site:

\[
CC = \frac{2S_{s}}{S_{j} + S_{k}}
\]  (9)
where $S_j$ is the number of species shared by the two sites, $S_j$ is the number of species in site $j$ and $S_k$ the number of species in site $k$. If the two sites are completely similar then $2S_j = S_j + S_k$ and $CC = 1$. If two sites are completely dissimilar then $2S_j = 0$ and $CC = 0$.

Sørenson's measure is effectively the same as Whittaker's except that it measures similarity, not dissimilarity (the two measures added together for a pair of sites should equal unity). So Sørenson's measure can accommodate variation in alpha diversity when sites are completely dissimilar, but is affected by the number of species in each site when sites do share some species. One way round this is to express the number of species shared by two sites by the maximum potential number of species that they could share. For example if one site contained five species and the other site two species the maximum number of species they could share would be two. In other words the maximum number of species that two sites could potentially share if there were no turnover is the number of species in the site with the fewer species:

$$CC_k = \frac{S}{S_{\text{min}}}$$

(10)

Where $CC_k$ is the modified Sørenson coefficient, $S$ is the number of species shared by two sites and $S_{\text{min}}$ is the number of species in the site with the fewer species (i.e. the maximum number of species two sites could share).

$CC_k$ can cope with nested dropout of species as well as large variations in the total number of species between pairs of sites. It is more time consuming to calculate since it requires calculating the similarity of individual pairs of sites. To calculate the similarity of a gradient of sites would require the $CC$ value for every pairwise combination of sites to be evaluated and a mean similarity calculated.

Whittaker's measure could have been modified in a similar way - with turnover expressed as the number of unique species over the number of potential species (derived by adding the number of species in each component site). The problem with this approach is that...
the total potential species will often exceed the number of species in the data set and so underestimate complementarity.

Species turnover was measured along fourteen east-west transects across a steep temperature gradient and the mean ‘beta diversity’ value calculated for each taxon using both Whittaker’s (equation 7) and Harrison’s (equation 8) measures. Habitats along this gradient ranged from montane in the west through highland, moist upland, drier upland, mistbelt, bushveld, coastal hinterland to coastal lowland in the east. This transect represented the major environmental gradient extracted by the CCA analysis. For the mammal data, beta diversity was measured at three different scales - half, quarter and eighth degree grid square - (the original mammal atlases were mapped at eighth degree grid square resolution) to determine whether scale had a significant effect on the turnover values obtained.

Additionally, $CC_i$ values were calculated for birds and plants for 168 pairs of sites taken from the same dominant habitat type and 168 pairs of sites taken from different dominant habitat types to test the prediction that sites from different habitat types should show greater complementarity that pairs of sites from the same habitat type.

**Identifying surrogates for wholesale diversity.**

To evaluate the ability of different potential surrogates at representing regional diversity in KwaZulu Natal, networks of sites were selected using the alternative surrogate measures. The effectiveness of the different networks at representing the diversity of other taxa was then evaluated.

Selected networks for each taxon comprised the top five percent of sites scored for richness of surrogate species sets. Five percent of sites in KwaZulu Natal is eight out of the total 166 sites and although it represents an arbitrary cutoff point for reserve selection, five percent is a working minimum since 5.8% of the land area of South Africa is currently designated as reserves (Siegfried 1989). The following surrogate measures were considered:
1. plant species richness
2. bird species richness
3. mammal species richness
4. endemic bird richness
5. endemic grass richness
6. bird family richness (higher taxon richness)
7. habitat specific passerine richness
8. environmental representativeness - a measure created by selecting one site, at random, from each of the eight habitat types (montane, highland, moist upland, drier upland, mistbelt, bushveld, coastal hinterland and coastal lowland). Only sites where the desired habitat type was the sole, dominant vegetation class were included. Six sets of eight environmentally representative sites were selected.

How well do subsets of sites represent the regional diversity of different groups?
Each priority site network was evaluated in terms of its ability to represent the different aspects of regional diversity for each of the different taxa. All surrogates were compared to one hundred random selections of eight sites.

RESULTS.
Correlations between site level richness for different taxa and diversity components and overlap in top sites.
There were significant positive correlations between the number of bird species in a site and the number of plant species \(r = 0.449, n = 166, P<0.001\) and also between the number of birds and the number of mammals \(r = 0.384, n = 166, P<0.001\), but there was no significant correlation between the number of mammal species and the number of plant species in each site \(r = 0.127, n = 166, NS\) (Table 3.2). It should be noted that the significance tests for these pairwise correlations are not independent of one another. One method of controlling for this non-independence is to adjust the significance level required using the Bonferroni method. Since all but one of the significant correlations were significant at a probability level of 0.1%, these correlations still remain significant.
to at least a 5% level, as does the one correlation that was significant at the 1% level prior to any adjustment for non independence. Of the top eight sites for bird and plant richness three sites were common to both taxa, and of the top sixteen sites, five were common to birds and plants (Figure 3.3). The richest areas for bird species in KwaZulu Natal are in the north-east bushveld and also along the east/south-east coastal lowlands. The richest areas for plant species are along the south-east coast but also in the south-west Drakensberg mountains. Three of the top eight and eight of the top sixteen bird and mammal sites were the same. But only one of the top sixteen sites for mammal and plant species richness overlapped. Top sites for mammals are mainly in the north-eastern dry uplands and in the north-east bushveld.

Table 3.2. Correlation coefficients (r) between species richness of quarter degree grid cells (n=166) for different groups in KwaZulu Natal.

<table>
<thead>
<tr>
<th></th>
<th>BIRD SPECIES RICHNESS</th>
<th>MAMMAL SPECIES RICHNESS</th>
<th>PLANT SPECIES RICHNESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIRD SPECIES RICHNESS</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>MAMMAL SPECIES RICHNESS</td>
<td>0.384***</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PLANT SPECIES RICHNESS</td>
<td>0.449***</td>
<td>0.127**(NS)</td>
<td>-</td>
</tr>
<tr>
<td>HABITAT SPECIFIC PASSERINES</td>
<td>0.386***</td>
<td>0.209**</td>
<td>0.540***</td>
</tr>
<tr>
<td>ENDEMIC BIRD RICHNESS</td>
<td>-0.108**(NS)</td>
<td>0.120**(NS)</td>
<td>0.247***</td>
</tr>
<tr>
<td>ENDEMIC GRASS SPECIES</td>
<td>0.057**(NS)</td>
<td>0.103**(NS)</td>
<td>0.647***</td>
</tr>
<tr>
<td>BIRD FAMILY RICHNESS</td>
<td>0.928***</td>
<td>0.417***</td>
<td>0.423***</td>
</tr>
</tbody>
</table>

*** P ≤ 0.001  
** P ≤ 0.01  
(N) Non significant

Richness of habitat specific passerines was significantly correlated with species richness for plants, birds and mammals (Table 3.2). Of the top sixteen sites for richness of habitat specific species, seven were top sites for bird richness (43.8%), nine for plant richness (56.3%) but only three for mammal richness (18.8%) (Table 3.3). Richness of habitat specific passerines was most strongly correlated to plant species richness (Table 3.2). Bird family richness was very strongly correlated with bird species richness (r = 0.928,
Figure 3.3. Concordance of top sites for bird, mammal and plant species richness in quarter degree squares in KwaZulu Natal, southern Africa. The top eight squares for each taxon are shown.
Figure 3.4. Relationship between bird species richness and bird family richness of quarter degree grid cells in KwaZulu Natal, southern Africa.

$r = 0.928, n = 166$
P<0.001) and top sites for bird family richness were virtually the same as those for bird species richness (Table 3.2, Figure 3.4).

Endemic birds were not generally good at reflecting richness patterns of the different taxa. Endemic bird richness showed a negative (though not significant) correlation with overall bird richness and was not significantly correlated with mammal richness in the region (Table 3.2). Endemic bird richness was significantly correlated with plant richness but the correlation was weaker than that of habitat specific passerines (Table 3.2). Similarly, endemic grasses were not significantly correlated to bird or mammal species richness, but they did show a positive correlation with plant species richness (Table 3.2).

So in terms of general patterns of richness, birds and plants show some of the same richness gradients as do birds and mammals. Plants and mammals show less similarity in richness patterns. Endemic birds have very different richness patterns to non-endemic bird species and mammal species, but show some overlap with plant species richness patterns. Endemic grasses show very similar richness patterns to endemic birds. Habitat specific bird richness is a good correlate of bird and plant species richness.

Table 3.3. Percentage overlap of top sites for bird, mammal and plant species richness with sites selected using alternative diversity criteria.

<table>
<thead>
<tr>
<th>PERCENTAGE OVERLAP OF TOP SITES</th>
<th>BIRDS</th>
<th>MAMMALS</th>
<th>HABITAT SPECIFIC PASSERINES</th>
<th>ENDEMIC BIRDS</th>
<th>ENDEMIC GRASSES</th>
<th>BIRD FAMILY RICHNESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIRDS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>TOP 8 SITES</td>
<td>-</td>
<td>-</td>
<td>12.5</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>TOP 16 SITES</td>
<td>-</td>
<td>-</td>
<td>43.8</td>
<td>12.5</td>
<td>12.5</td>
<td>100</td>
</tr>
<tr>
<td>MAMMALS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>37.5</td>
</tr>
<tr>
<td>TOP 8 SITES</td>
<td>37.5</td>
<td>-</td>
<td>12.5</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>TOP 16 SITES</td>
<td>50.0</td>
<td>-</td>
<td>18.8</td>
<td>6.25</td>
<td>0</td>
<td>50.0</td>
</tr>
<tr>
<td>PLANTS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>37.5</td>
</tr>
<tr>
<td>TOP 8 SITES</td>
<td>37.5</td>
<td>12.5</td>
<td>37.5</td>
<td>12.5</td>
<td>25.0</td>
<td></td>
</tr>
<tr>
<td>TOP 16 SITES</td>
<td>50.0</td>
<td>6.25</td>
<td>56.3</td>
<td>43.8</td>
<td>50.0</td>
<td>31.3</td>
</tr>
</tbody>
</table>
Spatial autocorrelation in the data sets.

However there are problems in looking for correlations between two spatially distributed variables since there may be spatial autocorrelation in the data set. Spatial autocorrelation was measured using the Moran Coefficient \((I)\). The Moran Coefficient \((I)\) for birds was 0.308. This indicates that the bird distribution data were more clustered than the expected value. Under the normality significance test \(z = 7.61\) and assuming randomisation \(z = 7.64\). The one-tailed critical value at a 0.001 significance level is 3.09. A one-tailed test was used because the data are expected to show clustering of like values. The data therefore exhibit strong, positive spatial autocorrelation. That is similar values of bird species richness tend to be more clustered than would be predicted by chance. For plants, \(I\) was 0.283, with \(z_N = 6.99\) and \(z_R = 7.03\), \(P<0.001\). So as with birds, the plant data appear more clustered than expected although not as strongly as the bird data. The \(I\) value for mammals was 0.354, with \(z_N = 8.73\) and \(z_R = 8.72\), \(P<0.001\). So the mammal distribution data also show strong, positive autocorrelation.

This means that similar values of species richness for each taxon tend to be more spatially clustered than would be predicted by chance. Adjacent values of species richness for each taxon therefore cannot be assumed to be independent and this will reduce the degrees of freedom in any correlation between species richness across the different taxa. In an attempt to control for this, correlation coefficients were calculated for bird, mammal and plant species richness with only non-adjacent grid squares included in the analysis. This generated four sets of data on which correlation coefficients could be calculated. When only non-adjacent were included there was no significant correlation between mammal and plant species richness for any of the data sets (\(r\) values of 0.008, \(n=45\); 0.056, \(n=40\); 0.329, \(n=43\) & 0.028, \(n=38\), all NS). Two of the correlation coefficients for plant and bird species richness were non-significant, although there was still a significant positive correlation for two of the data sets (\(r\) values of 0.337, \(n=45\), NS; 0.498, \(n=40\), \(P<0.01\); 0.550, \(n=43\), \(P<0.01\) & 0.357, \(n=38\), NS). Similarly, two of the correlation coefficients between mammal and bird species richness were non significant when only non-adjacent squares were included, although there was still a significant positive correlation for two of the data sets (\(r\) values of 0.247, \(n=45\), NS;
0.184, n=40, NS; 0.533, n=43, P<0.01 & 0.558, n=38, P<0.01). As with the previous correlations there is the problem of non-independence between the pairwise comparisons, although those correlations that are significant without any adjustment have significance levels high enough that they would remain significant even after adjustment (e.g. by the Bonferroni method).

Patterns of diversity and complementarity and their environmental correlates.

Eigenvalues of first four axes and the cumulative percentage variation in species data explained by these axes are given for the different taxa for both CA and CCA (Table 3.4). Values for CCA are only slightly lower than for unconstrained ordination for all of the taxa indicating that the environmental variables used do explain the major patterns of variation in species composition. All four axes for birds and mammals are significant at the 0.01 probability level, but most of the variance in the data is accounted for by axes 1 and 2 (Table 3.5). Axes 1, 2 and 3 are significant at the 0.01 level for Poa1 but only axis 1 is significant at this level for Poa2 (although the second axis is significant at the 0.05 level). To simplify comparisons only the first two ordination axes are considered, these represent between 7.6 (Poa2) and 26.5% (Passerines) of the biological variation (Table 3.5).

Interset correlations of the environmental variables with axes 1 and 2 are given in Table 3.6 (a & b). The same sets of environmental variables are important for all taxa. CCA axis one largely reflects the minimum mean July temperature gradient which separates species with an easterly, coastal distribution (high minimum mean July temperature) from species occurring in the highlands and western interior which is characterised by lower minimum temperatures. The second axis is a mean precipitation/latitude gradient which separates species occurring in high rainfall areas - mostly the extreme eastern and south-eastern coastal forest belt plus montane and mistbelt areas in the west from those species characteristic of drier upland interior areas and north-eastern bushveld savanna regions. Species with centroids on the right hand side of the ordination are those whose distributions are correlated with low mean, minimum July temperatures (i.e. species occurring in the higher altitude areas in the west). Species with centroids on the left side
Table 3.4. Eigenvalues & cumulative percentage variance explained by the first four axes extracted by correspondence analysis compared to canonical correspondence analysis of sites in KwaZulu Natal, southern Africa based on (a) passerine birds, (b) nonpasserine birds, (c) mammals (carnivores & ungulates), (d) grasses (POA1) & (e) grasses (POA2).

(a) PASSERINE BIRDS

<table>
<thead>
<tr>
<th>AXIS</th>
<th>EIGENVALUE</th>
<th>CUMULATIVE % VARIANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CA</td>
<td>CCA</td>
</tr>
<tr>
<td>1</td>
<td>0.258</td>
<td>0.243</td>
</tr>
<tr>
<td>2</td>
<td>0.122</td>
<td>0.107</td>
</tr>
<tr>
<td>3</td>
<td>0.073</td>
<td>0.050</td>
</tr>
<tr>
<td>4</td>
<td>0.064</td>
<td>0.045</td>
</tr>
</tbody>
</table>

(b) NONPASSERINE BIRDS

<table>
<thead>
<tr>
<th>AXIS</th>
<th>EIGENVALUE</th>
<th>CUMULATIVE % VARIANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CA</td>
<td>CCA</td>
</tr>
<tr>
<td>1</td>
<td>0.243</td>
<td>0.228</td>
</tr>
<tr>
<td>2</td>
<td>0.124</td>
<td>0.107</td>
</tr>
<tr>
<td>3</td>
<td>0.094</td>
<td>0.073</td>
</tr>
<tr>
<td>4</td>
<td>0.060</td>
<td>0.041</td>
</tr>
</tbody>
</table>

(c) MAMMALS

<table>
<thead>
<tr>
<th>AXIS</th>
<th>EIGENVALUE</th>
<th>CUMULATIVE % VARIANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CA</td>
<td>CCA</td>
</tr>
<tr>
<td>1</td>
<td>0.357</td>
<td>0.321</td>
</tr>
<tr>
<td>2</td>
<td>0.214</td>
<td>0.170</td>
</tr>
<tr>
<td>3</td>
<td>0.192</td>
<td>0.107</td>
</tr>
<tr>
<td>4</td>
<td>0.129</td>
<td>0.088</td>
</tr>
</tbody>
</table>

(d) GRASSES - POA1

<table>
<thead>
<tr>
<th>AXIS</th>
<th>EIGENVALUE</th>
<th>CUMULATIVE % VARIANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CA</td>
<td>CCA</td>
</tr>
<tr>
<td>1</td>
<td>0.552</td>
<td>0.471</td>
</tr>
<tr>
<td>2</td>
<td>0.330</td>
<td>0.223</td>
</tr>
<tr>
<td>3</td>
<td>0.256</td>
<td>0.150</td>
</tr>
<tr>
<td>4</td>
<td>0.234</td>
<td>0.107</td>
</tr>
</tbody>
</table>

(e) GRASSES - POA2

<table>
<thead>
<tr>
<th>AXIS</th>
<th>EIGENVALUE</th>
<th>CUMULATIVE % VARIANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CA</td>
<td>CCA</td>
</tr>
<tr>
<td>1</td>
<td>0.491</td>
<td>0.441</td>
</tr>
<tr>
<td>2</td>
<td>0.348</td>
<td>0.193</td>
</tr>
<tr>
<td>3</td>
<td>0.292</td>
<td>0.133</td>
</tr>
<tr>
<td>4</td>
<td>0.250</td>
<td>0.110</td>
</tr>
</tbody>
</table>
Table 3.5. Eigenvalues for the first four axes extracted by Canonical correspondence analysis (CCA) for each group (passerines, nonpasserines, mammals and grasses (POAI & POA2)) with their significance levels calculated from 99 Monte Carlo permutations.

<table>
<thead>
<tr>
<th></th>
<th>AXIS 1</th>
<th>AXIS 2</th>
<th>AXIS 3</th>
<th>AXIS 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EIGENVALUES</strong></td>
<td>0.243**</td>
<td>0.107**</td>
<td>0.05**</td>
<td>0.045**</td>
</tr>
<tr>
<td><strong>CUM % VARIANCE IN SPECIES</strong></td>
<td>18.4</td>
<td>26.5</td>
<td>30.3</td>
<td>33.7</td>
</tr>
<tr>
<td><strong>CUM % VARIANCE IN SPP-ENV</strong></td>
<td>41.2</td>
<td>59.3</td>
<td>67.8</td>
<td>75.5</td>
</tr>
<tr>
<td><strong>SUM UNCONST EIGENVALUES</strong></td>
<td></td>
<td></td>
<td></td>
<td>1.318</td>
</tr>
<tr>
<td><strong>SUM CANONICAL EIGENVALUES</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.589</td>
</tr>
</tbody>
</table>

**PASSERINES**

<table>
<thead>
<tr>
<th></th>
<th>AXIS 1</th>
<th>AXIS 2</th>
<th>AXIS 3</th>
<th>AXIS 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EIGENVALUES</strong></td>
<td>0.228**</td>
<td>0.107**</td>
<td>0.073**</td>
<td>0.041**</td>
</tr>
<tr>
<td><strong>CUM % VARIANCE IN SPECIES</strong></td>
<td>14.6</td>
<td>21.4</td>
<td>26.1</td>
<td>28.7</td>
</tr>
<tr>
<td><strong>CUM % VARIANCE IN SPP-ENV</strong></td>
<td>36.7</td>
<td>53.9</td>
<td>65.6</td>
<td>72.1</td>
</tr>
<tr>
<td><strong>SUM UNCONST EIGENVALUES</strong></td>
<td></td>
<td></td>
<td></td>
<td>1.567</td>
</tr>
<tr>
<td><strong>SUM CANONICAL EIGENVALUES</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.623</td>
</tr>
</tbody>
</table>

**NONPASSERINES**

<table>
<thead>
<tr>
<th></th>
<th>AXIS 1</th>
<th>AXIS 2</th>
<th>AXIS 3</th>
<th>AXIS 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EIGENVALUES</strong></td>
<td>0.321**</td>
<td>0.170**</td>
<td>0.107**</td>
<td>0.088**</td>
</tr>
<tr>
<td><strong>CUM % VARIANCE IN SPECIES</strong></td>
<td>12.2</td>
<td>18.7</td>
<td>22.8</td>
<td>26.2</td>
</tr>
<tr>
<td><strong>CUM % VARIANCE IN SPP-ENV</strong></td>
<td>33.4</td>
<td>51.1</td>
<td>62.2</td>
<td>71.3</td>
</tr>
<tr>
<td><strong>SUM UNCONST EIGENVALUES</strong></td>
<td></td>
<td></td>
<td></td>
<td>2.623</td>
</tr>
<tr>
<td><strong>SUM CANONICAL EIGENVALUES</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.962</td>
</tr>
</tbody>
</table>

**MAMMALS**

<table>
<thead>
<tr>
<th></th>
<th>AXIS 1</th>
<th>AXIS 2</th>
<th>AXIS 3</th>
<th>AXIS 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EIGENVALUES</strong></td>
<td>0.471**</td>
<td>0.223**</td>
<td>0.15**</td>
<td>0.107</td>
</tr>
<tr>
<td><strong>CUM % VARIANCE IN SPECIES</strong></td>
<td>5.5</td>
<td>8.1</td>
<td>9.8</td>
<td>11.1</td>
</tr>
<tr>
<td><strong>CUM % VARIANCE IN SPP-ENV</strong></td>
<td>25.7</td>
<td>37.9</td>
<td>46.0</td>
<td>51.9</td>
</tr>
<tr>
<td><strong>SUM UNCONST EIGENVALUES</strong></td>
<td></td>
<td></td>
<td></td>
<td>8.594</td>
</tr>
<tr>
<td><strong>SUM CANONICAL EIGENVALUES</strong></td>
<td></td>
<td></td>
<td></td>
<td>1.832</td>
</tr>
</tbody>
</table>

**POA1**

<table>
<thead>
<tr>
<th></th>
<th>AXIS 1</th>
<th>AXIS 2</th>
<th>AXIS 3</th>
<th>AXIS 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EIGENVALUES</strong></td>
<td>0.441**</td>
<td>0.193*</td>
<td>0.133</td>
<td>0.110</td>
</tr>
<tr>
<td><strong>CUM % VARIANCE IN SPECIES</strong></td>
<td>5.3</td>
<td>7.6</td>
<td>9.2</td>
<td>10.5</td>
</tr>
<tr>
<td><strong>CUM % VARIANCE IN SPP-ENV</strong></td>
<td>25.2</td>
<td>36.2</td>
<td>43.8</td>
<td>50.1</td>
</tr>
<tr>
<td><strong>SUM UNCONST EIGENVALUES</strong></td>
<td></td>
<td></td>
<td></td>
<td>8.376</td>
</tr>
<tr>
<td><strong>SUM CANONICAL EIGENVALUES</strong></td>
<td></td>
<td></td>
<td></td>
<td>1.751</td>
</tr>
</tbody>
</table>

**POA2**

<table>
<thead>
<tr>
<th></th>
<th>AXIS 1</th>
<th>AXIS 2</th>
<th>AXIS 3</th>
<th>AXIS 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EIGENVALUES</strong></td>
<td>0.243**</td>
<td>0.107**</td>
<td>0.05**</td>
<td>0.045**</td>
</tr>
<tr>
<td><strong>CUM % VARIANCE IN SPECIES</strong></td>
<td>18.4</td>
<td>26.5</td>
<td>30.3</td>
<td>33.7</td>
</tr>
<tr>
<td><strong>CUM % VARIANCE IN SPP-ENV</strong></td>
<td>41.2</td>
<td>59.3</td>
<td>67.8</td>
<td>75.5</td>
</tr>
<tr>
<td><strong>SUM UNCONST EIGENVALUES</strong></td>
<td></td>
<td></td>
<td></td>
<td>1.318</td>
</tr>
<tr>
<td><strong>SUM CANONICAL EIGENVALUES</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.589</td>
</tr>
</tbody>
</table>

**•• P <= 0.01 (FROM 99 MONTE CARLO PERMUTATIONS)**

**• P <= 0.05 (FROM 99 MONTE CARLO PERMUTATIONS)**
<table>
<thead>
<tr>
<th>Table 3.6a. Interset correlations of the first axis of a CCA with environmental variables arranged in rank order for each taxon.</th>
</tr>
</thead>
<tbody>
<tr>
<td>PASSERINES</td>
</tr>
<tr>
<td>------------</td>
</tr>
<tr>
<td>HIGHLAND</td>
</tr>
<tr>
<td>DHIGH</td>
</tr>
<tr>
<td>MOIST UP</td>
</tr>
<tr>
<td>NONCOAST</td>
</tr>
<tr>
<td>DRIER UP</td>
</tr>
<tr>
<td>MONTANE</td>
</tr>
<tr>
<td>DDU</td>
</tr>
<tr>
<td>DMU</td>
</tr>
<tr>
<td>DMONT</td>
</tr>
<tr>
<td>SEMINJUL</td>
</tr>
<tr>
<td>LATITUDE</td>
</tr>
<tr>
<td>MISTBELT</td>
</tr>
<tr>
<td>DMI</td>
</tr>
<tr>
<td>MEANPREC</td>
</tr>
<tr>
<td>DCH</td>
</tr>
<tr>
<td>COAST HI</td>
</tr>
<tr>
<td>DB</td>
</tr>
<tr>
<td>DCL</td>
</tr>
<tr>
<td>BIRD DIV</td>
</tr>
<tr>
<td>BUSHVELD</td>
</tr>
<tr>
<td>COAST LO</td>
</tr>
<tr>
<td>MINT JUL</td>
</tr>
</tbody>
</table>

MINT JUL is the mean minimum July temperature calculated for each quarter degree square.
SEMINJUL is the standard error of the mean minimum July temperature calculated from the one minute data for each quarter degree square.
COAST HI is coast hinterland habitat; COAST LO is coast lowland habitat; DCH is dominant coast hinterland & DCL is dominant coast lowland.
DMONT is dominant montane habitat; DMU is dominant moist upland; DDU is dominant dry upland; DMI is dominant mistbelt; DB is dominant bushveld.
BIRD DIV is the number of bird species recorded from each quarter degree square; MAMM DIV is the number of mammal species recorded from each quarter degree square.
PLNT DIV is the number of plant species recorded from each quarter degree square.
<table>
<thead>
<tr>
<th>Taxon</th>
<th>Axis 2</th>
<th>Axis 2</th>
<th>Axis 2</th>
<th>Axis 2</th>
<th>Axis 2</th>
<th>Axis 2</th>
<th>Axis 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEANPREC</td>
<td>0.753</td>
<td>0.738</td>
<td>0.624</td>
<td>0.675</td>
<td>MEANPREC</td>
<td>0.712</td>
<td></td>
</tr>
<tr>
<td>LATITUDE</td>
<td>0.680</td>
<td>0.654</td>
<td>0.566</td>
<td>DCL</td>
<td>0.616</td>
<td>DCL</td>
<td>0.562</td>
</tr>
<tr>
<td>DCL</td>
<td>0.467</td>
<td>0.592</td>
<td>0.534</td>
<td>LATITUDE</td>
<td>0.580</td>
<td>COAST LO</td>
<td>0.430</td>
</tr>
<tr>
<td>COAST LO</td>
<td>0.435</td>
<td>0.545</td>
<td>0.493</td>
<td>COAST LO</td>
<td>0.562</td>
<td>LATITUDE</td>
<td>0.412</td>
</tr>
<tr>
<td>COAST HI</td>
<td>0.331</td>
<td>0.324</td>
<td>0.372</td>
<td>PLNT DIV</td>
<td>0.364</td>
<td>PLNT DIV</td>
<td>0.245</td>
</tr>
<tr>
<td>DCH</td>
<td>0.305</td>
<td>0.317</td>
<td>0.359</td>
<td>MINT JUL</td>
<td>0.300</td>
<td>MONTANE</td>
<td>0.232</td>
</tr>
<tr>
<td>SEMINJUL</td>
<td>0.291</td>
<td>0.265</td>
<td>0.311</td>
<td>COAST HI</td>
<td>0.187</td>
<td>SEMINJUL</td>
<td>0.213</td>
</tr>
<tr>
<td>MISTBELT</td>
<td>0.260</td>
<td>0.217</td>
<td>0.163</td>
<td>DCH</td>
<td>0.141</td>
<td>DMONT</td>
<td>0.177</td>
</tr>
<tr>
<td>DMI</td>
<td>0.193</td>
<td>SEMINJUL</td>
<td>0.207</td>
<td>MISTBELT</td>
<td>0.111</td>
<td>SEMINJUL</td>
<td>0.078</td>
</tr>
<tr>
<td>MINT JUL</td>
<td>0.184</td>
<td>0.184</td>
<td>BUSHVELD</td>
<td>0.107</td>
<td>DMI</td>
<td>0.037</td>
<td>DHIGH</td>
</tr>
<tr>
<td>MONTANE</td>
<td>0.107</td>
<td>0.047</td>
<td>DHIGH</td>
<td>0.058</td>
<td>MONTANE</td>
<td>0.031</td>
<td>DMI</td>
</tr>
<tr>
<td>DMONT</td>
<td>0.080</td>
<td>0.023</td>
<td>DMI</td>
<td>0.049</td>
<td>MISTBELT</td>
<td>0.020</td>
<td>DCH</td>
</tr>
<tr>
<td>DHIGH</td>
<td>0.026</td>
<td>BUSHVELD</td>
<td>0.019</td>
<td>MONTANE</td>
<td>0.048</td>
<td>DMONT</td>
<td>0.014</td>
</tr>
<tr>
<td>HIGHLAND</td>
<td>-0.054</td>
<td>-0.024</td>
<td>Dmont</td>
<td>0.034</td>
<td>DHIGH</td>
<td>-0.091</td>
<td>MISTBELT</td>
</tr>
<tr>
<td>DMU</td>
<td>-0.068</td>
<td>DHIGH</td>
<td>-0.043</td>
<td>HIGHLAND</td>
<td>-0.122</td>
<td>BUSHVELD</td>
<td>-0.167</td>
</tr>
<tr>
<td>BUSHVELD</td>
<td>-0.084</td>
<td>HIGHLAND</td>
<td>-0.164</td>
<td>DB</td>
<td>-0.152</td>
<td>DMU</td>
<td>-0.197</td>
</tr>
<tr>
<td>BIRD DIV</td>
<td>-0.129</td>
<td>DMU</td>
<td>-0.171</td>
<td>DMU</td>
<td>-0.265</td>
<td>HIGHLAND</td>
<td>-0.218</td>
</tr>
<tr>
<td>MOIST UP</td>
<td>-0.183</td>
<td>MOIST UP</td>
<td>-0.268</td>
<td>DDU</td>
<td>-0.427</td>
<td>MOIST UP</td>
<td>-0.325</td>
</tr>
<tr>
<td>DDU</td>
<td>-0.424</td>
<td>DB</td>
<td>-0.387</td>
<td>MAMM DIV</td>
<td>-0.463</td>
<td>DDU</td>
<td>-0.392</td>
</tr>
<tr>
<td>DB</td>
<td>-0.427</td>
<td>DDU</td>
<td>-0.436</td>
<td>MOIST UP</td>
<td>-0.469</td>
<td>DB</td>
<td>-0.448</td>
</tr>
<tr>
<td>NONCOAST</td>
<td>-0.468</td>
<td>DRIER UP</td>
<td>-0.541</td>
<td>NONCOAST</td>
<td>-0.483</td>
<td>DRIER UP</td>
<td>-0.518</td>
</tr>
<tr>
<td>DRIER UP</td>
<td>-0.480</td>
<td>NONCOAST</td>
<td>-0.547</td>
<td>DRIER UP</td>
<td>-0.589</td>
<td>NONCOAST</td>
<td>-0.572</td>
</tr>
</tbody>
</table>

MINT JUL is the mean minimum July temperature calculated for each quarter degree square.
SEMINJUL is the standard error of the mean minimum July temperature calculated from the one minute data for each quarter degree square.
COAST HI is coast hinterland habitat; COAST LO is coast lowland habitat; DCH is dominant coast hinterland & DCL is dominant coast lowland.
DMONT is dominant montane habitat; DMU is dominant moist upland; DDU is dominant dry upland; DMI is dominant mistbelt; DB is dominant bushveld.
BIRD DIV is the number of bird species recorded from each quarter degree square; MAMM DIV is the number of mammal species recorded from each quarter degree square.
PLNT DIV is the number of plant species recorded from each quarter degree square.
of the ordination are lowland species, with coastal or bushveld distributions. Species centred in the top part of the ordination are those whose distributions are correlated with higher precipitation areas - either coastal lowland forests (top, left of ordination) or wetter, interior areas like mistbelt or montane regions (top, right of ordination). Species with centroids in the bottom half of the ordination are those with distributions predominantly in the drier areas - either bushveld (bottom, left) or dry, upland habitats (bottom, right).

Interestingly passerine, nonpasserine and mammal species richness are all most strongly correlated with mean minimum July temperature (Figures 3.5 & 3.6) while plant species richness is more strongly related to mean precipitation than it is to temperature (Figure 3.7). Bird and mammal species richness is greatest in areas with the highest mean, minimum July temperatures - that is in the bushveld and coastal lowland areas in the north and east of KwaZulu Natal. Plant species richness is correlated with mean annual precipitation, being greatest in more southerly, coastal and mistbelt areas of KwaZulu Natal.

A large proportion of the bird species have their centroids close to the centre of the biplot (Figure 3.8). In contrast the majority of endemic bird species centroids fall some distance away from the origin of the biplot (Figures 3.9 & 3.10). Most of the endemics are on the right hand side of the temperature axis with their distributions centred in highland/montane/drier upland areas or in mistbelt forests. For passerine birds only six endemics have their centres of distribution in bushveld, and four species in the coastal forest belt. Nonpasserine endemics show similar distributional patterns to their passerine counterparts. No nonpasserine endemics are predominantly bushveld species, three are coastal lowland species and eleven species are concentrated in the western upland, highland and montane areas.

Threatened birds on the other hand are not restricted to any particular environmental gradient. (Figures 3.11 & 3.12). The majority of species are not widespread (i.e. they are not close to the centre of the ordination diagram), but they are evenly distributed with
Figure 3.5. Ordination diagram based on canonical correspondence analysis of passerine distribution data in 166 quarter degree squares in KwaZulu Natal, with respect to the environmental variables latitude (LATITUDE); mean minimum July temperature (MINT JUL); s.e. of the mean minimum July temperature (SEMINJUL); mean precipitation (MEANPREC); drier upland (DRIER UP); dominant drier upland (DDU); highland (HIGHLAND); dominant highland (DHIGH); montane (MONTANE); dominant montane (DMONT); moist upland (MOIST UP); dominant moist upland (DMU); mistbelt (MISTBELT); dominant mistbelt (DMI); bushveld (BUSHVELD); dominant bushveld (DB); coastal hinterland (COAST HI); dominant coastal hinterland (DCH); coastal lowland (COAST LO); dominant coastal lowland (DCL); noncoastal grid squares (NONCOAST) plus total bird diversity (BIRD DIV). Nominal environmental variables (e.g. habitat types) are depicted as centroids, continuous variables as arrows. Species data is shown on separate diagrams for clarity.
Figure 3.6. Ordination diagram based on canonical correspondence analysis of mammal distribution data in 166 quarter degree squares in KwaZulu Natal, with respect to the environmental variables latitude (LATITUDE); mean minimum July temperature (MINT JUL); s.e. of the mean minimum July temperature (SEMINJUL); mean precipitation (MEANPREC); drier upland (DRIER UP); dominant drier upland (DDU); highland (HIGHLAND); dominant highland (DHIGH); montane (MONTANE); dominant montane (DMONT); moist upland (MOIST UP); dominant moist upland (DMU); mistbelt (MISTBELT); dominant mistbelt (DMI); bushveld (BUSHVELD); dominant bushveld (DB); coastal hinterland (COAST HI); dominant coastal hinterland (DCH); coastal lowland (COAST LO); dominant coastal lowland (DCL); noncoastal grid squares (NONCOAST) plus total mammal diversity (MAMM DIV). Nominal environmental variables are depicted as centroids, continuous variables as arrows. Species data is shown on separate diagrams for clarity.
Figure 3.7. Ordination diagram based on canonical correspondence analysis of grass (POAI) distribution data in 166 quarter degree squares in KwaZulu Natal, with respect to the environmental variables latitude (LATITUDE); mean minimum July temperature (MINT JUL); s.e. of the mean minimum July temperature (SEMINJUL); mean precipitation (MEANPREC); drier upland (DRIER UP); dominant drier upland (DDU); highland (HIGHLAND); dominant highland (DHIGH); montane (MONTANE); dominant montane (DMONT); moist upland (MOIST UP); dominant moist upland (DMU); mistbelt (MISTBELT); dominant mistbelt (DMI); bushveld (BUSHVELD); dominant bushveld (DB); coastal hinterland (COAST HI); dominant coastal hinterland (DCH); coastal lowland (COAST LO); dominant coastal lowland (DCL); noncoastal grid squares (NONCOAST) plus total plant diversity (PLNT DIV). Nominal environmental variables are depicted as centroids, continuous variables as arrows. Species data is shown on separate diagrams for clarity.
Figure 3.8. Ordination diagram based on canonical correspondence analysis of passerine distribution data in KwaZulu Natal with respect to various environmental variables (omitted for clarity - see figure 3.5). Each of the 195 breeding passerine landbird species is shown as the centroid (▲) representing the species' optima in 2-dimensional space in the ordination.
Figure 3.9. Ordination diagram based on canonical correspondence analysis of passerine distribution data in KwaZulu Natal with respect to various environmental variables (omitted for clarity - see figure 3.5). Each of the endemic, breeding, passerine landbird species is shown as the centroid (▲) representing the species’ optima in 2-dimensional space in the ordination. Not all species are labelled: TURD FIS (Turdus fisheri); ERYT SIG (Erythopygia signata); COSS DIC (Cossypha dichroa); MALA OLI (Malacnotus olivaceous); BATI CAP (Batis capensis); LANI FER (Laniarius ferrugineus); TCHA TCH (Tchagra tchagra); SIEG SIEL (Sigelus silens); APAL RUD (Apalis ruddi); COSS HUM (Cossypha humeralis); SERI CIT (Serinus citriniceps); HYPAR MAR (Hypargos margaritatus); NECT NEE (Nectarinia neergaardi); SERI SCO (Serinus scotops); SERI ALB (Serinus albogularis); PROM GUR (Promerops gurneyi); PARI LAY (Parisoma layardi); ANTH CHL (Anthus chloris); CHAE AUR (Chaetops aurantius); CALA MAG (Calandrella magnirostris); PYCN NIG (Pycnonotus nigricans); TELO ZEY (Telephorus zeylonus); ANTH CRE (Anthus crenatus); MYRM FOR (Myrmecocichla formicivora); CERT CUR (Certhilauda curvirostris); HIRU SPI (Hirundo spilodera); AMAD ERY (Amandina erythrocephala); CALA CON (Calandrella conirostris); MIRA CHE (Mirafraxa cheniana).
Figure 3.10. Ordination diagram based on canonical correspondence analysis of nonpasserine distribution data in KwaZulu Natal with respect to various environmental variables (omitted for clarity). Each of the endemic, breeding, nonpasserine landbird species is shown as the centroid (▲) representing the species' optima in 2-dimensional space in the ordination. STA WOOD (Stactolaema woodwardi); TAU CORY (Tauraco corythaix); FRA NAT (Francolinus natalensis); BUT RUFO (Buteo rufofuscus); GYP COPR (Gyps coprotheres); ANT PARA (Anthropoides paradisea); GER CALV (Geronticus calvus); CUR RUFU (Cursorius rufus); EUP CAER (Eupodotis caerulescens); EUP AFR (Eupodotis afroides); GEO OLIV (Geolcolaptes olivaceous); FRA AFRI (Francolinus africanus); MEL CANO (Melierax canorus); CIR MAUR (Circus maurus).
Figure 3.11. Ordination diagram based on canonical correspondence analysis of passerine distribution data in KwaZulu Natal with respect to various environmental variables (ommitted for clarity - see figure 3.5). Each of the threatened, breeding, passerine landbird species is shown as the centroid representing the species' optima in 2-dimensional space in the ordination. Globally threatened (▲) TURD FIS (Turdus fisheri); HIRU ATR (Hirundo atrocaerulea); COSS DIC (Cossypha dichroa); LIOP NIG (Lioptilus nigricapillus); ANTH CHL (Anthus chloris); SERI SYM (Serinus symonsi); OENA BIF (Oenanthe bifasciata); BATI FRA (Batis fratum); APAL RUD (Apalis ruddi); SERI CIT (Serinus citriniceps); HYPA MAR (Hypargos margaritatus); NECT NEE (Nectarinia neergaardi). Only in South African red data book (+) SPER FRI (Spermestes fringilloides); PLAT PEL (Platysteira peltata); SMIT CAP (Smithornis capensis); ZOST SEN (Zosterops senegalesis); ANTH BRA (Anthus brachyurus).
Figure 3.12. Ordination diagram based on canonical correspondence analysis of nonpasserine distribution data in KwaZulu Natal with respect to various environmental variables (omitted for clarity). Each of the threatened, breeding, nonpasserine landbird species is shown as the centroid representing the species' optima in 2-dimensional space in the ordination. Globally threatened (▲) STA WOOD (Stactolaema woodwardi); CIR FASC (Circaetus fasciolatus); GYP COPR (Gyps coprotheres); GER CALV (Geronticus calvus); EUP CAER (Eupodotis caerulescens); CIR MAUR (Circus maurus). Only in South African red data book (+) CAP NATA (Caprimulgus natalensis); COL DELE (Columbia delegorgui); MAC ALCI (Macheiramphus alcinius); AVI CUCU (Aviceda cuculoides); BUC LEAD (Bucorvus leadbeateri); FAL PERE (Falco peregrinus); POI ROBU (Poicephalus robustus); GYP BARB (Gypaetus barbatus); POL BELL (Polemaetus bellicosus); NEO DENH (Neotis denhami); TUR HOTT (Turnix hottentotta); TOR TRAC (Torgus tracheliotus); TER ECAU (Terathopius ecaudatus); LEP CRUM (Leptoptilis crumeniferus); TRI OCCI (Trigonoceps occipitalis); VAN LUGI (Vanellus lugubris); NEC MONA (Necrosyrtes monachus).
respect to the environmental variation (i.e. not clustered in one area of the ordination). However, all the highland and upland threatened species were globally threatened species, for example the Drakensberg Serin (*Serinus symonsi*), Yellowbreasted Pipit (*Anthus chloris*), Orangebreasted Rockjumper (*Chaetops aurantius*) and Black Harrier (*Circus maurus*). In contrast, those species that were only in the South African Red Data Book were predominantly coastal lowland and bushveld species, for example, the Pied Mannikin (*Spermestes fringilloides*), Wattle-eyed Flycatcher (*Platysteira peltata*), African Broadbill (*Smithornis capensis*), Yellow White-eye (*Zosterops senegalensis*) and Bateleur (*Terathopius ecaudatus*). These species were concentrated on the left of the ordination diagram where sites with coastal lowland, coastal hinterland and bushveld habitat were clustered.

As expected passerine species classified as habitat specific on the basis of field guide descriptions generally had greater ordination scores than non habitat specific species i.e. they were strongly correlated with the major ordination axes (Figure 3.13). Unlike endemic species, habitat specialist species had representatives in all the major environmental types in the region.

Like endemic birds, endemic grasses were strongly correlated with the cool, western high altitude areas (Figure 3.14). Seventy-five percent of all grassland endemics occurred in areas of below average mean minimum July temperatures. A large proportion of these species are montane. Unlike bird endemics, however more than half the endemic Poaceae occur in areas of above average species richness.

The majority of mammal species had their optimum distributions in bushveld or dry upland areas. Very few were coastal lowland or montane species. As with birds threatened mammals were not confined to any one environmental gradient or habitat type (Figure 3.15).
Figure 3.13. Ordination diagram based on canonical correspondence analysis of passerine distribution data in KwaZulu Natal with respect to various environmental variables (omitted for clarity - see figure 3.5). Each of the habitat specific, breeding, passerine landbird species is shown as the centroid (▲) representing the species' optima in 2-dimensional space in the ordination.
Figure 3.14. Ordination diagram based on canonical correspondence analysis of grass distribution data (POA2) in KwaZulu Natal with respect to various environmental variables (omitted for clarity - see figure 3.7). Each of the endemic species is shown as the centroid (▲) representing the species' optima in 2-dimensional space in the ordination.
Figure 3.15. Ordination diagram based on canonical correspondence analysis of mammal distribution data in KwaZulu Natal with respect to various environmental variables (omitted for clarity - see figure 3.6). Each of the threatened mammal species is shown as the centroid (▲) representing the species' optima in 2-dimensional space in the ordination: PARA SEL (Paracynictis selousi); PHIL MON (Philantomba monticola); CEPH NAT (Cephalophus natalensis); HIPP AMP (Hippopotamus amphibius); CIVE CIV (Civettictis civetta); POECALB (Poecilogale albimucha); OURE OUR (Ourebia ourebi); NEOT MOS (Neotragus moschatus); MELL CAP (Mellivora capensis); LYCA PJC (Lycaon pictus); PANT PAR (Panthera pardus); HIPPO NI (Hippotragus niger); DICE BIC (Diceros bicornis); ACIN JUB (Acinonyx jubatus); LOXO AFR (Loxodonta africana); DAMA LUN (Dama dama lunatus); FELI SER (Felis serval); PROT CRI (Proteles cristatus); FELI LYB (Felis lybica); HIPP EQU (Hippotagus equinus); HYAE BRU (Hyaena brunnea); FELI NIG (Felis nigripes).
Quantifying complementarity using turnover and similarity measures.

Measurement of species turnover.

Table 3.7 shows the turnover values using the Harrison et al. (1992) modified measure (equation 8) (and for comparison the original measure of Whittaker (1960), equation 7) for birds, mammals and plants across fourteen east-west gradients of increasing altitude (basically the first axis extracted by CCA). Turnover was low for both birds and mammals across all fourteen gradients. Although there was greater variation between transects in the turnover values for plants, turnover was on average three times greater for plants than for birds or mammals. Only on one transect was plant turnover less than bird species turnover. Mean turnover values for plants and birds were significantly different ($t_{\text{paired}} = 6.33$, $P<0.01$, $n = 14$) and also for plants and mammals ($t_{\text{paired}} = 5.93$, $P<0.01$, $n = 14$). However turnover values for birds and mammals were not significantly different ($t_{\text{paired}} = 1.10$, $P = 0.292$, $n = 14$).

For the purposes of reserve selection it is the level of turnover at the scale of the reserve area that is important. However the ‘sites’ used in this study were larger areas than could realistically be designated reserves, so to test the effect of scale on the turnover values, beta diversity was calculated at both a finer and a coarser level for the mammal data (since these data originally existed in eighth degree resolution). Turnover values were 3.39 (S.E. 0.446, n=5) at eighth degree resolution, 4.82 (S.E. 0.387, n=14) at quarter degree resolution and 6.30 (S.E. 0.381, n=3) at half degree resolution. When all plant subspecies and varieties were lumped there were 5738 species recorded from KwaZulu Natal (compared to 6111 species, subspecies and varieties). The value for mean beta diversity across the fourteen transects was 14.2 (S.E. 1.43) when turnover was calculated with subspecies and varieties treated separately and 13.9 (S.E. 1.40) when turnover was measured at the species level indicating that the high turnover recorded for plants was not a result of the lower taxonomic level used compared to birds and mammals.

Turnover was also measured for twelve north-south transects for birds, mammals and plants (Table 3.8). As with the east-west turnover values, plant turnover was
Table 3.7. Species turnover for birds, mammals & plants along fourteen east-west transects in KwaZulu Natal, measured using two alternative methods (Harrison et al. 1992 & Whittaker 1960 - see text for explanation).

<table>
<thead>
<tr>
<th>E-W TURNOVER</th>
<th>BIRDS</th>
<th>MAMMALS</th>
<th>PLANTS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HARRISON</td>
<td>WHITTAKER</td>
<td>HARRISON</td>
</tr>
<tr>
<td>TRANSECT 1</td>
<td>5.06</td>
<td>9.58</td>
<td>6.04</td>
</tr>
<tr>
<td>TRANSECT 2</td>
<td>3.78</td>
<td>8.31</td>
<td>5.04</td>
</tr>
<tr>
<td>TRANSECT 3</td>
<td>5.27</td>
<td>12.6</td>
<td>3.47</td>
</tr>
<tr>
<td>TRANSECT 4</td>
<td>3.71</td>
<td>10.7</td>
<td>4.90</td>
</tr>
<tr>
<td>TRANSECT 5</td>
<td>4.84</td>
<td>10.6</td>
<td>4.17</td>
</tr>
<tr>
<td>TRANSECT 6</td>
<td>2.02</td>
<td>9.62</td>
<td>3.54</td>
</tr>
<tr>
<td>TRANSECT 7</td>
<td>4.41</td>
<td>11.4</td>
<td>4.80</td>
</tr>
<tr>
<td>TRANSECT 8</td>
<td>4.39</td>
<td>10.1</td>
<td>8.08</td>
</tr>
<tr>
<td>TRANSECT 9</td>
<td>4.14</td>
<td>12.1</td>
<td>2.03</td>
</tr>
<tr>
<td>TRANSECT 10</td>
<td>3.41</td>
<td>12.2</td>
<td>4.81</td>
</tr>
<tr>
<td>TRANSECT 11</td>
<td>2.56</td>
<td>12.3</td>
<td>5.21</td>
</tr>
<tr>
<td>TRANSECT 12</td>
<td>7.72</td>
<td>16.0</td>
<td>6.12</td>
</tr>
<tr>
<td>TRANSECT 13</td>
<td>5.42</td>
<td>14.0</td>
<td>5.44</td>
</tr>
<tr>
<td>TRANSECT 14</td>
<td>3.70</td>
<td>14.5</td>
<td>3.76</td>
</tr>
<tr>
<td>MEAN</td>
<td>4.32</td>
<td>11.7</td>
<td>4.82</td>
</tr>
<tr>
<td>S.E.</td>
<td>0.369</td>
<td>0.566</td>
<td>0.387</td>
</tr>
</tbody>
</table>
significantly higher than both bird and mammal turnover ($t_{\text{paired}} = 6.28$ and 7.15 respectively, $P<0.01, n = 12$) but there was no difference between mammal and bird turnover ($t_{\text{paired}} = 0.01, P = 0.992, n = 12$). There was not a significant difference between turnover values for east - west transects compared to north - south transects.

Table 3.8. Species turnover values for birds, mammals and plants measured along twelve north-south transects in KwaZulu Natal using the beta diversity measure of Harrison et al. (1992).

<table>
<thead>
<tr>
<th>HARRISON MEASURE OF TURNOVER</th>
<th>BIRDS</th>
<th>MAMMALS</th>
<th>PLANTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRANSECT 1</td>
<td>6.08</td>
<td>3.56</td>
<td>7.48</td>
</tr>
<tr>
<td>TRANSECT 2</td>
<td>3.34</td>
<td>4.44</td>
<td>13.1</td>
</tr>
<tr>
<td>TRANSECT 3</td>
<td>2.09</td>
<td>3.98</td>
<td>7.66</td>
</tr>
<tr>
<td>TRANSECT 4</td>
<td>3.37</td>
<td>4.87</td>
<td>9.06</td>
</tr>
<tr>
<td>TRANSECT 5</td>
<td>2.54</td>
<td>1.54</td>
<td>11.3</td>
</tr>
<tr>
<td>TRANSECT 6</td>
<td>4.49</td>
<td>7.69</td>
<td>12.0</td>
</tr>
<tr>
<td>TRANSECT 7</td>
<td>1.76</td>
<td>4.92</td>
<td>8.18</td>
</tr>
<tr>
<td>TRANSECT 8</td>
<td>6.62</td>
<td>2.11</td>
<td>12.4</td>
</tr>
<tr>
<td>TRANSECT 9</td>
<td>8.62</td>
<td>2.50</td>
<td>10.4</td>
</tr>
<tr>
<td>TRANSECT 10</td>
<td>3.67</td>
<td>4.15</td>
<td>12.5</td>
</tr>
<tr>
<td>TRANSECT 11</td>
<td>2.73</td>
<td>3.57</td>
<td>11.0</td>
</tr>
<tr>
<td>TRANSECT 12</td>
<td>2.49</td>
<td>4.37</td>
<td>19.2</td>
</tr>
<tr>
<td>MEAN</td>
<td>3.98</td>
<td>3.98</td>
<td>11.2</td>
</tr>
<tr>
<td>S.E.</td>
<td>0.606</td>
<td>0.459</td>
<td>0.921</td>
</tr>
</tbody>
</table>

To test the prediction that high turnover could be caused by habitat specificity along environmental or habitat gradients, turnover was measured along the 14 east - west transects for habitat specialist passerines and generalist passerines separately. Turnover is expected to be greater for the specialist species than the more widespread generalist ones. Mean turnover for specialist species was 6.81 (S.E. 0.719) and for generalist species was 2.59 (S.E. 0.229) (Table 3.9). Turnover was significantly higher for specialist species than for generalists ($t_{\text{paired}} = 6.07, P<0.01, n = 14$).
Table 3.9. Mean turnover (with S.E.) for fourteen east-west transects in KwaZulu Natal calculated for habitat specific and habitat generalist passerine species separately using the Harrison et al. (1992) measure of species turnover.

<table>
<thead>
<tr>
<th></th>
<th>MEAN TURNOVER</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>HABITAT SPECIALISTS</td>
<td>6.81</td>
<td>0.719</td>
</tr>
<tr>
<td>HABITAT GENERALISTS</td>
<td>2.59</td>
<td>0.229</td>
</tr>
</tbody>
</table>

Turnover between sites from different habitat types.

Complementarity was predicted to be greater between sites drawn from different habitat types compared to sites from the same habitat type. Pairwise similarity using the modified CC coefficient (equation 10) was measured for 336 pairs of sites for birds and plants. 168 pairs of sites consisted of two sites from the same habitat and 168 were pairs from different habitat types. Mean pairwise similarity for sites taken from the same habitat was 81.2% for birds and 20.9% for plants (Table 3.10). Mean pairwise similarity for site pairs taken from different habitat types was 69.6% for birds and 13.3% for plants (Table 3.10). Similarity between site pairs from the same habitat was significantly higher than for sites from different habitats for both birds and plants ($t_{(paired)} = 9.27$ and $t_{(paired)} = 6.78$ respectively, $P<0.01, n = 168$). However it should be noted that site pairs drawn from the same habitat type are more likely to be geographically close compared to sites from different habitats. Therefore the similarity of sites from the same habitat relative to those from different habitats could be caused by spatial autocorrelation effects.

Table 3.10. Mean pairwise similarity (CCJ (x100)) for birds and plants, between 168 pairs of sites from the same dominant habitat type, and 168 pairs of sites from different dominant habitat types.

<table>
<thead>
<tr>
<th></th>
<th>BIRDS</th>
<th>PLANTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIFFERENT HABITATS</td>
<td>69.6</td>
<td>13.3</td>
</tr>
<tr>
<td>SAME HABITAT</td>
<td>81.2</td>
<td>20.9</td>
</tr>
</tbody>
</table>

Identifying surrogates for wholesale diversity.

The taxa which showed low levels of turnover between sites were better represented when small subsets of sites were selected. Figure 3.16 shows the species accumulation curves for the three taxa in KwaZulu Natal as sites are added to a reserve network starting with the most species rich site for the taxon. The vertical line indicates the
Figure 3.16. Species accumulation curves as sites are added to a site network starting with the most species rich site and selecting subsequent sites in decreasing rank order (up to a total of 30 sites). The vertical line indicates the number of species that would be represented in the top eight sites.
percentage of species in each taxon that would be represented if the top five percent of sites in the region were selected for that particular taxon. More than 60% of bird and mammal species (low turnover taxa) are represented in the first, most species rich site and 89.2% of all mammals and 88.2% of all birds could be incorporated in eight top sites (five percent of the region) (Figure 3.16). In contrast only 18.4% of plants occurred in the top scoring site for the taxon (Figure 3.16). Even after the top five percent of sites had been selected only 62.5% of plant species were represented (Figure 3.16). In order to represent a similar percentage of plant species to the percentage of birds and mammals represented in just one site would require ten times the number of sites.

As shown in the previous chapter, more than 80% of the threatened bird species occurred in bird hotspots and 87% of threatened mammal species occurred in mammal hotspots (Figure 3.17). In contrast, plant hotspots only encompassed 38% of the threatened plants (Figure 3.17). Similarly the percentage of threatened birds and mammals represented in hotspots for other taxa is high (more than 65% of threatened species occurred in hotspots for other taxa) and much lower (less than 30%) for plants (Figure 3.17). Six percent of threatened plant species occurred only in sites with below average species richness, whereas none of the threatened bird species was restricted to below average diversity sites.

**How do richness and spatial turnover patterns relate to the ability of sites at representing regional diversity?**

Table 3.11 shows the percentage of species for each taxon that are represented in site networks (eight sites) selected using potential surrogate measures. Both birds and mammals were well represented in hotspots (top five percent sites) for other taxa, with more than 80% of bird species occurring in hotspots for plants and in hotspots for mammals and 78.2 and 80% of mammals occurring in plant and bird hotspots respectively. In contrast plants were poorly represented in sites selected on the basis of other taxa - only an average of 37% of plants occurred in the bird and mammal hotspots. Random sites selection picked up an average of 25% of plant species compared to top sites for plants which incorporated 63% of all species.
Figure 3.17. Percentage of threatened species represented in the top eight sites for species richness in KwaZulu Natal, for each of the taxa, plants, birds & mammals.
Table 3.11. Percentage of species represented in the top eight sites selected using alternative diversity criteria.

<table>
<thead>
<tr>
<th></th>
<th>TOP BIRD SITES</th>
<th>TOP MAMMAL SITES</th>
<th>TOP PLANT SITES</th>
<th>TOP ENDEMIC BIRD SITES</th>
<th>TOP ENDEMIC GRASS SITES</th>
<th>TOP ENDEMIC SPECIFIC HABITAT PASSEERINE SITES</th>
<th>RANDOM NETWORK</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIRDS</td>
<td>88.2</td>
<td>82.8</td>
<td>86.7</td>
<td>60.9</td>
<td>59.3</td>
<td>81.4</td>
<td>75.4</td>
</tr>
<tr>
<td>MAMMALS</td>
<td>80.0</td>
<td>89.2</td>
<td>78.5</td>
<td>63.1</td>
<td>49.2</td>
<td>73.8</td>
<td>72.3</td>
</tr>
<tr>
<td>PLANTS</td>
<td>44.7</td>
<td>28.3</td>
<td>62.5</td>
<td>36.1</td>
<td>38.6</td>
<td>47.8</td>
<td>25.4</td>
</tr>
</tbody>
</table>

Habitat specific bird species were good at representing bird and mammal species richness (81 and 74% respectively) and were an improvement on bird and mammal richness as surrogates for plant species richness (48% plants). In contrast endemic birds and grasses were poor at representing bird and mammal richness (worse than the mean random selection) and were poor at representing plant species richness (36 and 39% respectively). Six different site networks incorporating all habitat types in KwaZulu Natal were an improvement on random selections of sites (although only just in the case of bird species). It should be noted that these comparisons are not independent of one another.

DISCUSSION.

A widely agreed aim of protected areas is to maintain a representative selection of the existing biotic diversity (Austin & Margules 1986, Ratcliffe 1977, Usher 1986, Vane-Wright 1996). However it cannot be assumed that it is possible to select a small subset of areas that will encompass the majority of biotic diversity across a range of different taxa. Demonstration that high diversity areas for different taxa coincide at a particular scale does not necessarily indicate that regional diversity for a variety of taxa can be adequately represented by selecting a small number of diverse sites. Many of the correlations in species richness across different taxa are scale dependent (Murphy & Wilcox 1986). There are also examples, at a variety of scales, where high diversity areas for different taxa do not coincide (eg. Siegfried & Crowe 1983, Murphy & Wilcox 1986, Kremen 1992, Prendergast et al. 1993). Systematic approaches to the selection of priority areas for conservation therefore need to establish the extent to which multiple
A second problem is that such evaluations will have to be made using a small subset of taxa since exhaustive inventories will not be available in the near future (Kremen 1994). It is therefore important to identify those elements of biodiversity which most accurately reflect the distribution patterns of a wide range of other elements. Most commonly these indicator elements have been well studied, conspicuous vertebrates like birds, and large mammals but these may not be the best indicator groups to use (Murphy & Wilcox 1986, Sutton & Collins 1991).

Is there concordance in the distribution of diversity across different taxa?
The extent to which a set of species rich sites for one taxon will also represent the diversity of other taxa depends partly on the degree to which the different taxa show similar patterns of species richness in a region. If areas of high diversity for two taxa are broadly similar it might be expected that selecting a set of species rich sites for one of the taxa will result in the conservation of a large proportion of the species in the second taxon.

Bird species richness in KwaZulu Natal was correlated with plant species richness and with mammal diversity. However there was no significant correlation between the mammal and plant species richness of sites. A number of other studies have also found positive correlations between areas of high species richness for different taxa. For example, species richness of the plant family Proteaceae is significantly correlated with the other major families and genera of Fynbos plants in the Cape Floristic Region, Southern Africa (Rebelo & Siegfried 1990). There are also significant rank correlations between the number of bird, plant, reptile and mammal species on islands in the Gulf of California and between birds, mammals and plants in canyons in San Diego, California (Ryti 1992). However these correlations may not be significant once the problem of spatial autocorrelation is taken into account (Legendre 1993). Values of species richness for birds, mammals and plants in KwaZulu Natal were all positively, spatially autocorrelated. This violates the statistical assumption that observations are pairwise...
independent, so in statistically testing the significance of any correlation between two variables, like bird and mammal species richness, the degrees of freedom will be less than the sample size suggests (Legendre 1993). The significant correlations between species richness for the different taxa may therefore be due to autocorrelation in the data and this may also be true for other studies that have demonstrated a 'significant' correlation between species richness for different taxa.

The strong degree of positive spatial autocorrelation observed in the KwaZulu Natal data is not unexpected. Species composition at any particular locality is likely to be influenced by the species composition of surrounding sites as a result of contagious biotic processes like migration, and at a finer scale by processes like reproduction, mortality, parasitism, predator-prey interactions, mutualisms and food availability etc. (Legendre 1993). Similarly physical variables are not uniformly or randomly distributed, but form patches and gradients and these in turn induce the formation of similar spatial responses in biological systems (Legendre 1993). An alternative cause of spatial autocorrelation could result from the manner in which the data were collected. If a species has been recorded from a particular square observers may expect the species also to exist in adjacent squares and may spend more time searching for the species in nearby squares than in more distant sites. As a result the species is more likely to be recorded in these adjacent squares than in distant ones even though it might still occur in these. If the same observers tend to visit adjacent or nearby squares which is likely then the species observed from these squares might reflect the ones which the observer is most familiar with or which he specialises in. This is another way in which positive spatial autocorrelation could arise in the data set.

Determining that spatial autocorrelation exists in a data set using for example the Moran coefficient does not allow the different causal hypotheses to be distinguished between. However, it is possible to modify the statistical method to accommodate spatial autocorrelation when it has been observed to exist in a data set. For example, Clifford et al. (1989) propose a method for testing the significance of the correlation coefficient between two spatial processes by estimating a modified number of degrees of freedom.
An alternative method to reduce the effects of spatial autocorrelation might be to restrict analysis to non-adjacent squares. While there could still be spatial autocorrelation between non-adjacent squares, the phenomenon will be greatest between adjacent squares. When correlation coefficients were calculated for the KwaZulu Natal data with adjacent squares excluded from the analyses, some of the correlations (in two of the four data sets) between bird and mammal, and between bird and plant species richness were not significant (whereas there were significant positive correlations in the complete data set). However some significant correlations did remain, after removing adjacent squares so it would appear that there is a degree of concordance between plant and bird richness and bird and mammal richness despite the existence of strong spatial autocorrelation.

**Patterns of diversity and complementarity and their environmental correlates.**

Nonetheless, patterns of species distribution for the different taxa in KwaZulu Natal were related to the same major gradients of environmental variation. These were a gradient from low to high minimum July temperatures which split species occurring in the warmer eastern zone from species confined to high altitude western areas, and a secondary gradient based on mean annual precipitation and latitude, separating species occurring in wet coastal evergreen forests and high altitude evergreen forests from those occurring in the dry interior upland grasslands and bushveld savanna areas.

Maximum species richness for birds was along the east coast and in north-east bushveld, for mammals in the northern dry uplands and bushveld and for plants in wet coastal and wet, high altitude areas. So although broad patterns of complementarity for the different taxa were the same, gradients of species richness did show some intra regional differences.

Within taxa, habitat specialists and threatened species were not restricted to any one environmental gradient, whereas endemic species were almost exclusively upland species, restricted to the Drakensberg Mountains in the west of KwaZulu Natal.

As a group, threatened birds were not confined to any particular environmental gradient.
However, the majority of individual species were not widespread, that is the distributions of individual threatened species tended to be strongly correlated with a particular environmental type but unlike endemic species, threatened species were found in a variety of habitats. However, all the highland and upland, threatened passerine species were globally threatened species. The Drakensberg Serin (*Serinus symonsi*) is endemic and occurs only in montane scrub and grassland in the Drakensberg range. The Yellowbreasted Pipit (*Anthus chloris*) is also an endemic confined to high altitude grassland in the Drakensberg. The Orangebreasted Rockjumper (*Chaetops aurantius*) is an endemic found on mountain slopes and high altitude grasslands in the Drakensberg. The Buffstreaked Chat (*Oenanthe bifasciata*) is endemic but has a wider distribution than the last three species, occurring on upland grassy hillslopes. The Bush Blackcap (*Lioptilus nigricapillus*) is endemic and is found in montane evergreen forests and mistbelt. The Blue Swallow (*Hirundo atrocaerulea*) occurs in eastern and south-eastern Africa. It occurs in upland grassland areas bordering mistbelt forest, is highly localised and endangered due to habitat loss (Sinclair et al. 1993). The Chorister Robin (*Cossypha dichroa*) is endemic and frequents mistbelt evergreen forests, wintering in coastal forests.

Those bird species that were only in the South African Red Data Book were predominantly coastal lowland species and were in many cases species with distributions at the south east edge of their range in South Africa. The Pied Mannikin (*Spermestes fringilloides*) occurs locally throughout the Afrotropical Region but in southern Africa only in Zimbabwe and Mozambique with its range extending south in a narrow forest strip along the extreme east coast of KwaZulu Natal. The Wattle-eyed Flycatcher (*Platysteira peltata*) and African Broadbill (*Smithornis capensis*) have similar distributions. The Yellow White-eye (*Zosterops senegalensis*) is Afrotropical extending into eastern bushveld areas of KwaZulu Natal at the southern edge of its range. The Short-tailed Pipit (*Anthus brachyurus*) occurs in central and south-eastern Africa. In southern Africa, the subspecies, *Anthus brachyurus brachyurus*, is confined to coastal and midland districts of KwaZulu Natal (Howard & Moore 1980). As with birds threatened mammals were not confined to any one environmental gradient or habitat type.
As expected passerine species classified as habitat specific on the basis of field guide descriptions generally had greater ordination scores than non habitat specific species. Like threatened birds, they also had representatives in all the major areas delimited by the environmental areas so are better indicators of the major gradients than the endemic species.

In contrast, endemic birds and grasses were predominantly clustered in western, high altitude areas in the Drakensberg in highland/montane/drier upland areas or in mistbelt forests, for example, the Drakensberg Serin (*Serinus symonsi*) which occurs in scrub and montane grassland, the Melodious Lark (*Mirafra cheniana*) which frequents upland grassland and the Chorister Robin (*Cossypha dichroa*) which is a mistbelt species. These areas were the poorest for overall bird species richness. Endemic species richness therefore shows exactly the opposite trend to overall bird species richness in KwaZulu Natal. For passerine birds only six endemics (nine percent) have their centres of distribution in bushveld, and four species (six percent) in the coastal forest belt, the areas supporting the highest bird species richness. This means that 77% of endemic terrestrial passerines occur predominantly in the lower temperature, high altitude west of KwaZulu Natal, where species richness is lowest. Similar patterns emerge for nonpasserine bird species. No nonpasserine endemics are predominantly bushveld species, three are coastal lowland species and eleven species are concentrated in the western upland, highland and montane areas.

Like endemic birds, endemic grasses are also strongly correlated with the cool, western high altitude areas. Seventy-five percent of all grassland endemics occur in areas of below average mean minimum July temperatures. A large proportion of these species are montane. Unlike bird endemics, however more than half the endemic Poaceae occur in areas of above average species richness.

For all the taxa analysed, complementarity will be greatest between sites from the high altitude Drakensberg in the west and the warmer coastal lowland or bushveld areas in the east. Turnover will also be high between high precipitation areas like coastal lowland
forests or mistbelt areas and drier interior uplands or bushveld. So patterns of complementarity within KwaZulu Natal would seem to be broadly similar for the different taxa. Plants show a different richness gradient to mammals and birds however.

When the most species rich sites for different taxa are compared, there was little correspondence between the top sites selected for each of the taxa in KwaZulu Natal. Three of the eight top sites were common to birds and mammals, three sites were common to birds and plants and only one site was common to mammals and plants. ‘Hotspot’ overlap is also low for a number of British taxa - between 0 and 60% (mean overlap 20.4%) in the top 5% of sites selected for bird, butterfly, dragonfly, liverwort and aquatic plant species richness (Prendergast et al. 1993). So, despite broad correlations in the patterns of richness for different taxa, at a finer scale there would appear to be little concordance between the most species rich sites for different taxa.

**What makes a good surrogate for wholesale diversity?**

Irrespective of whether there is concordance in alpha (site level) richness patterns, if the taxon used for site selection does not exhibit significant turnover between sites, whereas other taxa in the region do, the selected sites are not likely to represent the high turnover taxa very well. The highest diversity sites of the low turnover taxon may all be in a uniform habitat type and might not sample the full range of environmental variation, and therefore will miss many of the high turnover species in the region. Bird hotspots in KwaZulu Natal were all in a narrow belt along the eastern coast of the province and encompassed only three (coastal hinterland, coastal lowlands and bushveld) of the eight main habitat types designated for KwaZulu Natal. These sites selected using bird species (low turnover) richness as the selection criteria encompassed only 44.7% of plant species (high turnover) but 88.2% of birds and 80% of mammals. Mammal hotspots were concentrated in the north-east and did not include any montane, highland or coastal lowland habitat. Mammal species, which did not show a significant correlation in the distribution of alpha diversity with plants, were even worse as a surrogate for plant diversity. Sites selected on the basis of mammal species richness incorporated a mere 28.3% of plant species, but represented 89.2% of mammals and 82.8% birds. Plant
hotspots, in contrast, were much more dispersed across the province and included samples of all eight habitat types, from montane in the west to coastal lowland in the east. Low turnover taxa will be well represented in the hotspots for other taxa provided that the hotspots contain one or a few sites of above average alpha diversity for the taxon. Sites selected using plant richness encompassed 86.7% of bird species and 78.5% of mammals.

In Britain where turnover appears to be low for birds, butterflies and dragonflies the top 5% of sites scored for species richness of these taxa incorporated between 70 and 100% (mean 84.7%) of the total diversity of the other taxa (Prendergast et al. 1993). For example bird hotspots incorporated 87% of all bird species, 100% of butterfly species and 92% of dragonfly species (Prendergast et al. 1993).

It is therefore possible to identify certain traits which good indicator groups will possess. Good indicator taxa are predicted to have a broad geographic range in the region, but with individual species specialised, habitat specific and relatively sedentary (ie. high beta diversity) (Brown 1991, Sutton & Collins 1991, Pearson & Cassola 1992). The taxon should also be taxonomically and ecologically highly diversified, well known and easily observed with a stable taxonomy (Brown 1991, Sutton & Collins 1991, Pearson & Cassola 1992). Finally, patterns of richness in the indicator group should be highly correlated with those of other taxa in the region (Brown 1991, Sutton & Collins 1991, Pearson & Cassola 1992).

By definition target taxa should either correlate with patterns of environmental heterogeneity, or with the distribution patterns of species in other taxonomic groups (Brown 1991, Kremen 1994). Thus effective indicator groups are likely to be what Cody (1986) terms ‘beta rarities’. These are species that occupy a narrow range of environmental variation in a region i.e. are habitat specific. Kremen (1994) argues that clades which have undergone an evolutionary radiation in the region, and are therefore species rich with a large number of endemic species, are likely to be useful general surrogates or focal taxa because they will be good indicators of the environmental
variation and complexity in a region and they are likely to show congruence with other groups.

However not all regions will have taxa with large numbers of endemic species which could be used for site discrimination. For example South America has 440 bird species with ranges less than 50,000km$^2$, representing 25% of its total avifauna (Terborgh & Winter 1983). In contrast there are only 8 species (2% of total avifauna) with ranges less than 50,000km$^2$ in North America (Terborgh & Winter 1983).

More generally then, taxa in which individual species show a high degree of turnover between sites within a region (species that occupy a narrow range of environmental variation and are localised), but where the taxon as a whole has a wide distribution within the region, are predicted to be information rich. Selecting sites to represent these species will require sites that span the whole region and which sample the full range of habitat types. This is a similar approach to using habitat or landscape types as selection surrogates (e.g. Mackey et al. 1989) but does not require a decision about the appropriate scale to measure habitat diversity since the taxa will define this.

High beta diversity has been attributed to low mobility and habitat or environmental specificity (Cody 1986, Bond 1989, Cowling et al. 1989, Harrison et al. 1992). In addition smaller species are likely to show greater turnover since habitats exhibit fractal dimensions so smaller species will respond to a finer level of environmental variation (Murphy & Wilcox 1986, Brown 1991, Sutton & Collins 1991). Plants are more likely to be beta rarities because they are sedentary and often highly habitat specific (Major 1988). Many plant taxa are restricted to a specific substrate, such as patches of serpentine soil in California (Cody 1986, Major 1988). In the North-East Transvaal Escarpment, 81 plant species and infra specific taxa are endemic or near endemic to quartzite and related rock types and 32 species endemic to dolomite rocks and these two groups are nearly always mutually exclusive as to substrate preference (Matthews et al. 1993). Beta diversity has been shown to be high in the Fynbos flora of the Cape Floristic Region in South Africa (Kruger & Taylor 1979, Bond 1983, Cowling et al. 1989). A five
kilometre transect along a uniform coastal plain, across a fertility gradient had 3.9 complete community changes giving 4.9 community types (Cowling et al. 1989).

Certain invertebrate taxa and amphibians are also potentially good surrogates (Murphy & Wilcox 1986, Brown 1991, Sutton & Collins 1991, Daniels 1992). According to Daniels (1992) nearly half of the amphibian fauna of the Western Ghats in India are highly localised, being known from only a single one degree latitude band. Seventy-six percent of species are endemic to the region. In contrast birds show only 5% endemism and are much more widespread across the region (Daniels 1992). A poorer ability to disperse compared to birds and a dependency on certain environmental conditions probably explains the differences (Daniels 1992).

As well as possessing good indicator properties, beta rarities themselves are very important to conserve if an aim of conservation is to maximise ecological type diversity (Bond 1989). Conserving beta rarities will require sites that span environmental and habitat gradients in the region and will therefore be likely to represent a wide variety of the ecological diversity exhibited by species in the region.

So plants due to their lower mobility and greater substrate specificity are predicted to be better surrogates for biodiversity than birds or mammals and this was the case for taxa in KwaZulu Natal. Endemic species might also be expected to perform well as surrogates due to their restricted geographical range. However they will only be successful surrogates if endemic taxa as a whole have a widespread distribution throughout the region. In KwaZulu Natal endemic birds and grasses were largely concentrated in the Drakensberg highland region, and top sites for endemic species richness were all restricted to this area. Both plants and birds do show other centres of endemism (in the north-east bushveld and south-east coastal zone) which when considered together would pick up more turnover, but this areas are not as rich in endemic species as the Drakensberg.

Within birds, species that are habitat specific might generally be a better surrogate than
endemic species, since they are less likely to be restricted to one environmental type. When sites were selected in KwaZulu Natal using richness of habitat specific passerines, the resultant network was an improvement over the use of simple bird species richness at representing plant species and was markedly better than endemic bird richness at representing all three taxa in KwaZulu Natal. Higher taxon richness for birds was a good surrogate for bird species richness and has been demonstrated to be a good correlate of species level diversity for a number of other taxa, in a variety of regions and scales (Williams & Gaston 1994).

Selecting sites that together represented all major habitat types in KwaZulu Natal was predicted to be a good surrogate, particularly for plant species richness. Pressey & Nicholls (1989b) propose that when species distribution data are incomplete, selecting sites that together sample a given proportion of each environment type in the region should represent most species. Sites selected from different habitat types were significantly more complementary than sites from the same habitat type. However six sets of environmentally representative sites were on average poor at representing plants (though better than random) and birds (although better than endemic bird richness). The reason for the poor performance of environmental representativeness as a surrogate was that although they maximised complementarity between sites, in many cases the sites were of low species richness. Two sites showing no complementarity (ie identical species composition) but with twice as many species as two sites showing 100% complementarity would contain the same overall number of species. Using environmental representativeness ensures complementarity but cannot guarantee to pick up high richness sites. In the same way concentrating on species richness may not maximise complementarity and in some cases will be a poor surrogate for regional species richness. One solution is to combine information on spatial turnover and species richness patterns for the taxa that show greatest complementarity in a region. Algorithms that attempt to maximise complementarity and richness will be discussed in the following chapters.
How well can regional diversity be represented in small number of sites?
The amount of regional diversity that can be represented in a subset of sites - for example, reserves or protected areas - will depend of the extent of turnover that occurs at the site scale. If turnover is low, then a large proportion of the regional species richness will occur in the most diverse sites. In KwaZulu Natal, turnover between 'sites' for birds and mammals is very low and consequently more than 88% of the regional species richness for these taxa is represented in the top 5% most species rich sites. Likewise in Britain, where turnover for a variety of taxa (including birds, butterflies and dragonflies) has been shown to be of a similar magnitude to that for birds and mammals in KwaZulu Natal, between 87 and 92% of breeding bird, butterfly and dragonfly species occur in the top 5% of species rich sites for each taxon (Harrison et al. 1992, Prendergast et al. 1993).

In contrast plant species in KwaZulu Natal show much greater turnover between sites at the scale measured (on average three times the turnover of birds and mammals) and as a result a smaller amount of the total regional diversity (62.5%) occurs in the top 5% of sites scored for plant richness.

So taxa that show a large degree of turnover at the site level will be less easily represented in a small subset of sites. When turnover for several taxa on islands in the Gulf of California and canyons in San Diego was measured using a nestedness measure, turnover proved to be lower for birds compared to plants (and reptiles and mammals) (Ryti 1992). All bird species in the region could be represented in just two sites in both regions, whereas to represent all plant species required 10 and 13 sites respectively (Ryti 1992). A suite of sites selected to protect all plant species would therefore be larger than the set of sites required to conserve all bird species.

It cannot be assumed that a set of areas that are good for one taxon will also represent a similar proportion of the diversity in other groups. A useful approach for selecting priority areas in any region is to base selection on taxa that show high levels of turnover at the level at which sites are to be selected. This will require sites that sample the full
range of environmental and habitat variation in the region and it may be necessary to
select sites specifically to include threatened or rare species (or other aspects of diversity)
since these will not necessarily be represented in the most species rich sites. In addition
sites selected using a high turnover taxon are likely to maximise representation of other
high turnover taxa because a variety of habitat and environmental diversity will be
included in the selected sites, as well as encompassing the more easily conserved low
turnover groups, like birds and mammals in KwaZulu Natal. If sites are selected in
KwaZulu Natal on the basis of plant diversity there is virtually no compromise in the
amount of bird and mammal species that would be represented, at the same time the
number of plant species conserved is maximised. Conserving beta rarities will require
selecting sites along environmental gradients and this may also provide species with
dispersal routes should climatic conditions change in the future causing habitats to
‘migrate’ (Rebelo & Siegfried 1990).

While taxa with low turnover can easily be represented in a small number of sites (since
one site may contain most of the species in the region), this site will not necessarily
contain viable populations of the species or represent a variety of genetic diversity, so
that additional sites or the surrounding habitat matrix might be required to ensure the
persistence of the species.

Selecting sites on the basis of the current distribution of diversity assumes a snapshot
interpretation of species distributions (Rebelo & Siegfried 1990). Species are assumed
to be stable and not subject to extinction or climatic tracking in any site. Selecting sites
on the basis of beta rarities will however be an improvement on the use of general species
richness since the sites selected should sample a wide range of environmental variation
and so act as an ecological buffer should conditions change (Rebelo & Siegfried 1990).
CONCLUSIONS.

- There were significant positive correlations between the bird and mammal species richness of sites and also between bird and plant species richness.
- However, endemic species richness was concentrated in areas of low overall species richness.
- Patterns of spatial turnover were related to the same environmental gradients for the different taxa.
- Spatial turnover was low for birds and mammals, but higher for plant species.
- Within passerine birds, habitat specific species exhibit greater turnover compared to generalist species.
- High turnover groups require a larger number of sites to represent the same proportion of their regional diversity compared to low turnover groups.
- High turnover groups are the best surrogates to use for selecting sites to represent regional diversity because they result in the selection of sites that together sample a wider range of community types compared to low turnover groups.
CHAPTER FOUR.

A COMPARISON OF THE CONSEQUENCES OF USING TWO ALTERNATIVE SITE SELECTION METHODS TO REPRESENT THE DIVERSITY OF A REGION: SCORING VERSUS ITERATIVE SELECTION METHODS.

INTRODUCTION.

Once areas or potential reserves have been evaluated in terms of their biodiversity value it is then necessary to use this information to select a network of areas to represent the biodiversity of a region. In the past selection of reserves has been based on scoring or ranking procedures (Wright 1977, Usher 1986, Smith & Theberge 1987). Areas are given some kind of score which is commonly the species richness value but could be a more complex score based on a variety of factors such as rarity, irreplaceability, representativeness, size etc (Gehlback 1975, Perring and Farrell 1977, Margules & Usher 1981, Van der Ploeg 1986, Given and Norton 1993). The top ranking sites are then selected for the reserve network. Multivariate scores were developed since it is widely recognised that biodiversity involves more than just simple species richness (see Chapter 2). However scores combining a variety of different criteria have been criticised on the grounds that the weightings that have to be given to each criterion within the score are highly subjective (Götzmark et al. 1986). For example how do you balance the importance of rare species in a site over threatened species or overall species richness? Another problem with complex scores is that since the site value is encapsulated in a single number it is not obvious what components of biodiversity have contributed to a score. Two sites may have the same score but for very different reasons.

Simple scores based on just one criterion are more explicit but they cannot accommodate all facets of biodiversity. One way around this problem might be to score areas separately for a variety of biodiversity elements and then select either those areas that score highest for all elements or select the top scoring areas for each biodiversity element.
The main problem with site scores (and most particularly simple scores based on for example species richness) is that in most situations the top scoring sites will not maximise the representation of biodiversity in a region because they do not take account of complementarity between sites. As explained in the previous chapter if high diversity areas do not sample gradients of complementarity or turnover in a region then purely selecting the top sites for richness will not be a good way of representing the biological variation in that region. This is because all the top scoring sites will contain a similar complement of species, whereas lower scoring sites that are complementary with high scoring sites will have many species that are not found in high richness sites.

Ideally then, a site selection method needs to combine information on species richness and species turnover (complementarity). A network that represents the range of diversity in a region will contain high scoring sites within areas which contain complementary suites of species. One way in which this could be achieved would be to use a surrogate measure that reflects both the major patterns of richness and complementarity in the region, as discussed in the previous chapter.

An alternative method would be to use information on the distribution of species in the region and apply a numerical selection procedure that maximises both richness and complementarity. The idea of combining complementarity as well as richness is not new. Kirkpatrick (1983) first applied an algorithm that accounted for complementarity to the selection of areas for threatened species in Tasmania in an attempt to avoid the problem that when formulas were used to score sites some species were duplicated in several sites while others were not represented at all. Ackery and Vane-Wright (1984) used the concept of complementarity to identify the minimum set of sites to include at least one population of all 157 milkweed butterfly species in a process they called ‘critical fauna analysis’. A number of other algorithms that use information on site complementarity have now been developed (Margules et al. 1988, Pressey and Nicholls 1989a, Williams 1992, Kershaw et al. 1994, Csuti et al. in press). These iterative algorithms vary in detail. The simplest iterative algorithm, referred to here as a richness algorithm, operates in the following way. The first site to be selected is the one containing the most species.
The second site is the one which adds the highest number of unrepresented (i.e. new) species to the site network. This site may or may not be the second most species rich site in the region. A site with relatively few species might be highly complementary with the first sites selected and therefore contribute more new species than a higher richness site which is very similar in species composition to the first site. Subsequent sites are similarly selected to maximise the number of new, unrepresented species added to the network at each step. By selecting sites that contribute the most new species at each step the algorithm is ensuring that there is high complementarity in the species composition of selected sites, and at the same time is taking account of species richness. A site that contains species found in no other sites (and is therefore highly complementary with other sites) may not be selected if it is relatively poor in species, since a site which is less complementary with those already selected, but which is rich in species could add more new species to the network.

Since the number of sites that can be selected for conservation is likely to be limited by economic and other considerations a primary aim in reserve selection must be to represent all attributes (for example species) in the smallest number (or area) of sites possible, or to maximise the number of attributes represented in a predefined area of reserve. Iterative selection methods should incorporate more species per site network compared to the same number of sites selected by a scoring approach i.e. an iterative selection should be more efficient at representing species (Pressey & Nicholls 1989a). Iterative methods should also be able to incorporate all species in a smaller number of sites. Another potential advantage of using an iterative algorithm is that components of diversity that do not occur in the most species rich sites may be picked up early on in the site selection process if they occur in sites that are complementary with the richer sites.

In this chapter I compare the efficiency of selecting the top scoring sites compared to using an iterative richness algorithm for a variety of taxa showing different patterns of richness and spatial turnover at two contrasting scales. Additionally, for each taxonomic group, I look at whether site networks selected using an iterative species richness algorithm are better at representing those attributes of diversity that do not occur in the
most species rich sites for the taxon as whole. The effect of some sites already having been set aside for conservation (*ad hoc* reservations (Pressey 1994, Pressey & Tully 1994) on the subsequent inclusion of biodiversity is investigated for the antelope data.

In particular I investigate the following questions using distribution data for birds, mammals and plants in KwaZulu Natal and for Afrotropical antelopes:

1. How much more efficient is an iterative site selection at representing species richness compared to selection of the top-scoring sites and how does this relate to the patterns of species richness and spatial turnover in the region?
2. Are the same sites selected by a scoring method compared to an iterative method?
3. How is the efficiency of the site network affected if some sites have already been set aside for protection?
4. Are components of diversity that do not occur in the most species rich sites better represented by an iterative selection process?

These questions were used to test the following predictions:

If spatial turnover between sites is low for a given biodiversity attribute (e.g. species), then a selection of the top scoring sites for that attribute should encompass a large proportion of the regional diversity. Bird and mammal species in KwaZulu Natal show low turnover it is therefore predicted that top scoring sites for species richness will represent a large proportion of the regional species richness and may not be much less efficient than an iterative selection.

If spatial turnover between sites is high, and additionally species richness across sites is relatively homogeneous then there will be a large difference in the efficiency of a top scoring method of sites selection compared to an iterative one. In this case, top scoring sites are predicted to represent a smaller proportion of the regional richness compared to a situation of low spatial turnover, and an iterative selection method should be noticeably more efficient at representing the regional diversity. Site level species richness for Afrotropical antelopes was reasonably equable across sites, and given the large biogeographical area from which the data were drawn there is likely to be a high level of
spatial turnover between sites. An iterative site selection method is therefore expected to be more efficient compared to selection of top scoring sites at representing regional diversity, and this difference is predicted to be greater than for birds and mammals in KwaZulu Natal.

Where spatial turnover between sites is high, but species richness is highly heterogeneous, a top scoring method may perform as equally well as an iterative method. Plant species in KwaZulu Natal show a high level of complementarity between sites, but there is also a large amount of variation in the species richness of individual sites. For plant species an iterative selection method may not be much more efficient compared to a scoring method.

Biodiversity attributes that show different patterns of richness to overall species richness will not be represented by a method that selects the top scoring sites for species richness. In contrast, an iterative approach that maximises complementarity between the sites selected, is more likely to pick up those elements of diversity that show different patterns of distributions to the taxon as a whole. Endemic birds, threatened antelope and restricted-range antelope species all show different patterns of richness compared to general species richness. An iterative selection is therefore predicted to be substantially better at representing these species compared to selection of the top scoring sites for species richness for the whole taxon. For example, it is known that endemic bird species in KwaZulu Natal occur in sites that are poor in species (see Chapter 3). A selection of top scoring sites does not pick any sites that are rich in endemics. However the sites that are rich in endemics show high turnover in composition with the high richness east coast sites so it might be predicted that using an algorithm that takes complementarity into account will select an area from the west where the endemic species are concentrated.

METHODS.

Data.
The antelope data used were the same as in Chapter 2 and consisted of 99 species and subspecies in 249 sites in Subsaharan Africa. 41 of the species were classified as being
under some degree of threat, while the other 58 were considered 'not of concern'. Data on birds, mammals and plants in KwaZulu Natal are as analysed in Chapter 3.

**Selection of priority sites.**

Two alternative site networks were selected for each taxon separately (birds, mammals and plants in KwaZulu Natal and Afrotropical antelopes). The first consisted simply of the top sites scored for species richness. The second network was selected using an iterative richness algorithm.

WORLDMAP (Version 2.5, Williams 1992, see also Chapter 2) was used to select sites iteratively for the antelope data in the following way: the site with the highest species richness score was selected first. The scores for each remaining site were then recalculated with the species present in the selected site removed from all other sites where they occur. This means that the recalculated scores only take account of species not already included in the selected site network. The highest scoring site after the scores were recalculated was then the site adding the most new species, and this site was then selected. This iterative process was continued until all species were represented. Where there was a tie in the highest scores between sites, the site with the most species overall (represented plus unrepresented species) was selected. If there was still a tie at this point then a site was selected randomly from those sites with the highest score plus most species. For each selected site added to the list WORLDMAP calculates the cumulative percentage of species included in selected areas.

The KwaZulu Natal data set was too large to be processed using the WORLDMAP software available at the time of the analysis. Instead an iterative richness algorithm was written in FORTRAN and applied to the data. The algorithm was identical to the one used for the antelope data in WORLDMAP, except that a slightly different tie-breaking routine was used. Instead of selecting the sites with the most species overall and then a site at random, all ties were broken simply by selecting the first site in the sequence of those with the highest score. Sites were selected until all species had been represented.
To allow cross comparison the first twelve sites were selected by each method for the antelope species and the first eight sites for all KwaZulu Natal taxa (approximately five percent of the total area in each case). These sites were evaluated in terms of how well they represented overall species richness and also other diversity components. For the bird site networks, representation of all bird species, threatened species and endemic species was measured; for mammals, all species and threatened species; for plants, all species and threatened species; and for Afrotropical antelopes, all species, threatened species, restricted-range species and taxonomic diversity (as measured by dispersion weighted criterion in WORLDMAP). All these different groups are as defined in the previous two chapters. The selected sites networks were used to address the following questions:

**How much more efficient is an iterative site selection at representing species richness compared to selection of the top-scoring sites and how does this relate to the patterns of species richness and spatial turnover in the region?**

**Number of sites needed to represent all species.**

The different site selection methods (iterative versus scoring) were compared in terms of how efficiently they incorporated species. The number of areas needed to include all attributes can be expressed as an efficiency coefficient (Pressey and Nicholls 1989a):

\[
E = 1 - \left( \frac{X}{T} \right)
\]

where \(X\) is the number of sites needed to include all attributes (eg. species) and \(T\) is the total number of sites.

\(E\), the efficiency coefficient, can range from 0 to 1; the closer \(E\) is to 1, the greater the efficiency of the selected areas at representing all attributes. The coefficient enables easy comparison of different criteria for selecting reserve networks.
Numbers of species represented in priority networks selected by scoring and iterative methods.

Given that there will not be the resources to select sites to represent all the species in a region, a small subset of sites was selected for each taxon using scoring and iterative methods and evaluated in terms of the number of species that were represented. For each taxon five percent of the number of sites in the region were selected - this resulted in twelve sites for the Afrotropical antelopes and eight sites for birds, mammals and plants in KwaZulu Natal.

Are the same sites selected by a scoring method compared to an iterative method?
The first twelve sites selected by each method for the antelope and the first eight sites for each taxon (birds, mammals and plants) in KwaZulu Natal were compared to investigate differences between the location of site networks selected using scoring versus iterative methods.

How is the efficiency of the sites network affected if some sites have already been set aside for protection?
Where sites have already been set aside for conservation the requirement will be to select additional sites to complement the existing network. The pre-existence of sites may reduce the efficiency of a network so that more reserves will be needed in total than if sites could be selected from scratch. Obviously the extent to which an iterative selection procedure will be less efficient when sites have already been set aside will depend on the attributes contained within those sites. To examine this for the antelope data a varying number of sites were preassigned to the selection and the number of additional areas required to represent all species by complementarity was determined. A number of random selections of two, five, ten and fifteen sites were selected for both species rich and species poor sites. Species poor sites were defined as those scoring below the median value of four species for unweighted species richness, and species rich sites as those with four or more species. Once the preassigned sites had been excluded sites were selected iteratively until all species were represented, and the total number of sites needed was recorded. This was repeated for fifteen random selections of two, five and ten
preassigned sites and ten selections where fifteen sites were preselected. The mean number of sites required to incorporate all the species was then calculated for each. These were compared to the number of sites needed to represent all species when no sites were preselected. Fifteen such selections were made with no sites preassigned, and the mean number of sites needed calculated as for the other selections. Fifteen selections with no sites preassigned were made because if there were more than one top scoring site at any step in the iterative selection procedure (which happened frequently) then a site was selected at random. This meant that a number of different combinations of sites were possible, so the total needed to represent all species could potentially vary.

Are components of diversity that do not occur in the most species rich sites better represented by an iterative selection process?

Site networks (twelve sites for antelope and eight sites for the KwaZulu Natal taxa) selected by scoring and iterative procedures were evaluated in terms of their ability to incorporate different diversity components. For the antelope site networks, the number of threatened species and the percentage of total taxonomic diversity and restricted-range diversity that were incorporated in each network were measured. Taxonomic diversity (dispersion-weighted) and restricted-range diversity were calculated using WORLDMAP software as described in Chapter 2. For the bird data the number of endemic species and the number of threatened species in each network was measured and for plants and mammals the number of threatened species in each network was evaluated. The data for the southern African taxa were the same as that used in Chapters 2 and 3.

RESULTS.

How much more efficient is an iterative site selection at representing species richness compared to selection of the top-scoring sites and how does this relate to the patterns of species richness and spatial turnover in the region?

Number of sites needed to represent all species.

As expected an iterative approach was always more efficient than a scoring method, that is it required fewer sites to achieve representation of all the species (Table 4.1). Thirty-
seven out of a total 249 sites (14.9% of all sites) were required to include all antelope species when sites were selected iteratively. In contrast at least 216 (and up to 249) of the sites (87 - 100% of all sites) would be needed to include all antelope species if the top ranking sites were selected in sequence. Between them, the top 37 scoring sites contained only 68% of all species. The efficiency of the iterative selection was 0.851 compared to an average of 0.066 for top-ranking sites (the range in values is due to there being a number of sites with the same score for low score values - so 0.066 represents the mean efficiency for all possible sequences of tied sites which would result in complete representation) (Table 4.1).

Table 4.1. Number of sites needed to represent all species using two alternative sites selection methods (selection of top sites (scoring) and an iterative richness algorithm) & efficiency of iterative site selection methods compared to the selection of top scoring sites for birds, mammals and plants in KwaZulu Natal and Afrotropical antelopes.

<table>
<thead>
<tr>
<th>NUMBER OF SITES TO INCLUDE ALL:</th>
<th>BIRDS (570)</th>
<th>MAMMALS (65)</th>
<th>PLANTS (6111)</th>
<th>ANTELOPES (99)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITERATIVE</td>
<td>27</td>
<td>9</td>
<td>140</td>
<td>37</td>
</tr>
<tr>
<td>SCORING</td>
<td>143</td>
<td>122-127</td>
<td>164</td>
<td>216-249</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EFFICIENCY OF DIFFERENT NETWORKS</th>
<th>BIRDS</th>
<th>MAMMALS</th>
<th>PLANTS</th>
<th>ANTELOPES</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITERATIVE</td>
<td>0.837</td>
<td>0.946</td>
<td>0.157</td>
<td>0.851</td>
</tr>
<tr>
<td>SCORING</td>
<td>0.139</td>
<td>0.235-0.265</td>
<td>0.012</td>
<td>0-0.133</td>
</tr>
</tbody>
</table>

The difference in efficiency between iterative and scoring selections was also marked for birds and mammals in KwaZulu Natal (Table 4.1). All 570 bird species could be represented in just 27 sites (16% of the total area) whereas selection of top scoring sites required 143 sites (86% of the total) before all species had been represented. All 65 mammal species were represented in nine sites selected iteratively, compared to a minimum of 122 sites (and potentially 127 sites depending on the selection order of tied sites) when the species richness was used without reference to complementarity (Table 4.1). When sites were selected iteratively for plant species in KwaZulu Natal all 6111 species could be represented in 140 sites (84% of the total) (Table 4.1). Nonetheless an
iterative approach was more efficient at representing all species compared to a scoring selection. When top scoring sites were selected 164, sites were needed (99% of all sites) before all species were represented compared to 140 sites when an iterative algorithm was used (Table 4.1). So for plant species in KwaZulu Natal the difference in efficiency between a scoring and iterative selection method is less than for birds and mammals.

**Numbers of species represented in priority networks selected by scoring and iterative methods.**

The first twelve sites selected using the iterative algorithm represented 70 antelope species compared to only 53 species in the top ranking twelve sites for species richness (Table 4.2).

A network of eight sites selected iteratively incorporated 541 bird species (95%) compared to the top ranking eight sites which represented 503 species (88%) (Table 4.2). The top eight sites for mammal species richness represented 58 species (89%) whereas the first eight sites selected by the iterative algorithm incorporated 64 of the total 65 species (99%) (Table 4.2). Two of the first eight sites selected iteratively were of below average species richness for mammals. The top eight sites for plant species richness were very similar to the first eight sites selected iteratively in terms of the number of species represented. The top eight sites incorporated 3852 species compared to 3820 in the iterative selection (Table 4.2).

<table>
<thead>
<tr>
<th></th>
<th>ITERATIVE</th>
<th>SCORING</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIRDS</td>
<td>541 (95%)</td>
<td>503 (88%)</td>
</tr>
<tr>
<td>MAMMALS</td>
<td>64 (99%)</td>
<td>58 (89%)</td>
</tr>
<tr>
<td>PLANTS</td>
<td>3852 (63%)</td>
<td>3820 (63%)</td>
</tr>
<tr>
<td>AFROTROPICAL ANTELOPES</td>
<td>70 (71%)</td>
<td>53 (54%)</td>
</tr>
</tbody>
</table>

**Are the same sites selected by a scoring method compared to an iterative method?**

Only five of the first twelve sites selected by the two methods were the same for the
antelope data (Table 4.3). Serengeti and Kafue were the first two sites selected by both methods and Selous, Boma and Odzala all featured in both site networks. However seven of the highest scoring sites for antelope species richness were not selected by the iterative algorithm (in the first twelve sites). Of the 37 sites selected to represent all species iteratively 17 (49.9%) were not in the top 40 scoring sites.

Table 4.3. First twelve sites selected to represent Afrotropical antelopes using two alternative selection methods (scoring versus an iterative richness algorithm). The cumulative number of species that would be represented as sites are added to the network is also given.

<table>
<thead>
<tr>
<th>SCORING</th>
<th>ITERATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>SITE</td>
<td>NO. OF SPECIES</td>
</tr>
<tr>
<td>STEP</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Serengeti</td>
</tr>
<tr>
<td>2</td>
<td>Kafue</td>
</tr>
<tr>
<td>3</td>
<td>Selous</td>
</tr>
<tr>
<td>4</td>
<td>Boma</td>
</tr>
<tr>
<td>5</td>
<td>Gorongosa</td>
</tr>
<tr>
<td>6</td>
<td>Luangwa</td>
</tr>
<tr>
<td>7</td>
<td>Tsavo</td>
</tr>
<tr>
<td>8</td>
<td>Kruger</td>
</tr>
<tr>
<td>9</td>
<td>Zakouma</td>
</tr>
<tr>
<td>10</td>
<td>Manova</td>
</tr>
<tr>
<td>11</td>
<td>Ruaha</td>
</tr>
<tr>
<td>12</td>
<td>Odzala</td>
</tr>
</tbody>
</table>

Top sites for bird species richness were predominantly (six out of the eight) in the north-east bushveld and coastal region (Figure 4.1). The other two sites were in the south-east coastal hinterland and lowlands. In contrast the first few sites selected iteratively sampled a much greater range of the habitat types in the region. Only three of the first eight sites were in the north-east bushveld and coastal region. Three sites were montane and highland sites in the west of KwaZulu Natal, one site was in mistbelt habitat and one
Figure 4.1. Location of the first eight sites selected for birds in KwaZulu Natal by the two alternative selection methods (top-ranking sites for species richness & an iterative richness selection)

(a) Top sites

(b) Iterative selection
Figure 4.2. Location of the first eight sites selected for mammals in KwaZulu Natal by the two alternative selection methods (top-ranking sites for species richness & an iterative richness selection)
in the south-east coastal lowlands (Figure 4.2). An iterative selection included some sites that were of overall low species richness for birds, for example square CL in the Drakensberg, which with 189 species has less than half the number of species than that recorded from the most species rich quarter degree square.

Table 4.4. First eight sites selected to represent bird species in KwaZulu Natal using two alternative selection methods (scoring versus an iterative richness algorithm). The cumulative number of species that would be represented as sites are added to the network is also given. Sites are coded as letters corresponding to the west-east and north-south position of the quarter degree grid cell.

<table>
<thead>
<tr>
<th>STEP</th>
<th>SCORING</th>
<th>ITERATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SITE</td>
<td>NUMBER OF SPECIES IN SITE</td>
</tr>
<tr>
<td>1</td>
<td>ND</td>
<td>380</td>
</tr>
<tr>
<td>2</td>
<td>PD</td>
<td>378</td>
</tr>
<tr>
<td>3</td>
<td>NA</td>
<td>359</td>
</tr>
<tr>
<td>4</td>
<td>PA</td>
<td>359</td>
</tr>
<tr>
<td>5</td>
<td>IM</td>
<td>357</td>
</tr>
<tr>
<td>6</td>
<td>GL</td>
<td>326</td>
</tr>
<tr>
<td>7</td>
<td>PF</td>
<td>326</td>
</tr>
<tr>
<td>8</td>
<td>NF</td>
<td>321</td>
</tr>
</tbody>
</table>

Two of the first eight areas selected iteratively were of below average species richness for birds, and one of these areas was in the bottom five percent of sites scored for species richness. Only two of the first eight sites were common to the scoring and iterative selections for mammals in KwaZulu Natal (Figure 4.2, Table 4.5). Top sites were almost exclusively in northern dry uplands and north-east bushveld areas, although two of the top sites included mistbelt habitat (GK and KD). However the iterative selection included a coastal lowland site (RB), a highland/montane site in the west (AH) and a moist upland area (EE) in addition to those habitats represented in the top sites. Top sites for plant richness sampled most habitat types and the iterative selection of eight sites contained seven of the same sites (Figure 4.3, Table 4.6).
Figure 4.3. Location of the first eight sites selected for plants in KwaZulu Natal by the two alternative selection methods (top-ranking sites for species richness & an iterative richness selection)
Table 4.5. First eight sites selected to represent mammal species in KwaZulu Natal using two alternative selection methods (top sites versus an iterative richness algorithm). The cumulative number of species that would be represented as sites are added to the network is also given.

<table>
<thead>
<tr>
<th>STEP</th>
<th>SITE</th>
<th>NUMBER OF SPECIES</th>
<th>CUMULATIVE NUMBER OF SPECIES</th>
<th>TOP SITES</th>
<th>NUMBER OF SPECIES</th>
<th>CUMULATIVE NUMBER OF SPECIES</th>
<th>ITERATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>KD</td>
<td>40</td>
<td>40</td>
<td>KD</td>
<td>40</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>JD</td>
<td>38</td>
<td>42</td>
<td>PA</td>
<td>33</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>GK</td>
<td>37</td>
<td>49</td>
<td>GK</td>
<td>37</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>KC</td>
<td>36</td>
<td>54</td>
<td>EE</td>
<td>26</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>PD</td>
<td>36</td>
<td>56</td>
<td>LF</td>
<td>31</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>MG</td>
<td>36</td>
<td>57</td>
<td>RB</td>
<td>6</td>
<td>62</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>NF</td>
<td>35</td>
<td>58</td>
<td>FE</td>
<td>10</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>ND</td>
<td>34</td>
<td>58</td>
<td>AH</td>
<td>19</td>
<td>64</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.6. First eight sites selected to represent plant species in KwaZulu Natal using two alternative selection methods (top sites versus an iterative richness algorithm). The cumulative number of species that would be represented as sites are added to the network is also given.

<table>
<thead>
<tr>
<th>STEP</th>
<th>SITE</th>
<th>NUMBER OF SPECIES</th>
<th>CUMULATIVE NUMBER OF SPECIES</th>
<th>TOP SITES</th>
<th>NUMBER OF SPECIES</th>
<th>CUMULATIVE NUMBER OF SPECIES</th>
<th>ITERATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CJ</td>
<td>1123</td>
<td>1123</td>
<td>CJ</td>
<td>1123</td>
<td>1123</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>GL</td>
<td>1092</td>
<td>1870</td>
<td>IM</td>
<td>1061</td>
<td>1969</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>FK</td>
<td>1087</td>
<td>2254</td>
<td>GL</td>
<td>1092</td>
<td>2490</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>IM</td>
<td>1061</td>
<td>2759</td>
<td>NF</td>
<td>890</td>
<td>2880</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>KE</td>
<td>958</td>
<td>3081</td>
<td>CL</td>
<td>913</td>
<td>3186</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>JM</td>
<td>923</td>
<td>3345</td>
<td>KE</td>
<td>958</td>
<td>3462</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>CL</td>
<td>913</td>
<td>3593</td>
<td>GQ</td>
<td>860</td>
<td>3684</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>NF</td>
<td>890</td>
<td>3820</td>
<td>JM</td>
<td>923</td>
<td>3852</td>
<td></td>
</tr>
</tbody>
</table>
How is the efficiency of the sites network affected if some sites have already been set aside for protection?

Table 4.7 shows how the efficiency of an iterative selection procedure is affected by varying numbers and quality of sites already having been set aside for conservation. As expected when sites were already assigned and additional sites then selected iteratively to represent all species, the resulting network was less efficient than when all sites were selected iteratively (Table 4.7).

Table 4.7. Effect of having some sites already preassigned to the site network on the efficiency of an iterative selection of sites for Afrotropical antelopes.

<table>
<thead>
<tr>
<th>NO. SITES SET ASIDE</th>
<th>RICHNESS OF SITES</th>
<th>MEAN TOTAL NUMBER OF SITES NEEDED TO REPRESENT ALL SPECIES</th>
<th>N</th>
<th>S.E.</th>
<th>% ADDITIONAL SITES NEEDED RELATIVE TO WHEN NO SITES ARE PREASSIGNED</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-</td>
<td>36.5</td>
<td>15</td>
<td>0.133</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>LOW</td>
<td>38.1</td>
<td>15</td>
<td>0.165</td>
<td>4.38</td>
</tr>
<tr>
<td>2</td>
<td>HIGH</td>
<td>36.6</td>
<td>15</td>
<td>0.235</td>
<td>0.274</td>
</tr>
<tr>
<td>5</td>
<td>LOW</td>
<td>40.3</td>
<td>15</td>
<td>0.248</td>
<td>10.4</td>
</tr>
<tr>
<td>5</td>
<td>HIGH</td>
<td>38.1</td>
<td>15</td>
<td>0.274</td>
<td>4.38</td>
</tr>
<tr>
<td>10</td>
<td>LOW</td>
<td>43.7</td>
<td>15</td>
<td>0.305</td>
<td>19.7</td>
</tr>
<tr>
<td>10</td>
<td>HIGH</td>
<td>40.6</td>
<td>15</td>
<td>0.475</td>
<td>11.2</td>
</tr>
<tr>
<td>15</td>
<td>LOW</td>
<td>48.0</td>
<td>10</td>
<td>0.493</td>
<td>31.5</td>
</tr>
<tr>
<td>15</td>
<td>HIGH</td>
<td>43.8</td>
<td>10</td>
<td>0.629</td>
<td>20.0</td>
</tr>
</tbody>
</table>

Where N = the number of different site selections (i.e. site networks) made.

The network was significantly less efficient if the preselected sites were of low species richness (Table 4.8), compared to high richness and became less efficient as more sites were preassigned (Tables 4.7 & 4.9). When just two sites were set aside there was no significant difference in the number of sites needed to include all species compared to a situation where no sites had been preassigned when the preselected sites were high diversity \( (t = 0.370, \text{NS}) \). If the preassigned sites were high diversity, then the efficiency
of the network was only significantly reduced when five or more sites were preselected (over four percent more sites needed than when no sites were preassigned to the selection). When the preselected sites were low diversity the resulting network was significantly less efficient with only two sites preselected (Table 4.9). When fifteen low diversity sites were preselected, an average of 11.5 more sites were needed to incorporate all species (31.5% more sites) than when no sites were preselected.

Table 4.8. t-tests for the differences between the mean number of sites needed to represent all Afrotropical antelope species when high richness compared to low richness sites are preassigned to the site network for different numbers of preassigned sites (2, 5, 10 & 15). P values give the two-tailed significance level.

<table>
<thead>
<tr>
<th>NO. SITES SET ASIDE INITIALLY</th>
<th>t</th>
<th>d.f.</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>5.23</td>
<td>28</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>5</td>
<td>5.95</td>
<td>28</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>10</td>
<td>5.49</td>
<td>28</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>15</td>
<td>5.25</td>
<td>18</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table 4.9. t-tests for the difference between the mean number of sites needed to represent all Afrotropical antelope species with no sites preassigned to the network compared to when different numbers of sites, of varying richness are preassigned. P values give the two-tailed significance level.

<table>
<thead>
<tr>
<th>NO. SITES SET ASIDE INITIALLY</th>
<th>RICHNESS OF SITES SET ASIDE</th>
<th>t</th>
<th>d.f.</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>LOW</td>
<td>7.55</td>
<td>28</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>2</td>
<td>HIGH</td>
<td>0.370</td>
<td>28</td>
<td>NS</td>
</tr>
<tr>
<td>5</td>
<td>LOW</td>
<td>13.5</td>
<td>28</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>5</td>
<td>HIGH</td>
<td>5.25</td>
<td>28</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>10</td>
<td>LOW</td>
<td>21.6</td>
<td>28</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>10</td>
<td>HIGH</td>
<td>8.32</td>
<td>28</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>15</td>
<td>LOW</td>
<td>26.7</td>
<td>23</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>15</td>
<td>HIGH</td>
<td>13.7</td>
<td>23</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Are components of diversity that do not occur in the most species rich sites better
represented by an iterative selection process?
It was predicted that since an iterative selection takes account of complementarity as well
as richness, certain components of diversity that were not represented in high richness
sites might be better represented in an iterative selection.

Endemic birds were the major group to show a different pattern of richness to that of
birds as a whole in KwaZulu Natal (Chapter 3). Bird species richness is greatest in the
north-east and along the east coast of KwaZulu Natal (Chapter 3). Endemic bird richness
is greatest in the high altitude west of KwaZulu Natal (Chapter 3). Top sites for bird
species richness represented 40 of the 68 (59%) endemic bird species (Table 4.10). In
contrast the iterative selection picked up 61 (90%) of the endemic bird species, twenty-
one species more than in the top scoring sites (Table 4.10).

The iterative selection was also better at representing threatened bird species - an
additional ten species compared to the top site selection (Table 4.10). Threatened plants
and mammals were also better represented in the iterative selections compared to the top
sites (Table 4.10).

Table 4.10. Comparison of the ability (in terms of number and percentage of species)
of different site selection methods at representing threatened and endemic species for
birds, mammals and plants in KwaZulu Natal.

<table>
<thead>
<tr>
<th>SITE SELECTION METHOD</th>
<th>ITERATIVE RICHNESS</th>
<th>SCORING</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>THREATENED BIRDS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EIGHT SITES</td>
<td>75 (93%)</td>
<td>65 (80%)</td>
</tr>
<tr>
<td>SIXTEEN SITES</td>
<td>77 (95%)</td>
<td>74 (91%)</td>
</tr>
<tr>
<td><strong>ENDEMIC BIRDS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EIGHT SITES</td>
<td>62 (91%)</td>
<td>40 (59%)</td>
</tr>
<tr>
<td>SIXTEEN SITES</td>
<td>64 (94%)</td>
<td>51 (75%)</td>
</tr>
<tr>
<td><strong>THREATENED MAMMALS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EIGHT SITES</td>
<td>20 (95%)</td>
<td>19 (91%)</td>
</tr>
<tr>
<td>SIXTEEN SITES</td>
<td>21 (100%)</td>
<td>21 (100%)</td>
</tr>
<tr>
<td><strong>THREATENED PLANTS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EIGHT SITES</td>
<td>54 (44%)</td>
<td>48 (39%)</td>
</tr>
<tr>
<td>SIXTEEN SITES</td>
<td>85 (69%)</td>
<td>70 (56%)</td>
</tr>
</tbody>
</table>

Forty-one species of antelope were classified as being under some degree of threat.
When the twelve top scoring sites for species richness were selected only eight of these
threatened species were represented in the network (20% of the total) (Table 4.11). In contrast an iterative selection of twelve sites incorporated 18 of the threatened species (44%). Similarly the top scoring sites only represented 21% of the restricted-range species diversity compared to 47% in the iterative network (Table 4.11). More taxonomic diversity was also represented in the iterative selection, largely because this incorporated more species overall (Table 4.11).

Table 4.11. Comparison of the ability (in terms of % of species/diversity represented) of different site selection methods at representing threatened species, restricted-range species and taxonomic diversity for Afrotropical antelopes.

<table>
<thead>
<tr>
<th>SITE SELECTION METHOD</th>
<th>ITERATIVE RICHNESS</th>
<th>SCORING</th>
</tr>
</thead>
<tbody>
<tr>
<td>THREATENED SPECIES</td>
<td>44</td>
<td>20</td>
</tr>
<tr>
<td>RESTRICTED-RANGE RICHNESS</td>
<td>47</td>
<td>21</td>
</tr>
<tr>
<td>TAXONOMIC DIVERSITY</td>
<td>69</td>
<td>53</td>
</tr>
</tbody>
</table>

DISCUSSION.

How much more efficient is an iterative site selection at representing species richness compared to selection of the top-scoring sites and how does this relate to the patterns of species richness and spatial turnover in the region?

Number of sites needed to represent all species.

For the taxa in this analysis, an iterative approach to site selection was always more efficient at representing all attributes than a scoring approach. It is well established that an iterative approach is more efficient at representing attributes (Margules et al. 1988, Bedward et al. 1991, Kershaw et al. 1994), and this implies that a conservation policy of selecting high value sites with no attempt at achieving representativeness or complementarity between sites may not sample the full range of biodiversity at all well. For example, an iterative algorithm could represent all 122 snake species occurring in southern Africa in just 23 sites, whereas if the top scoring sites are selected then 548 sites need to be selected before all the species are represented (Lombard et al. 1995). Even
if only those sites that add new species out of the top 548 ranking sites are selected then 34 are needed to represent all species which is one and a half times more sites than the iterative selection (Lombard et al. 1995). Many important components of biodiversity (for example, threatened and endemic species) are not represented in countries or areas of high species richness (so would not be represented if only the most species rich areas are selected), and many of the species in areas of high species richness may be widespread and not under any threat (Thirgood & Heath 1994, Kershaw et al. 1995).

*Numbers of species represented in priority networks selected by scoring and iterative methods.*

Often it will not be possible to conserve enough sites to represent all species. When the number of sites is limited, use of an iterative richness algorithm is more efficient at representing species than selection of top scoring sites. The degree to which an iterative selection will more efficient depends on the number of sites that can be selected (since all selections will eventually converge in terms of the number of species represented), and also the level of complementarity between sites and the pattern of species richness across sites.

When complementarity is generally low for a taxon then top scoring sites will encompass a large amount of regional species richness, so when only a few sites can be selected there will not be a huge difference in the total number of species represented by the two methods. This was the case with the bird data. For example the top eight sites for bird species richness would represent 503 of the 570 bird species in the region. However all these sites are in the east of KwaZulu Natal and do not sample the major east-west gradient of species turnover (Chapter 3). If sites are selected from more complementary areas more than 540 species can be represented in eight sites, even though these sites are not the top eight most species rich.

If species richness across sites is fairly equable, but complementarity is high (and not even between sites) then there will be large differences in the efficiency of the different selections. In the case of the antelopes, where sites contained between one and seventeen
species (with many sites containing a similar number), the iterative selection of twelve sites contained seventy species compared to only fifty-three in the top sites.

Given that plant species in KwaZulu Natal show high complementarity across sites it might be expected that an iterative selection would contain many more species than the top sites. However the two selections were very similar in efficiency. This was partly because top sites for plant species were not confined to one area or habitat type (as was the case with birds and mammals in KwaZulu Natal) so complementarity was high between top sites. Another reason was due to the pattern of species richness between sites. If there is a big difference in the richness values of different sites then two very species rich sites that are not very complementary may between them contain more species than two low richness sites that may be completely different in species composition. Since all sites were reasonably complementary for plants, an iterative selection was virtually identical to the selection of top scoring sites for the initial sites selected.

Taxa that show a high level of spatial turnover will also require more sites overall to represent all the species in the region. Plant species in KwaZulu Natal exhibit a greater degree of spatial turnover compared to bird and mammal species (Chapter 3). One hundred and forty sites were needed to represent all plant species compared to only 27 to include all birds. Obviously there were many more plant species (6111) than birds (570), but if sites were selected iteratively to include all grass species (400 species) then 51 sites would be needed which is still considerably more than the number of sites needed to represent a larger number of bird species. Similarly, when woodlands were selected iteratively to represent bird and plant species in Norway, only 20% of the total area was needed to incorporate all bird species compared to 75% to represent all plant species which show a greater degree of spatial turnover in species composition between sites (Sætersdal et al. 1993).

Are the same sites selected by a scoring method compared to an iterative method? Sites selected by the iterative algorithm tended to sample a wider range of habitat types
than selection of the top scoring sites in KwaZulu Natal. Top sites for bird and mammal species richness were concentrated in particular habitats in KwaZulu Natal so that top sites for species richness were restricted to these areas. In contrast, since the iterative selection achieved greater complementarity, the sites selected came from a wider range of habitat types and in many cases included sites of overall low species richness.

**How is the efficiency of the sites network affected if some sites have already been set aside for protection?**

For the antelope data the efficiency of the network was reduced when some sites were automatically preassigned. The reduction in the efficiency at representing all the species was most marked when large numbers of sites and/or low diversity sites were preassigned. Nonetheless, when fifteen high scoring sites were set aside, on average 20% more sites were needed to include all the species than with no sites preassigned to the selection. When sites were selected iteratively in western New South Wales with the aim of representing all natural environment types, the efficiency of an iterative selection procedure was significantly reduced when the analysis was made with properties already acquired for conservation automatically included (Pressey & Nicholls 1989b). With conserved areas included in the analysis 31% more sites and a 44% larger area were needed to include all attributes (Pressey & Nicholls 1989b). The decrease in efficiency (and therefore increase in cost) of fully representing diversity following ad hoc reservations has also been observed by Rebelo & Siegfried (1992) and Pressey (1994).

When an iterative algorithm was used to represent plant species in southern Africa, all species could be incorporated in 877 of the total 3300 quarter degree grid squares (Rebelo 1994). Five hundred and sixteen of these 877 squares were in the existing reserve network along with a further 350 squares that were not selected by the iterative algorithm. If these 866 squares representing the existing reserves are taken as preassigned sites, a further 499 sites were required to represent all species (1365 sites in total) which is 1.56 times more sites than the configuration needed when sites were selected with none preassigned to the network (Rebelo 1994). This could have major implications given that areas set aside for conservation are often those sites which have
low value for commercial land use or human habitation, rather than having high conservation value (Pressey 1994).

Are components of diversity that do not occur in the most species rich sites better represented by an iterative selection process?

An iterative selection also went some way towards solving the problem of representing aspects of diversity not occurring or concentrated in the most species rich sites. Threatened species or endemic species for example may not occur in the sites which contain the most species. Additionally high richness sites may be degraded and contain invasive or opportunistic species that are of low conservation priority. Selection of the most species rich sites may therefore miss the highest priority species. In KwaZulu Natal the most species rich sites for birds were in the north-east and east coast areas. Many of the species contributing to the high richness of the north-east area were species at the southern limit of their range in Africa. In contrast endemic species richness was highest in the western, high altitude habitats. A network of top scoring sites for species richness was poor at representing endemic species, but good at representing edge of range species. Top sites for bird species richness only represented 40 of the 68 endemic bird species and this figure is in some ways optimistic since some of the endemic species that breed in the western highlands, move to the coast during winter so have eastern records (i.e. some endemic species were only represented in their wintering ranges). Although many threatened species of bird did occur in the top sites for species richness, a number of the globally threatened species had western distributions and like many endemics would not be included in a network of top sites. An iterative selection of sites was an improvement over top sites at incorporating both threatened and endemic bird species. This was because the selection process took account of the east-west complementarity in species distributions and therefore selected western sites as well as species rich eastern sites.

Threatened antelopes were not well represented by selection of the top scoring sites. Only twenty percent of the threatened species were represented by the top twelve sites for species richness. In contrast, an iterative selection represented more than twice this
number of threatened species because the sites selected were more complementary in terms of the species present. Similarly, restricted-range antelope species richness was better represented in an iterative selection, compared to the top scoring sites. Iterative selections were also an improvement over top sites at representing threatened plants and mammals and antelopes, although in the case of threatened mammals the improvement was marginal because threatened mammal richness was correlated with overall species richness (Chapter 3).

It is necessary to be aware of the limitations of such studies when setting real priorities for conservation. The data set on Afrotropical antelope took no account of species occurring outside the protected areas, and so might not reflect the true distribution of the species, especially if the protected areas were established to preserve unusual, rare or threatened species or habitats. Trying to achieve representation of all the species will be a meaningless exercise for conservation if evaluation is not carried out with reference to the complete distribution of each species - for the antelope data some species might occur outside the protected areas while others do not. The current level of protection and management of the reserves will also be relevant since a species' occurrence in a protected area may not necessarily guarantee its persistence. Ideally the evaluation of areas should be made at a biogeographically distinct level to avoid the problem of giving equal weight to species that also occur outside the region and those restricted to the region being evaluated. ‘Politically rare species’ may distort reserve configurations by overvaluing species that can be better preserved elsewhere, however regional and subcontinental conservation requirements will vary and it may be valid in some cases to value species that are rare in the region but common elsewhere (Rebelo 1994). In reality though, most practical conservation decisions are made at national or other political levels and in these instances an iterative selection procedure is more likely to incorporate those elements of biodiversity that are globally important than a scoring approach.

In the case of the KwaZulu Natal evaluations it is very unlikely that sites the size of quarter degree grid cells (660km²) could be set aside as conservation areas. Where information at finer scales is not available it may be possible to select representative sites
from within each larger selected cell which include the majority of species in the cell. Alternatively selected cells could be managed for conservation alongside other land uses. For taxa that show high spatial turnover in a region a large proportion of the total area will be required to represent all species. For these taxa alternative approaches to the designation of protected areas will be needed. For example the most important areas could be designated for pure conservation with other areas being managed partly for conservation or with some sort of environmental controls.

Under an iterative selection it is possible to identify which attributes in a site are contributing to the global representation goal and these could then be given priority within that site. For example, for birds in KwaZulu Natal several Drakensberg sites were selected. These sites contained some widespread species that had already been represented in other selected sites e.g. the Fork-tailed Drongo (*Dicrurus adsimilis*) occurs throughout KwaZulu Natal in all habitats and was recorded from 160 of the 166 grid cells. The Drakensberg sites also included species like the Rockjumper (*Chaetops aurantius*) which was only recorded from 15 squares and is confined to western Drakensberg uplands. The Rockjumper therefore contributes to the iterative value of the Drakensberg site whereas the Fork-tailed Drongo does not, so within the Drakensberg sites priority should be given to the conservation of species like the Rockjumper.

When information for all or many of the species in a region is lacking iterative selection procedures could still be of value if applied to groups of indicator species that have representatives in the majority of habitat types, collectively occur across the whole region and which show similar patterns of species richness to other groups. For example, from the previous chapter, habitat specific birds might provide a suitable surrogate measure.

Iterative algorithms are useful in conservation planning in that they enable flexibility in the site selection process. It is very easy to examine the consequences of different priority-area choices. Constraints can easily be incorporated into the selection algorithm. For example, it is possible to design an iterative algorithm to select sites from each representative habitat type in a region or to select sites that are as close together as
possible, or which ensure multiple representations of species (Nicholls & Margules 1993). Additionally species can be given some sort of weighting so as to increase the amount of taxonomic diversity represented at each step rather than just simple species richness for example (Vane-Wright et al. 1991). Iterative algorithms, unlike complex site scoring procedures, are explicit and the effect of changing the selection rules can easily be investigated.

CONCLUSIONS.

- An iterative algorithm that takes account of species complementarity across sites was always more efficient than a scoring method at representing all attributes.
- When insufficient areas to represent all attributes could be selected, an iterative approach was again more efficient at incorporating attributes.
- The efficiency of a site network was reduced when some sites were arbitrarily assigned to the network. This was particularly the case when the preassigned sites were of low species richness and the reduction in efficiency increased as more sites were preassigned.
- Iterative methods proved to be particularly advantageous in situations where certain attributes of biodiversity do not occur in the most species rich sites, for example, at representing endemic bird species in KwaZulu Natal and threatened antelopes in the Afrotropics.
CHAPTER FIVE.

A COMPARISON OF THE EFFICIENCY OF DIFFERENT ITERATIVE RESERVE SELECTION ALGORITHMS.

INTRODUCTION.

The goal of systematic conservation evaluation is to seek maximally efficient solutions for fully representative reserve networks, either as the minimum number of areas, or smallest configuration of sites, capable of providing sustainable protection for all species found in a given region (Pressey et al. 1993, Vane-Wright et al. 1994). This aim has lead to the development of a number of different iterative algorithms that attempt to achieve this goal. However algorithms that are designed to find the minimum (or near minimum) number of sites to include all species have been criticised on the grounds that conservationists should not be trying to minimise the area set aside for conservation, since the minimum area needed to include all species may not be sufficient to maintain the species over time or to account for changes in range or distribution. However, developments of simple, iterative algorithms can ensure multiple species representations, or that a certain minimum area is selected, so they do not necessarily just involve finding the smallest area to include all species once. Secondly it is unlikely that enough sites to represent all species in a region will be available for conservation. Iterative algorithms can however identify networks of sites that contain close to the maximum number of species that could be included within a region, which scoring approaches cannot guarantee.

Underhill (1994) points out that some iterative algorithms do not in fact minimise the number of sites needed to represent all species, and do not maximise the number of species represented in a fixed subset of sites. According to Underhill (1994), so called ‘greedy algorithms’ may produce ‘grossly suboptimal results’ and he proposes the use of optimisation models from operations research to avoid these problems. The reason why iterative algorithms cannot guarantee an optimum solution is that the site selection process proceeds in a stepwise fashion. This means that sites selected early on in the
process may, at a later point become redundant - that is the species that contributed to the site being selected could all be present in subsequent selections. This is why iterative richness algorithms are often referred to as greedy algorithms. The following is an example of a case where a richness algorithm would not find the optimum solution:

<table>
<thead>
<tr>
<th>SPECIES</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>sitel</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>site2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>site3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>site4</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

If a richness algorithm were used, the first site to be selected would be site one with four species. Sites two and three would then both add one unrepresented species each. In total sites one, two and three would all be required to represent all six species. It is obvious however that only sites two and three would in fact be necessary for full representation, site one becoming redundant.

In this chapter I discuss the merits and problems of alternative iterative selection algorithms with reference to the data on birds, mammals and plants in Natal. In addition I compare the efficiency of iterative algorithms compared to integer programming techniques (which can guarantee to find globally optimal solutions) using data on vertebrate distributions from Oregon, North America. Two broad classes of iterative algorithms are compared - richness algorithms and rarity algorithms. Two types of representation problem are investigated, the set covering problem (SCP) and the maximal covering problem (MCP) (Camm et al. in press, Church et al. in press). The SCP is to include all attributes (e.g. species) in the minimum number of sites possible. The MCP is to maximise the number of attributes represented in a fixed number of sites. MCP is probably the most realistic conservation problem since there is likely to be a limit on the number of sites that can be selected for conservation and the goal will be to select a limited number of sites to maximise the inclusive diversity. However the SCP might be relevant where for example the goal of a network is to represent all threatened species in an area or representatives from all ecological guilds.
Specifically this chapter considers the following questions:

1. Do different iterative algorithms vary significantly in their efficiency at representing all species in a region?
2. Are the same sites selected by the different algorithms?
3. How well do the different algorithms represent species richness when the number of sites selected is limited?
4. How close do iterative algorithms come to finding the optimum solution for a reserve network (as found by an integer programming solution)?

If iterative heuristics are consistently, grossly suboptimal at solving these sorts of problems then other methods of reserve selection may be needed. However if they are only marginally suboptimal then they could still provide a valuable tool for investigating priority area networks particularly if the integer programming techniques are very intensive in terms of computing time and capacity.

**METHODS.**

**Algorithms.**

*Iterative richness algorithms.*

As described in the previous chapter these algorithms select the site at each iteration which maximises the number of previously unrepresented species that are added to the site network. Two slightly different richness algorithms were used in this analysis. The antelope data were analysed in WORLDMAP (Version 2.5, Williams 1992) using the following richness algorithm:

**Richness I.**

1. Select the site with the most species
2. Select the site that adds the most new species to the network
3. If there is a tie between sites in the number of new species that would be added, select the site with the most species overall (represented + unrepresented)
4. If there is still a tie at this point select a site at random from those tied at step three.
The following richness heuristic was programmed in FORTRAN and applied to the Natal data for birds, mammals and plants:

**Richness II.**
1. Select the site containing the most species
2. Select the site that adds the most unrepresented species
3. If there is a tie between sites in the number of new species added, then select the first site in the sequence.

To accommodate the potential problem of site redundancy, a backward searching optimisation algorithm (*sensu* Rebelo 1994) (called backtracking) was applied following the iterative richness algorithm for the Natal data. This basically ran through the sites that had been selected using the richness algorithm and removed any sites where all the species present were also represented by other sites. In the above example, site one contains species that are all present in the other selected sites and so would be removed.

**Iterative rarity algorithms.**
These select sites on the basis of the number of sites in the region that the component species are found in, so that the least frequently occurring species in the region are picked up by the network first. Several different rarity algorithms were used.

**Rarity I**
This was developed from the heuristic algorithm described by Margules et al. (1988). Initial sites selected are those that contain unique species (i.e. found nowhere else in the region). Subsequent sites are selected to maximise inclusion of species occurring in two, then three, then four sites etc. until all species have been represented.

1. Select all sites containing species unique to that site (not occurring in any other sites). (These sites are ranked by the number of unique species that they contain and then according to the number of species they add to the network).
2. Starting with the least frequent, unrepresented species (i.e. those species
occurring in just two sites if there are any, if not in three sites etc.), select from
the sites in which they occur the site containing the most other unrepresented
species of equivalent frequency (ie. occurring in only two sites). If none of the
sites contain other unrepresented species of equal rarity then select the site that
adds the most unrepresented species (of any frequency) to the network.

3. If two or more areas contribute the same number of equally rare species, select
from these areas the site that contains the maximum number of unrepresented
species of any frequency.

4. If two or more sites contain the same number of unrepresented species on any
frequency, then select the first site encountered.

<table>
<thead>
<tr>
<th>SPECIES</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>site1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>site2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>site3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>site4</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

So in the above example, site three contains the only species which occurs in one site
(species F), so this site (three) will be selected first (this would also pick up species D
and E). That would leave species A, B and C unrepresented. Of these three species,
species A and C occur in just two sites, and site two contains both these species so this
site is selected second. This leaves none of the species unrepresented, so the network
consists of sites two and three.

**Rarity II**

This algorithm can be applied using WORLDMAP software and like the above it weights
site selection by the relative rarity of the fauna in each site. Instead of attempting to
represent individual species starting with the most infrequent, a score is calculated for
each site based on (i) the number of species and (ii) how infrequent the species are among
each site. Each species is scored as the inverse of the number of sites that it occurs in,
and the site score is simply the sum of the component species scores. So a species which
only occurs in one site will have a score of 1/1 and a species occurring in 100 sites will
score 1/100.
In WORLDMAP site scores are expressed as the percentage of the total scores for all species in the region:

\[
\text{RESTRICTED-RANGE SCORE} = \frac{\sum \frac{100}{X}}{\sum \frac{n_i}{X}}
\]

(1)

where \( X \) is the number of sites occupied by each species, \( n \) is the number of species in each site and \( n_i \) is the number of species in the data set.

Rare species therefore contribute more to a site score than commoner species. However the site scores are influenced by the absolute species richness of the site as well as the regional rarity of the component species. The selection algorithm proceeds by selecting the site with the highest initial score. The scores are then recalculated with the represented species removed, and the site with the highest new score is then selected. This process is continued until all species have been represented.

For the example described under Rarity I, the scores would be calculated as follows:

- species A occurs in 2 sites = 0.5
- species B occurs in 3 sites = 0.333
- species C occurs in 2 sites = 0.5
- species D occurs in 2 sites = 0.5
- species E occurs in 2 sites = 0.5
- species F occurs in 1 site = 1.0

So the sum of all species scores is 0.5 + 0.333 + 0.5 + 0.5 + 0.5 + 1.0 = 2.333.

the scores for each site would initially be:

- site 1: \([(0.333 + 0.5 + 0.5 + 0.5)/2.333] \times 100 = 78.6\)
- site 2: \([(0.5 + 0.333 + 0.5)/2.333] \times 100 = 57.1\)
- site 3: \([(0.5 + 0.5 + 1.0)/2.333] \times 100 = 85.7\)
- site 4: \([(0.333 + 0.50/2.333] \times 100 = 35.7\)
So site three has the highest score and would be selected first. This would remove species D, E and F and the recalculated site scores would be:

site 1: 35.7%
site 2: 57.1%
site 3: 0%
site 4: 35.7%

So site two would be selected second and at this point all species have been represented.

**Rarity III**

This is a modification of Rarity II which gives greater weighting to the regional rarity of site faunas relative to the species richness of sites. Instead of using the inverse of the number of sites a species occurs in, the inverse squared is used. So a species only found in one site will score $1/1^2$ for the site where it occurs, and a species found in 100 sites will score $1/100^2$ for those sites that it occurs in.

In both Rarity II and Rarity III ties were resolved simply by selecting the first tied site encountered.

Sites were selected iteratively for Afrotropical antelopes and for birds, mammals and plants in Natal using the above iterative algorithms to determine the number of sites needed to include all species at least once. In addition the number of species represented in the first twelve sites for antelopes and first eight sites selected for Natal taxa were compared for the different algorithms.

As part of a collaboration with Oregon Biodiversity Consortium aiming to compare the performance of various iterative selection algorithms with integer programming solutions, the algorithms described above were applied to an Oregon data set (Csuti et al. in press). This consisted of 426 terrestrial, vertebrate species distributed among 441 hexagon cells covering the whole of Oregon state (see Csuti et al. in press for details). Each of the 441 hexagon cells were 635 km² in area (Csuti et al. in press). Only terrestrial vertebrates
were considered in the analysis as distribution records for these species were the most comprehensive. Nonetheless, many of the hexagon cells had not been surveyed and lacked species occurrence records. Local experts were therefore asked to classify the probability of a species occurring in each hexagon cell. A species was recorded for a given hexagon if (i) a verified sighting has been made in the last ten years or (ii) the hexagon contained suitable habitat for the species, there had been verified sightings in nearby sites and in the opinion of the local expert it was highly probable that the species occurred in the site (Csuti et al. in press). The results of the iterative algorithms were compared to an integer programming solution using a branch and bound algorithm implemented using the software package LINDO. Details of this technique are given in Csuti et al. (in press). Since the optimal solution to the set covering problem can be guaranteed by solving an integer programming problem the results from this algorithm are a baseline against which to compare the performance of the iterative algorithms described above.

RESULTS.

Do different iterative algorithms vary significantly in their efficiency at representing all species in a region?

Afrotropical antelopes.
The richness algorithm was the least efficient algorithm at representing all species in the region (Table 5.1). Nonetheless all species could be represented in 37 sites which is just 15% of the total number of sites. Both the rarity algorithms used could represent all 99 species in 34 (14%) of sites. The efficiency of the richness selection was therefore 85% and the rarity selections, 86% (Table 5.2).

Birds in KwaZulu Natal.
As with the antelopes, the richness algorithm was less efficient than all the rarity algorithms at representing all species (Table 5.2). When richness was used, 27 sites were needed to represent all species whereas the rarity algorithms could include all species in just 24 sites (Table 5.1). When backtracking was applied, two sites proved to be redundant in the richness selection, but this still meant that 25 sites would be needed to
include all species. No sites were redundant to the rarity networks.

*Mammals in KwaZulu Natal.*
The richness selection took nine sites to include all 65 mammal species (Table 5.1). All three rarity algorithms could represent all the species in just eight sites. When backtracking was applied to the selected sites, one site proved to be redundant in the richness selection so that all species were represented in eight of the nine sites.

*Plants in KwaZulu Natal.*
All algorithms required the same number of sites to represent all plant species (Table 5.1). 140 of the 166 sites were needed to achieve full representation and this is 84% of the total area.

**Are the same sites selected by the different algorithms?**

*Afrotropical antelopes.*
For the antelope data, De Hoop in South Africa, with three species unique to the site was the first area to be selected by both rarity algorithms (Table 5.3). However, with a total of only four species, De Hoop did not rank in the top 40 sites scored for unweighted richness and was not in the top ten sites to be selected iteratively by the richness algorithm. Bale Mountains in Ethiopia has a total of six species and was not selected until after 25 steps when the richness algorithm was used. With one unique species the site is essential for representation of all the species, however Bale also contains the two most widespread species in the data set - occurring in 112 and 75 sites respectively and these may be unnecessarily duplicated, by their inclusion in earlier site selections.

Both Rarity I and Rarity II algorithms selected the same 34 sites, although not in the same order of priority (Table 5.3). A number of high richness sites were not included in either of the rarity selections. For instance neither of the rarity algorithms included Kafue which along with Serengeti is the most species rich site. Kafue is in the first two site selections using species richness, but the fact that it is not needed to represent all species shows that all its species occur in other sites and its inclusion early on in the richness
selection is reducing the efficiency of the network. Kafue is completely redundant in the richness network - if it were removed all species would still be represented in the 36 remaining sites.

*Birds in KwaZulu Natal.*

Complete site networks for all the algorithms were very similar. Only four of the sites selected by Rarity I were not in the richness selection and all of the rarity selections shared between 22 and 23 of their total 24 sites. There were some differences in the first eight sites that were selected by the different algorithms (Figure 5.1). However, all algorithms resulted in the selection of some sites from the north-east of the region and some sites from the western, highland areas in the first eight sites selected (Figure 5.1).

*Mammals in KwaZulu Natal.*

All the rarity algorithms selected the same eight sites (Figure 5.2). These were also the same sites as selected by the richness algorithm (minus the redundant site), except for one site which was a tie (and therefore flexible) with a site selected by the rarity algorithm. Two sites (KD and JD) both contribute the same species to the network, but KD was selected by the richness algorithm because it contains two more species than JD, and JD was selected by the rarity algorithms because it was the first site in the sequence.

*Plants in KwaZulu Natal.*

All the different algorithms selected the same one hundred and forty sites. As with the bird data there were some differences in the first eight sites selected by the different algorithms. The Richness and Rarity II algorithms both selected two sites in the north-east of the region. Rarity I and Rarity III only selected one site from the north-east in the first eight sites selected. All the algorithms resulted in the inclusion of western, highland sites and south-east coastal sites in the first eight sites selected. Five of the first eight sites selected by the richness algorithm were in the first eight sites for Rarity I and III and six in the first eight sites for Rarity II (Figure 5.1).
Figure 5.1. Location of the first eight sites selected by the different iterative algorithms for birds in KwaZulu Natal.
Figure 5.2. Location of the first eight sites selected by the different iterative algorithms for mammals in KwaZulu Natal.
Table 5.1. Number of sites needed to represent all species using four different iterative algorithms. Numbers in brackets denote the number of sites that were 'redundant' following the backtracking procedure which was applied to the richness algorithms for the KwaZulu Natal taxa.

<table>
<thead>
<tr>
<th>ITERATIVE ALGORITHM</th>
<th>RICHNESS</th>
<th>RARITY I</th>
<th>RARITY II</th>
<th>RARITY III</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANTELOPES</td>
<td>37</td>
<td>34</td>
<td>34</td>
<td>-</td>
</tr>
<tr>
<td>BIRDS</td>
<td>27 (2)</td>
<td>24</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>MAMMALS</td>
<td>9 (1)</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>PLANTS</td>
<td>140 (0)</td>
<td>140</td>
<td>140</td>
<td>140</td>
</tr>
</tbody>
</table>

Table 5.2. Efficiency of the different iterative algorithms.

<table>
<thead>
<tr>
<th>ITERATIVE ALGORITHM</th>
<th>RICHNESS</th>
<th>RARITY I</th>
<th>RARITY II</th>
<th>RARITY III</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANTELOPES</td>
<td>0.851</td>
<td>0.863</td>
<td>0.863</td>
<td>-</td>
</tr>
<tr>
<td>BIRDS</td>
<td>0.837</td>
<td>0.855</td>
<td>0.855</td>
<td>0.855</td>
</tr>
<tr>
<td>MAMMALS</td>
<td>0.946</td>
<td>0.952</td>
<td>0.952</td>
<td>0.952</td>
</tr>
<tr>
<td>PLANTS</td>
<td>0.157</td>
<td>0.157</td>
<td>0.157</td>
<td>0.157</td>
</tr>
</tbody>
</table>

How well do the different algorithms represent species richness when the number of sites selected is limited?

Afrotropical antelopes.

If external factors impose a limit on the number of sites that can be conserved (less than the minimum number needed to conserve all species), then species richness is more efficient at maximising species number for the antelope data (Table 5.4). In order to include 50 species (approximately 50% of all species) when richness is used, six sites are needed, whereas when Rarity I is used eight sites are required to include 50 species and with Rarity II, nine sites. Similarly to include 75 species, unweighted richness took 15 sites whereas Rarity I needed 24 sites and Rarity II, 22 sites.

Ten of the first twelve sites selected by the two rarity algorithms were the same (Table 5.3) and both networks represented 59 species after twelve site selections (Table 5.4).
<table>
<thead>
<tr>
<th>SELECTION ORDER</th>
<th>RICHNESS</th>
<th>ITERATIVE ALGORITHM</th>
<th>RARITY I</th>
<th>RARITY II</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Serengeti</td>
<td>De Hoop</td>
<td>De Hoop</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Kafue</td>
<td>Bale Mts.</td>
<td>O-Rime-O-Achim</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Odzala</td>
<td>West Pretorius</td>
<td>Meru</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Boma</td>
<td>Arawale</td>
<td>West Pretorius</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Comoe</td>
<td>O-Rime-O-Achim</td>
<td>Boni Dodori</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>O-Rime-O-Achim</td>
<td>Niokolo-koba</td>
<td>Luangwa</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>West Pretorius</td>
<td>Southern</td>
<td>Gamba</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Selous</td>
<td>Gamba</td>
<td>Kangandala</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Meru</td>
<td>Boni Dodori</td>
<td>Southern</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Haut Dodo</td>
<td>Dahlac Marine</td>
<td>Niokolo-koba</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Boni Dodori</td>
<td>Lame-Burra</td>
<td>Lake Bangweulu</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>De Hoop</td>
<td>Senkelle</td>
<td>Gashaka-Gumti</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Kruger</td>
<td>Meru</td>
<td>Bale Mts.</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Q. Elizabeth</td>
<td>Luangwa</td>
<td>Dahla Marine</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Yangudi Rassa</td>
<td>Gashaka-Gumti</td>
<td>Lame-Burra</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Gash-setit</td>
<td>Lake Bangweulu</td>
<td>Ruwenzori Mts.</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Luangwa</td>
<td>Ruwenzori Mts.</td>
<td>Senkelle</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Manovo</td>
<td>Kangandala</td>
<td>Arawale</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Niokolo-koba</td>
<td>Zanzibar Island</td>
<td>Yangudi Rassa</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Gamba</td>
<td>Kilimanjaro</td>
<td>Serengeti</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Singou</td>
<td>Zeraf</td>
<td>Zeraf</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Gashaka-Gumti</td>
<td>Singou</td>
<td>Selous</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Southern</td>
<td>Lochinvar</td>
<td>Gash Setit</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>C. Kgaligadi</td>
<td>Yangudi Rassa</td>
<td>Lochinvar</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Kpelle</td>
<td>Serengeti</td>
<td>Haut Dodo</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Bale Mts</td>
<td>Boma</td>
<td>Singou</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Kilimanjaro</td>
<td>Yob</td>
<td>Kilimanjaro</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Kangandala</td>
<td>Selous</td>
<td>Sokoke-Arabuko</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Lake Bangweulu</td>
<td>Manova</td>
<td>Boma</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Sokoke-Arabuko</td>
<td>Gorongosa</td>
<td>Maiko</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Gambella</td>
<td>Haut Dodo</td>
<td>Manova</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>Lochinvar</td>
<td>Kpelle</td>
<td>Gorongosa</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>Arawale</td>
<td>C. Kgaligadi</td>
<td>Kpelle</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>Dahlac Marine</td>
<td>Maiko</td>
<td>C. Kgaligadi</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>Lame-Burra</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>Ruwenzori Mts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>Senkelle</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

198
Birds in KwaZulu Natal.

When the site network was restricted to just eight sites there were some differences in the number of species that were represented (Table 5.4). As with the antelopes, the richness algorithm represented more species than the rarity algorithms when the number of sites was limited. However the difference was not very great. The richness algorithm incorporated 541 species in eight sites compared to between 534 and 538 by the rarity algorithms. Unlike the antelopes, a number of the most restricted bird species did occur in the most species rich sites, and these sites featured in all the iterative selections. Richness and rarity selections shared between four and five of their first eight sites (Figure 5.1). Two of the top five percent of richest bird sites were in the first eight sites selected by the Rarity I algorithm and five of the first eight sites selected by Rarity I were in the top ten percent of richest sites.

Plants in KwaZulu Natal.

The number of species represented in the first eight sites was very slightly higher for the richness algorithm than the rarity algorithms (Table 5.4). Both Rarity I and Rarity III which are most heavily weighted towards restricted species, selected the same eight sites initially and in almost the same sequence (Figure 5.3). These eight sites incorporated 3658 species compared to 3852 by the richness algorithm. Rarity II was intermediate, representing 3831 species in the first eight sites, although seven of these sites were the same as the other rarity algorithms. The difference was that Rarity II selected site NF whereas the other rarity algorithms picked site AH. NF has 890 species but only 19 unique species compared to AH which has 658 species and 33 unique species. AH scored higher for the algorithms most weighted towards restricted species (Rarity I and II) whereas NF had a higher score when richness was given greater weighting (Rarity II).
Figure 5.3. Location of the first eight sites selected by the different iterative algorithms for plants in KwaZulu Natal.
Table 5.4. Number of species that are represented by each of the alternative iterative algorithms when the number of sites that can be selected is limited.

<table>
<thead>
<tr>
<th>ITERATIVE ALGORITHM</th>
<th>RICHNESS</th>
<th>RARITY I</th>
<th>RARITY II</th>
<th>RARITY III</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANTELOPES</td>
<td>70</td>
<td>59</td>
<td>59</td>
<td>-</td>
</tr>
<tr>
<td>BIRDS</td>
<td>541</td>
<td>538</td>
<td>536</td>
<td>534</td>
</tr>
<tr>
<td>MAMMALS</td>
<td>64</td>
<td>65</td>
<td>65</td>
<td>65</td>
</tr>
<tr>
<td>PLANTS</td>
<td>3852</td>
<td>3658</td>
<td>3831</td>
<td>3658</td>
</tr>
</tbody>
</table>

How close do iterative algorithms come to finding the optimum solution for a reserve network (as found by an integer programming solution)?

Species accumulation curves for the different algorithms are given in Figure 5.4. All algorithms could include the 426 species in Oregon in less than thirty sites (Table 5.5). As in all the other examples, a richness algorithm was the least efficient method for representing all attributes (27 sites). The rarity algorithms that gave greatest weight to restricted species (Rarity I and Rarity III) both included all species in 24 sites. Rarity II which used species richness more than the other rarity algorithms took 25 sites to represent all species. All algorithms, but most particularly Rarity I and III performed well compared to the optimum solution of 23 sites provided by the integer programming solution (Table 5.5).

Table 5.5. Number of sites needed to include all species in the Oregon data set using four different iterative algorithms (Richness, Rarity I, Rarity II & Rarity III) compared to the optimal solution as derived from integer programming.

<table>
<thead>
<tr>
<th>ITERATIVE SELECTION ALGORITHM</th>
<th>RICHNES S</th>
<th>RARITY I II</th>
<th>RARITY II III</th>
<th>INTEGER PROG</th>
</tr>
</thead>
<tbody>
<tr>
<td>NUMBER OF SITES</td>
<td>27</td>
<td>24</td>
<td>25</td>
<td>24</td>
</tr>
</tbody>
</table>
Figure 5.4. Species accumulation curves for the different iterative algorithms (Richness, Rarity I, Rarity II & Rarity III) for the Oregon data set.
In contrast the richness algorithm was good at representing species when the number of sites was limited to less than the number required to include all species (Table 5.6). Selections of six and twelve sites both represented only two species fewer than the optimum number (from the integer programming solution). After just six site selections 91% of the total species in Oregon were represented and this rose to 96% after twelve selections. Beyond fifteen selections the richness algorithm was only accumulating one species at each selection step up to the last site.

The three rarity algorithms were all inferior to the richness algorithm when it came to representing absolute numbers of species in a limited number of sites (Table 5.6). Rarity I which was most heavily weighted towards rare species had included 85% of species after six selections and 92% after twelve selections. Rarity II, which, of the rarity algorithms, was the most heavily weighted towards species richness of sites, incorporated 87% of species after six selections and 94% after twelve. Rarity III was slower than Rarity II at picking up species richness early on in the selection process because it was more heavily weighted towards rare faunas, but later selections were more efficient at accumulating species.

<table>
<thead>
<tr>
<th>ITERATIVE SELECTION ALGORITHM</th>
<th>RICHNESS</th>
<th>RARITY I</th>
<th>RARITY II</th>
<th>RARITY III</th>
<th>INTEGER PROG</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIX SITES</td>
<td>388</td>
<td>361</td>
<td>372</td>
<td>361</td>
<td>390</td>
</tr>
<tr>
<td>TWELVE SITES</td>
<td>408</td>
<td>390</td>
<td>401</td>
<td>403</td>
<td>410</td>
</tr>
</tbody>
</table>

**DISCUSSION.**

Do different iterative algorithms vary significantly in their efficiency at representing all species in a region?

Comparing iterative selections, rarity algorithms proved to be generally more efficient at representing all species than a species richness algorithm. However, even for large species data sets (more than 400 species) the difference in efficiency was small. Richness
algorithms took three sites more than rarity algorithms to include all species for antelopes and birds, and one site more for mammals. All the algorithms selected the same site network for plants. This was because the large number of plant species only found in one site constrained the sites that could be selected and additionally, many of the sites with the most restricted species were also rich in more widespread species. One thousand four hundred and seven plant species (23% of all species) were recorded from just one site. All sites containing unique species are needed for complete representation (i.e. they are not flexible) and to represent all these species required 138 sites. This left two species unrepresented which required two more sites for them both to be represented giving a total of 140 sites to include all 6,111 species. All algorithms selected these 140 sites.

In a similar study Rebelo (1994) used species level data from PRECIS to identify the configuration of sites needed to represent all plant species recorded from southern Africa. PRECIS contains data on 19,460 taxa from southern Africa (3,950 quarter degree squares). There was a strong positive correlation between the species richness and endemic species richness of the grid squares. Using a rarity algorithm, a total of 877 squares (22% of the total) were needed to represent all species once. This low overall percentage is because more than half the squares have less than 50 species recorded, particularly those squares in Namibia and Botswana. Half the taxa could be represented in just sixteen squares (0.4%) and 90% in 202 grid squares (5%). However of the 877 squares needed to represent all species, 90% were irreplaceable (i.e. not flexible). Ninety-four of the 166 Natal squares (57%) were selected, so even when subspecies and varieties were combined (which tend to occur in fewer sites than the species as a whole) a large proportion of the area is needed.

It is important to consider the objective of the site network since this will determine which selection method is most suitable. Different iterative algorithms vary in their efficiency at incorporating various aspects of biotic diversity, and the efficiency of the different algorithms also varies according to the number of sites that can potentially be conserved (Williams et al. 1991, Williams 1993). For an iterative approach, if the aim
of the site network is to incorporate all the species in the minimum number of areas then a method that initially selects sites containing rare or restricted species will be most efficient. In any selection designed to incorporate every species (or any other attribute), those species that are most ‘rare’ (defined here as species recorded from only one or two sites) will be limiting on the selection. The sites where these species occur are essential in the selection, however if these sites have low overall diversity scores then they will not be selected if a simple species richness algorithm is used until near the end of the selection procedure. Any sites which contain many species, but whose species are common and also occur in the essential limiting sites will be selected early on by a richness algorithm. This will tend to reduce the efficiency of the selection as there may be ‘unnecessary’ duplication of widespread species which occur in sites containing unique species which are selected late on by the richness algorithm. While duplication of species is highly desirable for conservation, it will be the rarest species that will be most difficult to represent multiply and so a restricted-range algorithm designed to represent all species a certain minimum number of times will be more efficient than using a species richness measure. The rarity algorithms all initially select sites to incorporate the ‘rare’ species that are limiting and so ultimately are more efficient at representing every species than selecting sites on the basis of unweighted species richness.

Are the same sites selected by the different algorithms?
Generally, all the different iterative algorithms selected very similar sets of sites, the major difference was in the order in which individual sites were selected. For the KwaZulu Natal taxa, all the algorithms selected sites that sampled the major east-west gradient of species turnover identified in Chapter 3.

How well do the different algorithms represent species richness when the number of sites selected is limited?
Often it will not be possible to conserve enough sites to represent all species. When only a small number of areas could be selected, and therefore not all the species could be included, species richness was generally more efficient at incorporating absolute numbers of species in a small number of sites. A universal feature of the iterative algorithms used
was that the number of species represented rises very rapidly as initial sites are selected, so that the majority of species are represented in a relatively small number of sites. Ninety-five percent of bird species were represented in the first eight sites selected by the richness algorithm, even though it took 27 sites to represent all species. Similarly 63% of plant species were represented in the first eight sites selected by the richness algorithm and 86% in the first 30 sites. Despite this, 140 sites were needed to incorporate all species.

However if 'rare' species (that is species only occurring in a few sites) do not occur in the most species diverse sites, then these species will not be represented in the initial sites selected by a richness algorithm. The rarity algorithms are better at representing these 'rare' species in a limited number of sites, although this may be at the expense of absolute numbers of species. For Afrotropical antelopes, sites containing rare species in many cases did not have many common species (see Chapter 2). Selecting sites on the basis of antelope species richness required more than twice as many sites to represent all species occurring in just one site ('restricted species') than the rarity algorithms. Both Rarity I and II algorithms represented all antelope species occurring in only one site ('restricted species') after 18 selections. The richness algorithm needed all 37 sites to represent all species restricted to just one site. If only 18 sites could be set aside, then using unweighted species richness to select sites would result in the loss of 60% of the 'restricted species'. If 'restricted species' are a priority for conservation then use of a rarity algorithm might be desirable. Although if the aim of a reserve network is to represent a certain subset of the species e.g. restricted-range or threatened species then the most efficient approach would be to apply a richness algorithm to this subset of species. The lack of concordance between restricted-range species richness and absolute species richness for Afrotropical antelopes may have been because all the sites in the database were conservation areas. These might have been established solely to protect and manage particular rare or threatened species and so may represent rather artificial distributions of species. In a more natural situation where sites are drawn from contiguous areas the sites with many common species might also contain most of the rare species.
Rarity II is a compromise between using richness and rarity as sites are scored not only on absolute number of species, but also according to how geographically restricted their fauna is, so that the sites containing species found in only that site are more likely to be selected early on, ensuring that there will be greater complementarity of widespread species later on in the selection. It might be expected that since Rarity II also takes account of site species richness as well as rarity, that it would represent more species than Rarity I in a limited number of sites. However this was only true for plant species in KwaZulu Natal. Initially sites containing unique species will dominate site scores in Rarity II and so sites containing unique species will be selected first (the same as in Rarity I). Only after the very restricted species have been represented will site scores reflect differences in species richness between sites. In the case of the antelope data, 20 species occurred in just one site. Three of these species occur in De Hoop which is why it is selected first by both rarity algorithms. The other seventeen do not occur together in any site, so eighteen sites are needed to include all species restricted to one site. These sites are selected first by both the rarity algorithms. Once these rarest species have been represented, the differences between the rarity algorithms become apparent. After 24 selections Rarity I had represented 75 species compared to 81 by Rarity II and after 30 selections Rarity I had 88 species compared to 95 in the Rarity II network. This reflects the increasing influence of species richness in the Rarity II selection once the most restricted species have been represented. Rarity II did represent more plant species than the other rarity algorithms when the number of sites selected was limited to just five percent of the area, but there was very little difference between the algorithms when it came to representing bird and mammal species in KwaZulu Natal.

How close do iterative algorithms come to finding the optimum solution for a reserve network (as found by an integer programming solution)?

Iterative selections cannot guarantee to find the optimal configuration of sites because they cannot foresee that some species might be selected at later iterations (Rebelo 1994). However Underhill’s (1994) criticism of iterative algorithms as providing ‘grossly suboptimal results’ does not appear to be valid. The results from the Oregon study indicate that rarity algorithms come very close to the minimum number of sites needed
to represent all species, and even the so-called greedy richness heuristic only took four sites more (27 compared to 23) than the integer programming solution. This may be an important difference when it comes to the acquisition of sites for conservation, however factors such as land availability and economic considerations will be much more important at a practical level. This means that any site selection method needs to be flexible and be able to adjust to such constraints. Whilst it is possible to build such factors into integer programming solutions, programs can take many days to run compared to iterative algorithms which can process the same data in seconds (Possingham et al. 1993, Pressey et al. 1996).

Saederstal et al. (1993) applied a rarity heuristic and a linear programming technique to the selection of woods in Norway with the aim of representing (a) all 47 bird and (b) all 321 plant species. Both the rarity heuristic and the linear programming solutions could represent all birds in twelve sites and all plants in 32 sites, out of a total of 60 woods. However the size of the woodlands varied. The 32 sites selected by the integer programming technique covered 71.4% of the total area compared to the 32 rarity algorithm sites which covered just under 75% of the total woodland area. The twelve bird sites represented 27.9% of the total area under the integer programming solution and 40% of the area for the iterative method. This latter difference occurred due to the inclusion of one very large site by the iterative algorithm, and the majority of birds were represented in around 25% of the area. The differences in the total area selected by the two methods in this study may just be sampling differences because it does not appear that the linear programming formulation included any constraint to minimise the area of woodland selected, only to minimise the number of woods. If this was the case then the rarity heuristic performed as well as the integer programming method.

Pressey et al. (1996) have also compared the performance of iterative algorithms with linear programming solutions. Selecting sites to represent all land types in New South Wales, Australia, they found that use of a ‘good’ iterative algorithm required just 4 additional sites (57 compared to 53) to represent all 248 land types over an optimal solution. In terms of the area required, the iterative selection included about ten percent
more land relative to the optimal solution.

The amount by which an iterative algorithm selection deviates from the optimal solution will depend on the numbers of species and sites in the analysis as well as patterns of richness and complementarity. If the majority of species are narrowly distributed, occurring in just one or a few sites in total, then the sites selected by a rarity algorithm (and to a lesser extent richness algorithms) and the optimal selection are likely to be very similar (Pressey et al. 1996). This is because sites containing unique species have to be selected if all species are to be represented and this will force the selection methods to converge. This explains why all iterative selections took the same number of sites to represent all plant species and selected the same set of sites.

Similarly richness and rarity algorithms will produce similar results where there is not much complementarity between sites as was the case with the bird and mammal data sets. In these examples many of the restricted-range species are found in the most species rich sites. Also when a richness algorithm is used the first site selected picks up a large amount of the regional species richness (67% of the birds and 62% of the mammals) after a few selections most of the unrepresented species are restricted-range species and so these determine the rest of the sites selected.

Iterative algorithms are a useful tool for identifying networks of areas for biodiversity because they highlight the relevance and implications of every addition or removal of a site from the network (Rebelo 1994). They also indicate which sites are flexible and which are essential for a particular biodiversity goal, and the algorithm rule can be changed to accommodate desired features - for example multiple species representations, threatened species, representation of all habitat types or maximising the connectivity of sites in the network.
CONCLUSIONS.

- Rarity based iterative algorithms were more efficient than iterative richness algorithms at representing all species, although the differences were very small.
- When insufficient sites to represent all species could be selected, richness algorithms were generally more efficient at representing overall numbers of species, although this was dependent on the taxon.
- Fewer restricted-range antelope species were represented by the first sites selected by the richness algorithm compared to the rarity algorithms, so in certain cases use of a richness algorithm could compromise the representation of rare species.
- An integer programming solution was more efficient at representing all attributes compared to all the iterative algorithms, however in most cases the difference in efficiency was very small (e.g. some of the rarity algorithms only took one more site than the optimal solution). Iterative algorithms do not generally appear to produce ‘grossly suboptimal’ results, and have advantages in terms of requiring less computer time to process the data.
CHAPTER SIX.

GENERAL DISCUSSION.

A major aim of conservation today is the maintenance of biodiversity. This will require prioritisation of species and areas, since constraints on land use and costs prohibit the conservation of all areas and species (Usher 1986, Gaston 1994, Pressey 1994, Williams et al. 1996a, Williams et al. 1996b, Pressey & Tully 1994). At a practical level, maintaining a representative selection of biodiversity will necessitate the prioritisation of areas on the basis of their biodiversity value (Usher 1986, Pressey & Nicholls 1989a). However, the biodiversity value of an area depends on the overall conservation goal, and it is often not obvious how to practically implement a given goal. For example, protecting a representative sample of the current biotic diversity (where diversity can have a variety of different meanings e.g. species richness or taxonomic diversity), safeguarding species with traits that may be correlated with susceptibility to extinction (e.g. endemics or restricted-range species) and protecting those species that are currently categorised as under short term threat of extinction are all valid conservation goals. However, priority areas for conservation may vary depending on which of these three approaches is taken. The challenge is to determine how to reconcile different goals and values, and to minimise conflicts between them (Gaston 1994). This requires the development of quantitative methods which force people to make values explicit, and which enable the representation of biodiversity as efficiently as possible (Williams et al. 1996b).

In this thesis the designation of priority areas using different approaches was investigated for Afrotropical antelopes and for birds, mammals and plants in South Africa, in an attempt to derive some general rules concerning the representation of biodiversity in a region. Three major problems face the designation of priority areas for the conservation of biodiversity. First, there is the question of what to conserve and the extent to which different aspects of diversity can be represented in the same sites. Second, is the problem of how to measure the value of areas given that data will not be available for all taxa.
Third, once areas have been evaluated, what is the best method for selecting areas to achieve a given conservation goal?

When sites were selected on the basis of different biodiversity attributes there was often little agreement between the priority sites selected using the different methods. Sites selected by each approach were also sometimes poor at representing the diversity components ranked highly by other approaches. For Afrotropical antelopes and for plants in KwaZulu Natal, top sites for species richness were poor at representing threatened species, so for these two groups a focus on preserving species rich areas could result in the loss of species in the short term. In contrast, species rich sites for birds and mammals in KwaZulu Natal proved to be good at incorporating threatened species, although in the case of south African birds, many of the threatened species occurring in the most species rich sites were only nationally and not globally threatened. Therefore, selecting sites on the basis of simple species richness will often not be sufficient to represent those species that are most threatened with extinction in the short term.

An alternative approach would be to prioritise areas on the basis of the threatened species that they contain. Sites selected for threatened species were universally reasonable at representing species richness, although were not as good as site networks selected for species richness. This would suggest that areas rich in threatened species are an appropriate focus for conservation, and that this would not bias the representation of ‘wholesale’ biodiversity. However, accurate information on the likelihood of short term extinction may not be available for all taxa or areas. One way around this problem might be to prioritise areas on the basis of species that are rare in some way, since rare species may be more vulnerable to extinction (Gaston 1994, Mace & Kershaw in press).

A restricted geographical range represents one form of rarity, and restricted-range species may be more susceptible to threatening processes than widespread species (ICBP 1992). Areas rich in restricted-range species were good at representing all diversity components for all the taxa in this study. However, in this study, restricted-range was defined at the scale at which the site evaluation was made. In the case of the KwaZulu
Natal data set this was a political rather than biogeographical unit. Species that were rare in KwaZulu Natal might have been widespread elsewhere and vice versa. Networks selected using endemic species richness (i.e. species that are globally restricted-range but not necessarily restricted-range within KwaZulu Natal) were poor at representing all other diversity components. The general rule here appears to be that richness of species that occur in only a few sites within the region being evaluated may be the most efficient way to prioritise areas in order to represent a large proportion of the regional biodiversity.

Given that not all components of diversity can be represented in the same set of sites it is important for any selection procedure that precise aims and their consequences are understood. A combination of different approaches, emphasising different aspects, may be the best compromise for preserving a full range of biotic diversity.

Application of systematic methods to the selection of priority areas for conservation requires the use of suitable surrogates for overall biodiversity as well as decisions about what to conserve. The attributes which can be conserved in a small subset of sites, and the identity of suitable surrogates will be largely influenced by the patterns of spatial distribution exhibited by the different taxa in a region. Taxa that are the best surrogates for encompassing a representative range of diversity across a variety of taxa can be predicted by looking at the relative importance of the different components of diversity (species richness and spatial turnover) for different taxa in the region. Taxa that show little spatial turnover or complementarity between sites are likely to be poor surrogates for diversity in other taxa. In contrast, high turnover taxa are more likely to be good surrogates for diversity across a range of taxa. In KwaZulu Natal, birds and mammals show little turnover. Consequently, they were poor surrogates for diversity in higher turnover groups. Plants show much greater species turnover between sites and were better surrogates for diversity across a range of taxa. Amongst birds, those that are habitat specific are better surrogates for plant diversity, but endemic species are poor surrogates for overall diversity, being concentrated mainly in the west in high altitude areas. Maximising environmental representativeness, ensures that sites are
complementary, but will not necessarily pick up many species because the sites may be relatively species poor.

Once sites have been evaluated for biodiversity value there is the question of how this information should be used to select conservation areas. Selection of sites that have high scores for a given criterion (e.g. species richness) does not necessarily take account of species complementarity between sites. Top scoring sites may contain similar suites of species and collectively not represent many of the region's species. If good information on species distributions is available then iterative selection procedures that take account of complementarity between sites as well as site richness can be used. An iterative selection of sites was considerably more efficient at representing all attributes than selection of the top-scoring sites for a variety of taxa. When insufficient sites to represent all species could be selected iterative selections incorporated more species than top scoring site networks. However the efficiency of an iterative selection procedure was reduced when some areas were already part of the reserve network, particularly if those sites were of overall low diversity. Iterative selections went some way to solving the problem of different components of diversity not co-occurring. For example, iterative selections were better at representing threatened and endemic species than purely selecting the top scoring sites for species richness.

Iterative area selection algorithms that initially select species that occur in few sites ('rarity' algorithms) were consistently more efficient at representing all species than richness algorithms. However, even for large data sets (with thousands of species) the differences in efficiency were not great, and in reality it is likely that constraints on land acquisition will be a much more of a problem than slight differences in the efficiency of alternative algorithms. For a data set of Oregon vertebrates rarity algorithms came close to achieving the minimum number of sites to include all species as found by an integer programming solution. A richness algorithm was not far from being optimal at representing species richness in a subset of sites. Considering there is the need for flexibility in reserve selection, an iterative approach that achieves close to the optimum site configuration and which allows rapid data processing (and therefore rapid
examination of different networks) is a useful conservation tool (Pressey et al. 1996).

Since none of the criteria for selecting reserves maximises all aspects of biodiversity under all circumstances it is necessary to be clear about the objectives of a reserve network when deciding on a method for site selection.

**Data limitations and recommendations for further work.**

Most studies of the spatial relationships between taxa at the scales used in this thesis are hampered by a lack of data and serious biases in the consistency of the data across sites. Bird distribution data, collected as part of a long term atlassing project (as in the bird data used for KwaZulu Natal), constitute some of the best available data at national scales. Nonetheless, there will still be biases in the data, for example, some grid squares are more inaccessible than others and therefore are visited less frequently and by fewer observers. Similarly some species are more likely to be overlooked for example, due to secretive behaviour, because they occur in certain habitat types (e.g. forest compared to grassland), because they are easily confused with other species, because they are rare in some way (e.g. low abundance), or even because they are common and therefore may be ignored. The problem of collecting intensity is far worse for other taxa like plants. Although KwaZulu Natal is one of the better collected areas in southern Africa in terms of plant specimens, there were still quarter degree grid squares in the region that in 1984 had fewer than ten specimens collected, whilst other squares had more than two thousand records (Gibbs-Russell et al. 1984). Assessment of the relative value of different sites for biodiversity will need to account for collecting intensity (Nelson et al. 1990).

One way around these data deficiencies might be to model species distributions using information on known climate and habitat relationships (e.g. Austin 1991b, Busby 1991). The data on vertebrate species distributions in Oregon were incomplete and a method of interpolating the probable presence of a species between confirmed records was used to generate the distribution maps (Csuti et al. in press).

Another potential problem with studies documenting patterns of diversity is determining
the appropriate scale for evaluation. Choosing the wrong biogeographic scale at which to evaluate and select sites could result in the conservation of local biodiversity patterns at the expense of regional or global biodiversity. This is a major problem since data are often only consistent at political units and because most conservation action will take place within a political framework. There are also valid arguments for conserving local elements of biodiversity, such as representing local variation in diversity and protecting species which local people value (Hunter & Hutchison 1994). One way around this problem might be to first identify those elements of biodiversity within the political unit which are globally or regionally important and then apply the kinds of quantitative selection methods described in this thesis to those elements only. For example, within KwaZulu Natal, site selection could be based around those species that are globally threatened or are true endemics. In this way areas of relatively low species richness, but which are centres of endemism and support large numbers of threatened species (for example, the Drakensberg Mountains) would be given high priority.

Designation of priority areas for the conservation of biodiversity does not mean that other areas are not valuable for conservation (Gaston 1994). The successful representation of biodiversity will also require consideration of landscape level conservation. Priority sites cannot be viewed in isolation from the surrounding habitat matrix since this will lead to fragmentation of the landscape, and additionally discrete sites may not be sufficient to conserve some elements of diversity, for example, widespread or migratory species. Methods need to be developed to address the issue of wider scale conservation, otherwise a focus on individual sites could cause degradation of areas not given conservation priority.

Obviously it will not be possible to conserve the whole of a quarter degree grid square (660km²) - the unit of site evaluation used in the South African analyses. Many similar studies have used areas that in practice could not be set aside purely for conservation (e.g. Prendergast et al. 1993). Given that fine resolution data may not be available it is therefore necessary to identify those elements of diversity that are most important within the large areas that are selected as high priority by the quantitative methods, and within
these areas, focus on conserving the important elements. There is also the problem of how does the way the data are organised influence the areas that are selected? The selection of conservation areas may be sensitive to the spatial scale at which biodiversity value is measured (Stoms 1994). For example if the quarter degree grid were shifted 300km east in KwaZulu Natal how would this affect the sites selected? Similarly if the analysis were carried out on eighth degree or half degree grid squares do the areas selected change dramatically? In terms of evaluating potential areas for conservation it will be desirable to use a scale that maximises the between-site variation in biodiversity value. Further work is necessary to determine the influence of such scale effects on site evaluation procedures (Pressey & Logan 1994, 1995).

Conservation of biodiversity is an evolving process which is improved as better information becomes available. Unfortunately there is not enough time to wait until detailed information for all taxa is obtained. The challenge is therefore to find objective and explicit methods for designing networks of areas to represent biodiversity that can be used by those involved in the practical aspects of area selection.
LITERATURE CITED.


covering location problem. *Biological Conservation.*


of Six Rare Bird Species in the Kruger National Park, South Africa. *Bird Conservation International* 6, 117-137.


south-western Cape Province of South Africa. *Journal of Biogeography* 18, 509-518.


McIntyre, S. (1992). Risks associated with the setting of conservation priorities from rare
plant species lists. *Biological Conservation* 60, 31-37.


