This is a repository copy of *Using presence-absence data to establish reserve selection procedures that are robust to temporal species turnover.*

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/1341/

---

**Article:**

DOI:10.1098/rspb.2000.1087

---

**Reuse**
Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher’s website.

**Takedown**
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
Using presence–absence data to establish reserve selection procedures that are robust to temporal species turnover

Ana S. L. Rodrigues1*, Kevin J. Gaston1 and Richard D. Gregory2†

1Biodiversity and Macroecology Group, Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK
2British Trust for Ornithology, Nunnery Place, Thetford, Norfolk IP24 2PU, UK

Previous studies suggest that a network of nature reserves with maximum efficiency (obtained by selecting the minimum area such that each species is represented once) is likely to be insufficient to maintain species in the network over time. Here, we test the performance of three selection strategies which require presence–absence data, two of them previously proposed (multiple representations and selecting an increasing percentage of each species’ range) and a novel one based on selecting the site where each species has exhibited a higher permanence rate in the past. Multiple representations appear to be a safer strategy than selecting a percentage of range because the former gives priority to rarer species while the latter favours the most widespread. The most effective strategy was the one based on the permanence rate, indicating that the robustness of reserve networks can be improved by adopting reserve selection procedures that integrate information about the relative value of sites. This strategy was also very efficient, suggesting that the investment made in the monitoring schemes may be compensated for by a lower cost in reserve acquisition.

Keywords: biodiversity; complementarity; effectiveness; efficiency; monitoring; turnover

1. INTRODUCTION

Maintaining the diversity of biological features in a given region over the long term is the prime objective in the establishment of a network of nature reserves. Indeed, concern about the likely persistence of features in reserves has always been a feature of practical reserve selection exercises. This has often been expressed by focusing resources on a few target species, usually the most endangered ones (e.g. Thomas 1991; Madsen et al. 1998), with the expectation that, in maintaining them, other features of interest will be preserved as well. However, it has been demonstrated that, at least in some situations, this approach may not be effective in conserving non-target species (e.g. Kerr 1997).

One approach to addressing the simultaneous concerns about persistence and the conservation of a multitude of species has been to use reserve selection procedures which select the best sites on the basis of an index that incorporates one or several variables thought to be of importance (so-called scoring procedures). The appeal of such indices is that they can easily integrate a diversity of concerns, including biological, social and economic ones. The long-term persistence of features has usually been addressed by focusing resources on one or more of the species of interest (e.g. Gotmark et al. 1986; Brown et al. 1995; Turpie 1995). The drawback to these procedures is the risk that the areas selected are unnecessarily duplicating some attributes while leaving other features of interest totally unrepresented in the reserve network (Pressey & Nicholls 1989).

Reserve selection strategies based on the complementarity principle (Vane-Wright et al. 1991) have been developed in response to the recognition that resources for the protection of biodiversity will always be limited. These aim at a high efficiency (sensu Pressey & Nicholls 1989) in the representation of all biodiversity features of interest with the minimum cost. The most commonly used of these procedures searches for the minimum area such that all features are represented at least once in the reserve network (e.g. Margules et al. 1988; Sætersdal et al. 1993; Couti et al. 1997; Pressey et al. 1997; Howard et al. 1998). The underlying rationale is that features cannot be protected by reserves if they do not occur in the network in the first place (Margules et al. 1988).

However, this simple representation strategy may not be sufficient to assure the long-term persistence of features in the network. Indeed, previous studies have found that some years afterwards minimum sets failed to retain all the species that justified their selection (36% of species lost during an 11-year interval from the minimum set of limestone pavements that represented each plant species once, UK (Margules et al. 1994), 16% of species lost during a 63-year interval from the minimum set of lakes that represented each plant species once, Finland (Virolainen et al. 1999), and an average of 8% species loss during ten-year intervals from the minimum number of census plots that represented each bird species once, UK (Rodrigues et al. 2000a)).

Different strategies for improving the robustness of complementarity-based reserve selection procedures have been suggested in the literature. Making use of abundance data, Kershaw et al. (1994) and Turpie (1995) proposed that only those sites where species occur above a certain abundance value should be considered for selection, while Nicholls (1998) proposed the establishment of a minimum population size as a required
representation target for each species. Furthermore, the results obtained by Rodrigues et al. (2000a) suggested that more robust networks can be obtained by selecting the sites at which each species occurs at relatively high local abundance. Proposed strategies based on presence–absence data include multiple representations (e.g. Pressey & Nicholls 1989; Lombard et al. 1995; Williams et al. 1996) and the representation of all species in a minimum fraction of their range in the study area (e.g. Nicholls & Margules 1993; Pressey & Tully 1994; Pressey et al. 1997). A more elaborate approach proposed by Williams (1998) consists of excluding records for particular species in areas where their viability seems likely to be poor as assessed using niche-based modelling of the local habitat suitability.

Here we consider three families of reserve selection strategies based on presence–absence data: multiple representations, selecting an increasing percentage of each species’ range, and selecting the sites where species exhibited a higher permanence rate in the past. Considering species as features of biodiversity, we used data from the Common Birds Census (CBC) in the UK to examine how these strategies affect the efficiency and effectiveness (a measure inversely related to the gap between the representation target required and the one attained by the network; Rodrigues et al. 1999) of reserve networks in maintaining species over time in comparison with a single representation strategy.

2. DATA AND METHODS

The CBC, which is run by the British Trust for Ornithology, has been the primary scheme by which populations of common breeding birds have been monitored in the UK (for a comprehensive description of the history and methodology of the CBC, see Marchant et al. (1990)). Although it provides information on the abundance of each species in each site, in this analysis we used presence–absence data only.

We used the CBC data collected between 1976 and 1991 in a variable number of farmland and woodland sites. We considered six pairs of years with a ten-year interval in between: 1976–1986 through to 1981–1991. For each pair, we analysed only those sites with good quality information in both years and which had been visited at least twice in the previous five years. Only those species for which presence–absence had been recorded in all years (77 species) were considered.

Three families of reserve selection strategies were tested.

(i) Single and multiple representations. Single representation: select the minimum area such that each species is represented in at least one site (a). This corresponds to the most commonly used complementarity-based approach in the recent reserve selection literature. Multiple representations: select the minimum area such that each species is represented in at least n sites (or the maximum number of sites where the species occurs if less than n): (b) n = 2, (c) n = 3, (d) n = 4 and (e) n = 5.

(ii) Percentage of range. Select the minimum area such that each species is represented in at least p percentage of its range in the study area: (a) p = 10%, (b) p = 20%, (c) p = 30%, (d) p = 40% and (e) p = 50%. As an approximation, the range of each species was given by the total area of sites where the species occurs.

(iii) Permanence rate. A permanence rate is calculated for each species in each site, given by the fraction of years in which the species was recorded at the site in relation to the total number of years in which the site was visited in the period between five years before and the first year of a pair of years separated by a ten-year interval (e.g. for the 1976–1986 interval, the permanence rate for each species in each site is given by the number of times the species was recorded at the site from 1971 to 1976 in relation to the total number of visits in those years). Select the minimum area such that each species is represented at least at the site or one of the sites where it has the higher permanence rate registered.

In each pair of years, the first year’s data were used to select a reserve network following a specific strategy. The last year’s data were then used to evaluate the network according to (i) its efficiency (Pressey & Nicholls 1989), which is higher when the total area occupied is smaller, and (ii) its effectiveness (Rodrigues et al. 1999) over time, which is higher when the fraction of species absent from the network ten years afterwards (the representation gap) is smaller (throughout this paper, we use ‘effectiveness’ in this sense).

The optimal solution was found in each case and then four near-optimal solutions. This was done by first determining the optimal solution and then solving the problem after adding an additional constraint that excludes the optimal solution previously found (Camm et al. 1996; Rodrigues et al. 2000b). In this way, the optimization algorithm finds another optimum (if it exists) or the nearest best solution. By repeating this procedure, a sequence of five solutions with non-decreasing areas (but all close to the minimum) was obtained for each problem. The average area and average representation gap of those five solutions was obtained. The average area is still very close to the optimal value while the average effectiveness is a value that is more representative of the performance of a given strategy and less likely to have been determined by chance.

For each pair of years, the average efficiency and effectiveness were also determined for 100 randomly selected networks of (approximately) a pre-defined area for each pair of years. This was done by selecting sites randomly without replacement until the total area was approximately that predefined (obviously, given variability in the areas of sites, random networks will seldom be precisely this area). We applied this procedure to a wide range of areas in order to establish a null relationship between efficiency and effectiveness (random model).

All minimization problems were solved exactly as integer linear programming problems (see Rodrigues et al. in 2000b) using CPLEX (ILOG 1999).

The selection units considered in this study are census plots, which are much smaller than most nature reserves. Although this implies that the turnover rates observed are probably much faster than the ones occurring in reserves during a ten-year interval, turnover is also known to increase over time (Russell et al. 1995) and reserves are expected to prevail far longer than ten years. It was assumed that the turnover rates observed within ten-year intervals in the CBC plots exhibit similar patterns to the ones observed in reserves over longer periods (see Rodrigues et al. 2000a).

Although the CBC concerns species that are ‘common’ in the UK, many are rare in this data set (e.g. in 1981 nearly 40% of the species had a range of less than one-quarter of the total study area). The CBC data are used here as an exemplar data
set in order to explore general reserve selection strategies and, therefore, the results should not be interpreted as an attempt to propose a new reserve network in the UK.

3. RESULTS

For each pair of years, the corresponding random model illustrates the expected effectiveness of a network of a given area selected randomly for each pair of years (figure 1). As predicted, the larger the area, the lower the representation gap ten years afterwards (in the limit when all sites are selected a representation gap of zero is obtained). This has consequences for the interpretation of the results of the strategies tested—just by increasing the area (lowering the efficiency) of a network a higher effectiveness is anticipated.

By definition, the maximum efficiency in a reserve network is achieved by the minimum set that represents every species at least once (i.e. the single representation strategy, strategy (i(a))). However, this strategy always resulted in large representation gaps (figure 1). Increasing the minimum number of representations required for each species (strategies (i(b)), (i(c)), (i(d)) and (i(e))) always required the selection of a larger reserve network area but generally resulted in a considerably higher effectiveness of the networks obtained (figure 1). In particular, increasing the number of representations from one to two always resulted in a reduction of the gap to less than half its initial value (although requiring on average more than twice the area).

The general pattern of results obtained for the family of strategies involving the selection of some percentage of the range of each species is similar to that of the single and multiple representation family—a higher representation target always corresponds to a lower efficiency and generally to a higher effectiveness (figure 1). However, in
Figure 2. The relationship between the local permanence rate (fraction of years in which a species was recorded at the site in relation to the total number of years in which the site was visited between 1976 and 1981) and the subsequent extinction probability (the probability of the species becoming extinct in that site between 1981 and 1991). Likelihood ratio test: \( \chi^2 \text{-test, } n = 2205 \), \( \chi^2 \text{-test} = 304.661 \) and \( p < 0.001 \). The dot sizes refer to the number of correspondent data values (as there is only a small number of possible values for the permanence rate). Only data on the 1981–1991 pair of years has been used to prevent non-independence caused by overlapping time-series.

some cases an inversion was obtained in the pattern of decreasing effectiveness, i.e. an increasing representation gap in spite of the larger area (figure 1a,c,e,f).

Selecting sites with a high permanence rate consistently resulted in networks with maximum effectiveness (zero representation gap) and a high efficiency (always higher than the one obtained by the strategy of representing each species twice) (figure 1).

4. DISCUSSION

The results obtained in this study are consistent with previous findings (Margules et al. 1994; Virolainen et al. 1999; Rodrigues et al. 2000a) in suggesting that a single representation strategy for selecting reserve networks is not sufficient for ensuring the maintenance of species in the long term. A high level of efficiency is attained at the cost of low effectiveness. However, the results obtained using the three families of strategies tested indicate that, by compromising some efficiency, it is possible to obtain reserve networks that are more robust to species temporal turnover.

A multiple representation strategy seems to be a safer investment than one based on a percentage of area. The results in the first case—a general tendency to a decrease in the representation gap when the target is increased (figure 1)—are as expected since, by requiring multiple representations, there is a higher probability that each species will persist over a ten-year period in at least one of the selected sites. On the other hand, some of the results found in the family of strategies based on percentage of range are apparently counter-intuitive, as they suggest a significant risk of obtaining a simultaneously less efficient and less effective network when increasing the representation target (figure 1a,c,e,f). These contrasting results are probably a consequence of the fact that the first family of strategies places more emphasis on rare species while the second in fact gives priority to the most widespread ones. For a multiple representation strategy, a target of \( n = 3 \) sites, for example, means selecting all of the sites that are occupied by species which occur in a total of three or fewer sites and a decreasing proportion of the total number of sites occupied by each of the more widespread species as this number increases. On the contrary, a target of \( p = 30\% \) for a percentage of range strategy means requiring very large areas for very widespread species and very small areas for very rare ones. When the goal is to generate networks that are robust to temporal species turnover, the first approach is expected to perform better, since rare species are known to be more prone to local extinction (Gaston 1994; Rodrigues et al. 2000a) and may therefore require a higher investment in order to persist in a reserve network. The inversions observed in the results for a percentage of range strategy (figure 1a,c,e,f) perhaps follow from a reduction in the total area allocated to some of the rarer species when increasing the required representation target. For a small target (say 10\%), only one or two sites at which they occur need to be included for both rare and common species to be sufficiently represented. For increasing targets, most of the sites added are likely to be required to meet the target for the most common species, as the rare ones will still achieve the percentage of area required with representation at only one site. If, when increasing the target, some sites that contain rare species but that are unnecessary for achieving their representation target (which happened to have been selected previously) are replaced by others that contribute only to representing more common species, the probability of extinction of these rare species in the network increases and may result in a lower effectiveness despite the larger area being selected.

A practical problem when using the multiple representation strategy is knowing the adequate degree of replication needed in order to attain a high effectiveness without unnecessarily compromising the efficiency of the network. In practice, the adequate target for each species may have to be decided on a case-by-case basis, according to the information available and the specific goals established for the network.
A drawback of the multiple representation strategy is that it considers all sites where a species is present to have the same value for its persistence. Although the risk is lower when setting higher representation targets, there is a danger in this approach that all sites of a network where the species is represented are inadequate for its long-term survival while the best sites are left unprotected (Turrpie 1995). Strategies that target sites where species are more likely to persist are therefore expected to perform better in terms of long-term effectiveness (Williams 1998). Indeed, in this data set the strategy based on permanence was the most effective (figure 1), which can be explained by a lower local extinction probability in the sites with a higher permanence rate (figure 2).

Deciding which sites should preferentially be selected for each species can be based on abundance data, as in Rodrigues et al. (2000a), or on presence–absence information about species persistence over time, as here. Both types of information are often unavailable to planners and can be expensive to obtain. Obtaining an accurate estimate of the abundance of each species in each site at a given time requires a substantial investment when compared with the one needed to obtain presence–absence data. For example, the time input required for a full CBC analysis to obtain the number of breeding pairs of each species in each plot is estimated to be at least 3.5 times that of presence–absence data (11 versus four days per plot per year to carry out the fieldwork and analysis; presence–absence data could be obtained with less investment, but with greater uncertainty about which species use a site). In a related study, Gregory et al. (1994) estimated the time input required for a full CBC to be 6.9 times that of using point counts (55 versus 8 h per plot per year).

Permanence rates require having information on the presence–absence of species in each of the sites in a series of years, which involves a period of monitoring before deciding which reserves to select. Nevertheless, the collection of presence–absence data requires less expertise and may be more attractive to volunteers (Bart & Klosiewski 1989). In addition, it may be easier to recruit volunteers to work more intensively over some years than to concentrate the same investment in a single year in order to collect abundance data.

Whether using abundance or time-series of presence–absence data, there is an additional cost attached to directing surveys at obtaining information about the best sites for the conservation of each species over and above that of simply determining the spatial occurrence of species. However, including this information in the reserve selection procedures may result in both higher effectiveness and higher efficiency (i.e. lower cost) in the implementation of the network. Indeed, in our analyses we found that it was a better strategy to select the best site for each species than to invest in multiple, but blind, redundancy (figure 1). In practical reserve selection exercises, the gain from using more information is avoiding the cost of acquiring unnecessarily large reserve systems and may well compensate for the resources invested in the monitoring schemes needed to acquire that information (Balmford & Gaston 1999).

K.J.G. is a Royal Society university research fellow. A.S.R. is funded by a Portuguese FCT/PRAXIS studentship under the Science and Technology Sub-Programme of the Second Community Support Framework. We thank the thousands of volunteers who have carried out the CBC surveys on which the bird population estimates were based. The CBC is funded by a partnership between the British Trust for Ornithology and the Joint Nature Conservation Committee. Two anonymous referees made helpful comments on the manuscript.

REFERENCES


ILOG 1999 CPLEX 6.5. Gentilly, France: ILOG.


