Threatened and endemic species: are they good indicators of patterns of biodiversity on a national scale?

Aletta Bonn*, Ana S. L. Rodrigues and Kevin J. Gaston
Biodiversity and Macroecology Group, Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK
*Correspondence, E-mail: a.bonn@sheffield.ac.uk

Abstract
Endemic and/or threatened species are often targeted to set conservation priorities. It is tempting to assume that a reserve network focusing on these species will be an effective umbrella for overall species richness of a country. For South Africa and Lesotho we tested whether complementary networks selected for threatened and/or endemic bird species satisfactorily represent all bird species, both in terms of capturing areas where other species are present or areas where they are more abundant (and, presumably, more viable). We found that areas selected for threatened and endemic species perform considerably better than areas selected at random. However, they do not guarantee the representation of overall bird species diversity, particularly not in peak abundance locations. Although nationally threatened and endemic species are important conservation targets, our results indicate that reserve networks focusing solely on these species may not be sufficient to preserve overall species diversity in a country.

Keywords
Complementary networks, conservation planning, endemism, peak abundance locations, species richness, threat.

INTRODUCTION
At the global scale, several schemes have been employed for identifying areas that may be particularly important for the long-term maintenance of biodiversity. As decision criteria, these schemes have variously used data on patterns of species richness, endemism, threat or taxonomic uniqueness of species, as well as habitat features (for a recent advocacy of the first three criteria as forest biodiversity indicators by the United Nations Environment Programme see Kapos et al. 2001). They have led to the recognition of, for example, biodiversity hotspots (Mittermeier et al. 1998; Myers et al. 2000), centres of plant diversity (Davis & Heywood 1994–97), endemic bird areas (Bibby et al. 1992; Balmford & Long 1994; Stattersfield et al. 1998), and the most valuable ecoregions (Olson & Dinerstein 1998). To varying degrees, such schemes have influenced thoughts and actions in the placement and designation of new protected areas.

The majority of these global prioritization schemes have focused particular attention on the patterns of occurrence of threatened and endemic species. The rationale is simple. The former group, by definition, are already at a disproportionate risk of loss in the near future, and thus their persistence requires immediate conservation action. Endemic species may also be threatened (because of their small extant geographical ranges, and the low numbers which tend to be associated with this; Leigh 1981; Goodman 1987). They will commonly be at risk from stochastic events, and give rise to much of the regional or local ‘flavour’ of species assemblages. In addition, a pragmatic view is that threatened and endemic species may serve best to attract public attention as well as funding (Williams et al. 2000).

In prioritizing areas for conservation at the national scale, the focus on threatened and endemic species has commonly been retained (see Smith & Theberge 1986 for a review). Thus, schemes based on these criteria have been explored for, e.g. the United States (Dobson et al. 1997; Abbitt et al. 2000), Greece (Troumbis & Panayotis 1998), India (Daniels et al. 1991), the Nicobar islands (Sankaran 1997), Namibia (Barnard et al. 1998) and South Africa (Drinkrow & Cherry 1995).

A common presumption, shared by both global and national schemes for determining priority areas for conservation based on the occurrence of threatened and/or
endemic species, is that these areas will also prove appropriate to protect most of the other species within the target taxon. At global and large biogeographical scales, some consideration has been given to how realistic this might be (e.g. Bibby et al. 1992). At the national scale it has received rather more limited attention (but see Williams & Humphries 1994; Williams et al. 1996). Indeed, it has been more common to evaluate how well species richness hotspots also represent rare and/or endemic species (Prendergast et al. 1993; Cofré & Marquet 1999; but see Lombard 1995; Dobson et al. 1997; Reyers et al. 2000; for across-taxon evaluation of hotspots and complementary networks).

This paucity of national scale studies of the wider effectiveness of networks chosen on the basis of threatened or endemic species is particularly problematic, because the concepts of threat and of endemism do not necessarily translate readily from the global to the national scale (Gärdenfors 2001). This is so for a number of reasons including: nationally threatened species may not be threatened globally; globally threatened species may not be amongst the most threatened nationally; species that are narrowly endemic on a global scale may not be endemic to any one nation; and species that are endemic to a country may not be globally narrow endemics.

In this paper, using data for the birds of South Africa as a case study, we determine how effective priority area networks selected for (a) threatened species (b) endemics or (c) a combination of the two are in representing all species in the assemblage, comparing the effects of using either presence/absence and abundance data. This is an analysis on a national scale, which means that, for the aforementioned reasons, the results obtained do not necessarily translate into conclusions applicable at the global scale.

**DATA**

The Southern African Bird Atlas Project (Harrison et al. 1997) has provided the most comprehensive information currently available on the distribution of birds in Southern Africa. Several previous studies have used this data set as a basis for planning studies (e.g. Reyers et al. 2000; Fairbanks et al. 2001). Data were mainly collected between 1987 and 1992, at a spatial resolution of a quarter-degree grid for Lesotho, Namibia, South Africa, Swaziland, and Zimbabwe and on a half-degree grid for Botswana. Observers visiting each cell recorded the presence of identified species on checklists, with breeding and non-breeding records being considered equivalent. A variable number of checklists were submitted for each cell ($\mu = 39, \sigma = 88.0$). Based on these data, reporting rates were calculated for each species in each cell as the proportion of checklists submitted for that cell on which the species was recorded. We used the data for South Africa and Lesotho (1858 cells), excluding marine, vagrant, marginal and introduced or escaped species from the analysis (651 species).

For each species, peaks of reporting rates were defined as cells with reporting rates $\geq 80\%$ of the maximum value observed for that species. We assumed a positive association between peaks of reporting rates and peaks of each species’ relative abundance, as observed by Robertson et al. (1995) for Namibian birds. These peaks of abundance corresponded on average to $5.8\%$ of the total number of records for each species.

Three groups of species were distinguished (hereinafter called target species).

1 Threatened species. One-hundred and three bird species in South Africa and Lesotho are either globally ($n = 33$) or nationally threatened ($n = 97$) (Barnes 2000). Nine of the globally threatened bird species were not listed in the national Red Data Books. Among them, the Taita Falcon (Falco fasciinucha) is globally rare and has only a single breeding pair record in the study region, whereas the other eight species are endemic to South Africa and Lesotho and more abundant. Three of them are relatively common and widespread within this region, e.g. the Ground Woodpecker (Geocolaptes olivaceus) and the Buffstreaked Chat (Oenanthe bifasciata), whereas others have a restricted distribution, such as the Mountain Pipit (Anthus boeschi) or the Drakensberg Siskin (Pseudochloroptila symnasi). Most of the latter, however, are not considered as nationally threatened even with small populations, because of a relatively secure conservation status or inaccessibility of their habitat (Harrison et al. 1997).

2 Endemic species. Fifty-nine species were classed as endemics or near endemics to South Africa and Lesotho (Barnes 2000), of which $37\%$ are also threatened. Barnes (2000) defined near endemics as those species with at least $90\%$ of their breeding range confined to South Africa, Lesotho and Swaziland.

3 Threatened and endemic species. In total, threatened and endemic species comprised 140 species.

Three nationally threatened and three endemic subspecies listed by Barnes (2000) were not included in the analysis because the records of 1987–92 (Harrison et al. 1997) did not differentiate between their ranges and those of more common subspecies.

**ANALYSES**

**Area selection for the target species**

Complementary networks (Pressey et al. 1993; Cabeza & Moilanen 2001) were obtained representing the target
species groups in sets with the minimum possible number of areas. First, area selection was performed using presence/absence data, with representation targets of one, two, three and four occurrences (or the maximum possible) of each species (see Rodrigues et al. 2000a). Second, sets of sites were selected using the reporting rate data, with representation targets of one, two, three and four peaks of abundance (or the maximum possible) of each species. Because each approach has a number of possible optimal solutions, a set of 10 randomly selected optimal solutions was obtained for each of the above problems.

**Evaluation of complementary networks performance**

The effectiveness of each network (selected for threatened species, endemic species, or both) in the representation of all bird species was evaluated using two different measures. First, we determined the overall number of bird species represented (out of 651) in each network using presence/absence data. This is the most traditional measure of effectiveness (Rodrigues et al. 1999), but it may be misleading if the species are only represented at sites that are inadequate to ensure their persistence (Gaston et al. 2001). As a second measure of effectiveness, we took viability considerations into account and determined the number of species represented in at least one of their respective peaks of abundance. The underlying assumption is that peak abundance locations indicate more suitable habitat for the species under conservation concern (but see Pulliam 1988) and their selection provides a higher effectiveness in retaining species over time (Rodrigues et al. 2000b).

**Evaluation of further scenarios**

The results of these area selections were evaluated against three other scenarios.

1. **Null model I – random selection of sites.** The first scenario does not incorporate any information on species’ distributions. For a given number of sites (1–140) randomly selected without replacement, we determined (i) the number of species represented in at least one locality or (ii) in at least one peak abundance location. Each random selection was replicated 100 times.

2. **Maximum covering location problem (MCLP, Church et al. 1996).** This scenario assumes perfect knowledge about all species’ distributions. For a given number of sites, we selected the network that maximizes either (i) the number of species represented in at least one locality or (ii) in at least one peak abundance location.

3. **Null model II – reserve networks selected for sets of randomly selected species.** This scenario aims at providing insight into whether the networks selected to represent the target species perform better in representing overall species richness in relation to randomly selected sets of the same number of species (i.e. 59, 103 and 140 species). Here, 10 replicates were obtained using 10 independently selected sets of randomly selected species for each case.

All problems were solved optimally with C-PLEX software (ILOG 2001).

**RESULTS AND DISCUSSION**

The results for South African birds indicate that focusing network selection on threatened and/or endemic species will not guarantee the representation of all other species (Figs 1–3). However, quite different relative performances were obtained depending on which scenarios the networks are compared with, the target group used in network selection, and whether the results were evaluated in terms of simple species representation or representation in peaks of abundance.

**Performance of networks selected for target species in relation to a random site selection**

Networks selected for the three target groups tended to perform significantly better in representing overall species numbers than a random site selection (null model I; Figs 1–3). This confirms that it is better to have some information on species distributions, even if simply presence/absence, than no data at all (Gaston & Rodrigues in press). The notable exceptions were the area networks for endemic species with respect to representation of species’ occurrences (Fig. 2a): especially when representation targets were set to two or more representations, a random site selection gave similar or even better results.

**Performance of networks selected for target species in relation to the MCLP**

Networks selected for the three target groups did not perform as well as the best possible site selection (MCLP), which used information for all species. This is particularly so when networks selected based on occurrence data were evaluated in terms of representing species in peaks of abundance (Figs 1–3b).

**‘Indicator capacity’ of networks selected for target species**

The total number of species represented by networks selected for the target groups was always significantly higher than the corresponding number of target species (Figs 1–3).
This is true both for representation at occurrence (Figs 1–3a) and at peak abundance locations (Figs 1–3b). For example, the area network which represented the 103 threatened species at least once in one peak abundance location captured on average the peak abundance locations of 430 species, i.e. 327 species for which no information was used (Fig. 1b).

Figure 1: Performance of networks selected targeting at 103 threatened species in terms of the overall number of species represented in (a) occurrence and (b) peak abundance localities. Networks selected using information on threatened species are represented in filled symbols: circles for networks representing each species at least 1, 2, 3 and 4 times (from left to right); triangles for networks representing each species at least 1, 2, 3 and 4 times (from left to right) in peaks of abundance. These networks’ performance was evaluated against: (i) null model I, sites selected at random (grey line without symbols); (ii) the results for the maximum covering location problem which includes information for all species (grey line with crosses); (iii) the results for networks selected using information on sets of 103 randomly selected species (open symbols; circles for networks representing each species at least 1, 2, 3 and 4 times; triangles for networks representing each species at least 1, 2, 3 and 4 times in peaks of abundance). For the networks selected using information on the threatened and the randomly selected species, error bars provide the standard deviation for 10 possible solutions or 10 replicates, respectively. Confidence intervals for null model I are indistinguishably small. For reference, horizontal lines indicate the total number of species (651; dashed) and the number of species tested (103; solid).

Figure 2: Performance of networks selected targeting at 59 endemic species in terms of the overall number of species represented in (a) occurrence and (b) peak abundance localities. Symbols as in Fig. 1, but here reserves based on randomly selected species (open circles) used information on 59 species, and the horizontal solid line corresponds to this number of species also.

This is true both for representation at occurrence (Figs 1–3a) and at peak abundance locations (Figs 1–3b). For example, the area network which represented the 103 threatened species at least once in one peak abundance location captured on average the peak abundance locations of 430 species, i.e. 327 species for which no information was used (Fig. 1b).

Perhaps surprisingly, however, networks designed for the same number of randomly selected species represented overall species diversity as well as or even better than the target species groups (Figs 1–3). This was particularly so for peak abundance locations (Figs 1–3b), where randomly selected sets of species predicted overall bird species diversity much better than the threatened (Fig. 1b) and, especially, than the endemic (Fig. 2b) species. For example, the proportion of species represented in peaks of abundance...
increased from 41% to 60%, when area selection was targeted at at least one peak abundance location of randomly selected species instead of endemic species (Fig. 2b). Most likely, these results are a consequence of the fact that both endemic and threatened species do not constitute a representative sample of the diversity of birds in South Africa. Instead, they are biased toward particular habitat types (see below), which probably affects their indicator capacity for all species. By contrast, a randomly selected set of species is more likely to be a representative subset and therefore a better indicator across all species.

It may be tempting to consider that these results demonstrate a poor indicator capacity from the target species, and to suggest that randomly selected species should be used instead. However, although data on a broad subset of species (across all taxa, habitat associations and range sizes) are likely to be very valuable for nature conservation, in practice biological data do not tend to be available for random sets of species, but rather for particular species groups or habitats. More importantly, selections based on random sets of species are biased toward under-representing the threatened and endemic species, which are conservation priorities in their own right. For example, the networks selected to represent 59 randomly chosen species at least once, included 86% of all bird species but only 69% of the endemics and 66% of the threatened species. The difference was smaller but still significant if a greater number of randomly chosen species is used for area selection ($n = 140$, representation: 91% all species, 78% endemic species, 75% threatened species) or when using information on peaks of abundance for area selection.

**Threatened vs. endemic species**

Networks targeting threatened species (Fig. 1) always captured more overall species diversity than networks designed using information on endemic species (Fig. 2). Thus, a network targeted at threatened species in one occurrence location represented 96% of all species, in comparison to 63%, when targeted for endemics. Naturally, the better the information input into the area selection approach, i.e. the higher the number of species used (59 in Fig. 2, 103 in Fig. 1 and 140 in Fig. 3), the closer the results will be to the best possible site selection (MCLP). Keeping in mind this statistical relationship, networks targeted at threatened species (Fig. 1) nevertheless performed disproportionately better than those targeted at endemic species (Fig. 2). This is especially apparent when evaluated for peak abundance representation (Figs 1–3b).

Why is this so? Centres of endemism and threatened species richness do not necessarily coincide with those of overall species diversity (Prendergast *et al*. 1993; Williams & Humphries 1994; Lombard 1995; Gaston 1996). Comparison of maps of the richness of all bird species, threatened species and endemic species in South Africa and Lesotho show differing regional peaks for each group (Fig. 4; for further vertebrate taxa in this region see also Lombard 1995). In general, the north-east of South Africa is very species-rich (e.g. Kruger National Park), which may be attributed mainly to climatic conditions and vegetative cover (van Rensburg *et al*. 2002). Species richness is also high along the coast and within the fynbos biome in the south. However, at least part of the distribution of peaks of threatened and endemic species is explained by the presence of the South Africa boundary, particularly for endemic species.

It is expected that the more species that exist in a region the higher the number of those that are likely to be threatened, and indeed centres of richness in threatened species correlated reasonably well with overall species
richness (Fig. 5a, Table 1), as also found in previous studies (Rebelo 1994; Williams et al. 1996; Cofré & Marquet 1999). However, some nationally threatened species (by definition, often with restricted ranges) tend to concentrate near the boundaries if their ranges extend only marginally into the countries (e.g. Dobson et al. 1997; see Rodrigues & Gaston 2002 for a discussion). Indeed, the peak of threatened species in South Africa is located on the eastern border with Mozambique, which may reflect the subtropical subtraction effect, i.e. the southern range extension of tropical species into north-eastern South Africa (Lombard 1995).

On the other hand, owing to the superimposition of a geopolitical scale onto a biogeographical concept, nationally endemic species tend by definition to be located away from countries’ boundaries (a species whose range is adjacent to a boundary is likely to occur also in the neighbouring country, therefore not being an endemic). This is clearly the case in our analyses, as bird endemism in South Africa is highest in the Cape region in the south and the Drakensberg escarpment in the south-east. Depending on whether those areas away from the boundaries are species-rich or not, peaks of endemism may or may not be well correlated with peaks of richness, although biological gradients may contribute to strengthen or weaken this pattern. This may help to explain apparently contradictory results which found either a coincidence (e.g. Cowling & Samways 1995 and references therein; Kerr 1997) or a dissimilarity (Williams & Humphries 1994; Cofré & Marquet 1999; Kessler et al. 2001) between regions of high species richness and regions with high endemism. In South Africa, peaks of endemism do not correspond to the richest habitats, which, as explained above, are mainly on the north-eastern part of the country. As a result, the pattern of richness of endemic species relates very poorly with that of overall species.
Although patterns of congruence (or rather the lack of it) in richness, rarity and endemism hotspots do not necessarily predict complementarity patterns of priority area networks (Lombard 1995; Reyers et al. 2000), the distortions caused by the presence of an artificial geopolitical boundary may help to explain why neither threatened nor endemic species provide an ‘indicator capacity’ for overall species diversity as strong as a random set of species (Figs 1 and 2), as well as the better relative performance of threatened species (Fig. 1). On the other hand, that the bias for threatened species is in the opposite direction to the one for endemic species, may explain why a network selection that targets both groups performs relatively well (Fig. 3).

Peak abundance vs. occurrence data

Performance of priority area networks for the target species was much better regarding occurrence representation (Figs 1–3a) than peak abundance representation (Figs 1–3b) of all species. Thus, for example, an area network representing each of the threatened species in at least one occurrence locality, on average also included occurrence records of 96% of all species, but peak abundance records for only 17% of these species. So, although the occurrence representation rate might suggest a successful network selection, the latter figure casts doubts on viability at chosen sites.

With respect to performance of overall species representation in occurrence locations, there was no marked difference between networks selected to represent the target species in occurrence locations (circles, Figs 1–3a) and in peak abundance locations (triangles, Figs 1–3a). However, the networks selected using presence/absence data performed very poorly when evaluated against peaks of abundance of overall species, whereas networks selected

Table 1

<table>
<thead>
<tr>
<th></th>
<th>All species occurrence</th>
<th>All species peaks of abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r</td>
<td>P</td>
</tr>
<tr>
<td>Occurrence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Threatened spp.</td>
<td>0.82</td>
<td>0.001</td>
</tr>
<tr>
<td>Endemic spp.</td>
<td>0.17</td>
<td>n.s.</td>
</tr>
<tr>
<td>Threatened &amp; endemic spp.</td>
<td>0.61</td>
<td>0.000</td>
</tr>
<tr>
<td>Peaks of abundance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Threatened spp.</td>
<td>0.19</td>
<td>0.027</td>
</tr>
<tr>
<td>Endemic spp.</td>
<td>−0.04</td>
<td>n.s.</td>
</tr>
<tr>
<td>Threatened &amp; endemic spp.</td>
<td>0.06</td>
<td>n.s.</td>
</tr>
<tr>
<td>All spp.</td>
<td>0.19</td>
<td>0.001</td>
</tr>
</tbody>
</table>

©2002 Blackwell Science Ltd/CNRS
to represent abundance peaks for the targeted species groups tended to capture also the peaks of abundance of other species (Figs 1–3b). In this way, the proportion of species represented in a peak abundance location quadrupled from 17% to 66%, when area selection for threatened species was targeted at one peak abundance location instead of a simple occurrence location.

This may be explained by the fact that patterns of species richness do not necessarily coincide with patterns of peaks of species abundance (Table 1, Fig. 4). Thus, although occurrence data showed a very different distribution pattern between threatened or endemic species and overall species richness, their peak abundance locations both coincided to some degree with peaks of abundance of all species (Table 1). This congruence of peaks of abundance largely explains the difference in performance between networks selected using presence/absence data and using peaks of abundance data. The resulting networks may, however, be significantly larger (in our case, around four times larger) for a single representation target. Therefore, if viability concerns are to be addressed, larger and thus more costly priority area networks need to be accepted.

CONCLUSION

Preserving overall biodiversity is an explicit goal of the Convention on Biological Diversity (ISCBFD 1994), of which preserving species diversity is a necessary (albeit not sufficient) condition. While threatened and endemic species are important conservation targets in their own right, the results of the previous analyses indicate that reserve networks designed to preserve them do not guarantee that overall species diversity (even within the same taxon) will be preserved as a by-product. Given that usually there is not available information for all species, selecting priority areas for conservation by targeting endemic and/or threatened species may be a first step towards an effective reserve network. Awareness of the potential bias associated with these species may help in directing additional conservation efforts to other species/habitats in order to fill the potential gaps in the resulting reserve networks.

ACKNOWLEDGEMENTS

A.B. is supported by a postdoctoral DAAD research grant (German Academic Exchange Service) and A.S.R. by a Portuguese FCT/PRAXIS studentship under the Subprogramme Science and Technology of the 2nd Community Support Framework. We are grateful to L. Underhill and the Avian Demography Unit (University of Cape Town) for providing access to the data and to the numerous volunteers that contributed to its collection. J.J. Lennon provided software addressing spatial autocorrelation and P. Koleff gave advice on statistics. We also thank S. Jackson, O. Petechey, J. Rolff and three anonymous referees for valuable comments on the manuscript.

REFERENCES


Editor, A. Y. Troumbis
Manuscript received 14 June 2002
Manuscript accepted 16 July 2002

©2002 Blackwell Science Ltd/CNRS