Sensitivity of selection procedures for priority conservation areas to survey extent, survey intensity and taxonomic knowledge

Stefanie Freitag[1,2*] and Albert S. Van Jaarsveld[1]

[1]*Department of Zoology & Entomology, University of Pretoria, Pretoria 0001, South Africa*
[2]*Scientific Services, Kruger National Park, Private Bag X402, Skukuza 1350, South Africa*

Many procedures exist for identifying sets of sites that collectively represent regional biodiversity. Whereas the mechanics and suitability of these procedures have received considerable attention, little effort has been directed towards assessing and quantifying the effects of varying data inputs on their outcomes. In the present paper, we use sensitivity analysis to evaluate the impacts of varying degrees of (i) survey intensity, (ii) survey extent and (iii) taxonomic diversity on iterative reserve selection procedures. A comprehensive distribution database of the mammalian fauna from the Transvaal region of South Africa is systematically perturbed before implementation of a site selection algorithm. The resulting networks of sites are then compared to quantitatively assess the impact of database variations on algorithm performance. Systematic data deletions result in increased network variability (identity of selected sites), decreased numbers of frequently selected sites, decreased spatial congruence among successive runs and a rapid increase in the number of additional sites required to represent all species present in the region. These effects become particularly evident once data sets are reduced to below 20% of the original data. Consequently, a mixed survey strategy that balances survey effort with survey extent and maximizes taxonomic knowledge is more likely to ensure appropriate planning outcomes.

Keywords: sensitivity analysis; biodiversity; reserve selection algorithms

1. INTRODUCTION

Selection procedures for priority conservation areas provide the most efficient approach for sampling, sustaining and protecting regional biotas (Pressey *et al.* 1993; Margules *et al.* 1994; Williams *et al.* 1996; Pressey 1997). However, the ability to accurately sample regional biotas is influenced by our definition and measurement of biodiversity (Noss 1990; Pressey 1990; Pressey *et al.* 1993; Gaston 1996), and is largely data dependent (Nicholls & Margules 1993; Margules *et al.* 1994; Margules & Redhead 1995; Freitag *et al.* 1996; Haila & Margules 1996). On the other hand, pressures emanating from land transformation rates demand that existing biodiversity data be used as effectively as possible to make urgent conservation decisions (Davis *et al.* 1990; Lawton *et al.* 1994). Whereas additional biodiversity survey data can only be obtained at considerable cost (Burbridge 1991), the injudicious use of limited data sets could have extensive long-term conservation implications.

Iterative heuristic algorithms are widely used to derive priority conservation areas from information on the presence and/or absence, or the extent, of features such as species, habitat types and environmental domains (Pressey *et al.* 1993, 1996; Faith & Walker 1996). Differences in the outcomes of area selection procedures are a result of different data inputs (Freitag & Van

Jaarsveld 1995), the initial selection rules employed in different algorithms and the sequence in which selection rules are used in algorithms (Freitag *et al.* 1997). The latter components can be standardized within an algorithm, but biodiversity data inputs can vary in a number of ways: survey extent, survey intensity and taxonomic knowledge. Data interpolation techniques may be employed when the coverage of biodiversity surveys is inadequate (Austin & Margules 1986; Margules & Austin 1994; Margules & Redhead 1995) and species richness estimates may be derived from higher-taxon richness patterns (Balmford *et al.* 1996). Aggregate biodiversity scores such as species richness are, however, of limited value for priority area selection procedures (Pressey *et al.* 1993; Williams *et al.* 1996). Thus, there appears to be little prospect of useful substitutes for intensive surveys and improved taxonomic knowledge for increasing the planning value of regional biodiversity data, both of which can only be improved at considerable cost (Burbidge 1991; Cracraft 1995; Haila & Margules 1996; Lawton *et al.* 1998; Van Jaarsveld *et al.* 1998). Consequently, as many regions are considered to be 'data deficient' (Belbin 1993), more comprehensive and systematic global biodiversity surveys have been called for (Haila & Margules 1996; Van Jaarsveld *et al.* 1998).

An alternative pragmatic approach towards dealing with inadequate data sets is to quantify the risks associated with using limited data sets in priority area selection procedures. Sensitivity analyses can be used to

evaluate the reliability and robustness of priority area selection procedures by comparing the outputs derived from data that have been perturbed in a controlled and systematic manner (Lodwick *et al*. 1990; Stoms *et al*. 1992). In the present paper, we describe a sensitivity analysis procedure for priority area selection algorithms. An iterative 'near-minimum set' selection algorithm (Freitag *et al*. 1997) is used for the complementarity-based representation of mammalian fauna from the northeastern region of South Africa. The sensitivity of outputs to typical uncertainties inherent in biodiversity databases, namely survey extent, survey intensity and taxonomic knowledge, is assessed. The goal is to determine if priority area outputs seem valid over a range of reasonable assumptions about the nature of biodiversity data uncertainties (Stoms *et al*. 1992)

2. METHODS

(a) *Study area and databases*

The study covers northeastern South Africa (Transvaal region), which presently encompasses the provinces of Gauteng, Mpumalanga, Northern Province and part of North West Province. It is bordered to the north by Zimbabwe, to the west by Botswana, to the east by Moçambique, to the southeast by KwaZulu–Natal Province and Swaziland, and to the south and south west by Free State Province.

Large numbers of mammalian faunal surveys have been conducted in the region (Freitag *et al*. 1996), although sampling density is uneven (Freitag *et al*. 1998). Mammal data were collated from a wide range of sources and species' distribution records generalized to $15\,\mathrm{min} \times 15\,\mathrm{min}$ grid cells (*ca*. $25\,\mathrm{km}^2$; $n = 474$). These grid cells are the geographic sites used in the area selection and sensitivity analyses described below. The database includes distribution records for 192 discrete species and subspecies (see Freitag & Van Jaarsveld (1995) and Freitag *et al*. (1996) for species lists and data sources).

(b) *Iterative priority area selection algorithm*

A 'near-minimum' iterative priority area selection algorithm (described in Freitag *et al*. (1997)) was employed. The algorithm can be described as a 'weighted near-minimum set' selection algorithm rather than a more conventional rarity-based algorithm (Margules *et al*. 1988; Nicholls & Margules 1993). The algorithm prioritizes species in order of regional conservation importance using the following criteria: relative endemicity, taxonomic distinctiveness, vulnerability and regional occupancy (Freitag & Van Jaarsveld 1997; Freitag *et al*. 1997). Although such heuristic algorithms have been shown to be sub-optimal (Underhill 1994; Camm *et al*. 1996; Church *et al*. 1996), they do fulfil three important roles: they are indicative rather than prescriptive, have fast running times and can provide answers to proportional area problems (Pressey *et al*. 1997). In addition, the principle of complementarity, namely the contribution of unrepresented features by a site to a network of sites representing regional biodiversity (Pressey *et al*. 1993), is easily implemented. We apply the heuristic algorithm in a comparative and indicative role to assess the implications of varying three components of database quality on priority area outputs.

(c) *Sensitivity analyses*

The sensitivity of the iterative algorithm outputs to variations in (i) regional survey extent (assessed by the numbers of grid

Table 1. *Percentage exclusions conducted during sensitivity analyses and actual numbers of excluded grid cells, species and data records*

(Total numbers (n) of grid cells containing data, species for which data are available and data records within the distribution data set are also provided.)

% excluded	no. excluded grid cells	species	data records
<1	1	1	1
1	3	2	65
5	17	10	323
10	34	19	645
20	69	38	1290
40	138	77	2580
60	206	115	3871
n	344	192	6451

cells for which data exists), (ii) survey intensity (as reflected by the number of data records for the region), and (iii) taxonomic diversity (determined by the number of species for which survey data exists in the region) was determined by perturbing the database systematically prior to algorithm implementation. The input data set was perturbed by randomly deleting a specified percentage of the total number of data records ($n = 6451$), grid cells containing distribution data ($n = 344$) or species for which data exists ($n = 192$). Exclusions were systematically performed, beginning with the random exclusions of single ($<1\%$) data records, grid cells and species. The degree of random exclusion was increased in a stepwise manner to 1, 5, 10, 20, 40 and 60% of the total number of data records, grid cells or species, respectively (table 1). The priority area selection algorithm was subsequently implemented on the perturbed data set and this process repeated 500 times for each exclusion/deletion level.

The impacts of these data exclusions were evaluated in terms of the following.

1. The numbers of grid cells selected from the perturbed data sets to achieve complete representativeness, i.e. priority area network size and efficiency (Pressey *et al*. 1993).
2. The number of times that specific grid cells were selected in successive runs of the algorithm, defined in terms of RASI scores for individual grid cells. RASI is an acronym for 'relative algorithm specific index' (Freitag *et al*. 1997) and reflects the number of times a specific grid cell is selected, expressed as a percentage of the number of algorithm runs at a specific deletion level ($n = 500$).
3. The degree of spatial congruence (Gaston 1996; Van Jaarsveld *et al*. 1998) among a randomly selected subset of the selected networks ($n = 30$ of 500) per deletion level. Spatial congruence is expressed as the mean number of grid cells shared across all possible pairwise comparisons of the 30 randomly selected networks (from each set of 500 runs).
4. The mean number of additional grid cells required to complete species representation based on the non-deleted input data set after pre-selecting grid cells chosen from the perturbed data sets ($n = 30$ per deletion level).

3. RESULTS

Perturbations of the data set through grid cell, species or data record exclusions all indicate that increased data

Table 2. *Outcomes of sensitivity analyses for the random exclusion of grid cells, species and data records in terms of the numbers of grid cells used (i.e. the number of different grid cell identities), most common network size and size range across replicate runs (n=500)*

exclusions	% excluded	no. of grid cells used	most common network size	network size range	no. of runs
grid cells	0.3	44	23	22–25	344
	1	45	23	22–25	500
	5	66	23	20–26	500
	10	83	23	19–28	500
	20	107	23	18–29	500
	40	140	22	17–28	500
	60	183	21	15–29	500
species	0.5	36	23	22–26	192
	1	38	23	21–26	500
	5	48	23	19–26	500
	10	61	22	19–26	500
	20	81	21	15–24	500
	40	97	17	12–22	500
	60	115	13	9–18	500
data records	0.02	50	23	22–25	6451
	1	59	23	22–26	500
	5	91	24	18–28	500
	10	105	24	20–29	500
	20	131	24	19–29	500
	40	174	28	21–35	500
	60	208	33	27–40	500

exclusion levels lead to increased variability in the identities of grid cells selected in successive runs of the algorithm (table 2). Thus, the pool of grid cells from which priority area networks were chosen in the course of 500 random deletion replicates increased from 44 to 183 (grid cell deletions), from 36 to 115 (species deletions) and from 50 to 208 (record deletions).

Scatterplots of RASI scores for species exclusions (figure 1a) indicate that at low levels of species exclusion (<1%), a number of grid cells have very high RASI scores (>90%), i.e. they are selected repeatedly in the 500 algorithm runs. As the deletion level increases to 60%, the proportion of high-scoring grid cells is sharply reduced, whereas the proportion of low-scoring grid cells increases, i.e. large numbers of grid cells are selected infrequently during the 500 random replicate runs. Similar trends are shown for the exclusion of grid cells (figure 1b) and data records (figure 1c), although the shapes of these curves vary slightly. The decrease in the number of high-scoring RASI grid cells is marked for all three deletion scatterplots. The largest decrease in maximum RASI scores (from 100 to ±40%) was found for the species exclusion curve (figure 1a), and the least pronounced decrease (from 100 to 96%) in the maximum RASI score (figure 1c) for increased levels of data record deletions. The above trends are accompanied by an increase in the variability of the sizes of priority area networks selected by the algorithm (table 2 and figure 2).

Increasing grid cell and species deletion levels result in a decreased number of grid cells required to form representative networks (figure 2). Although this decrease is not as marked for grid cell exclusions (23 to 21), it is particularly evident with increasing species deletions, where mean network size decreases from 23 to 13 grid cells (table 2 and figure 2). However, data record

exclusion shows a reversed trend, namely, that average network size increases with increased record deletion (table 2 and figure 2), from 23 to 33 grid cells when record exclusion is increased from 0.02% to 60%.

The degree of spatial congruence among multiple networks shows a similar downward trend across increasing species, data record and grid cell deletion levels (figure 3). Grid cell congruence is very high (>80%) when deletion levels are not more than 5% of the database, but this decreases to between 30 and 40% at the 60% deletion level (figure 3).

Figure 4 shows the mean numbers of additional grid cells required to complete species representation (based on the entire undeleted data set) after pre-selecting grid cells already chosen from the perturbed sensitivity analysis data sets. It shows up to a sixfold increase in the mean number of additional grid cells required to complete representation when comparing the scenarios generated at 10% and 60% deletion levels. This change is greatest for the species deletion, i.e. reduced taxonomic information scenario (figure 4).

Collectively these results indicate that data deletion levels exceeding 10–20% rapidly increase network variability, result in less grid cells being consistently selected in algorithm runs (RASI score), cause reduced spatial overlap among successive runs and require increasing numbers of grid cells to eventually represent all species from the region.

4. DISCUSSION

These data suggest that the reduction of input data sets and the subsequent implementation of minimum set algorithms result in systematic variations in outcomes of priority area selection procedures. The response and
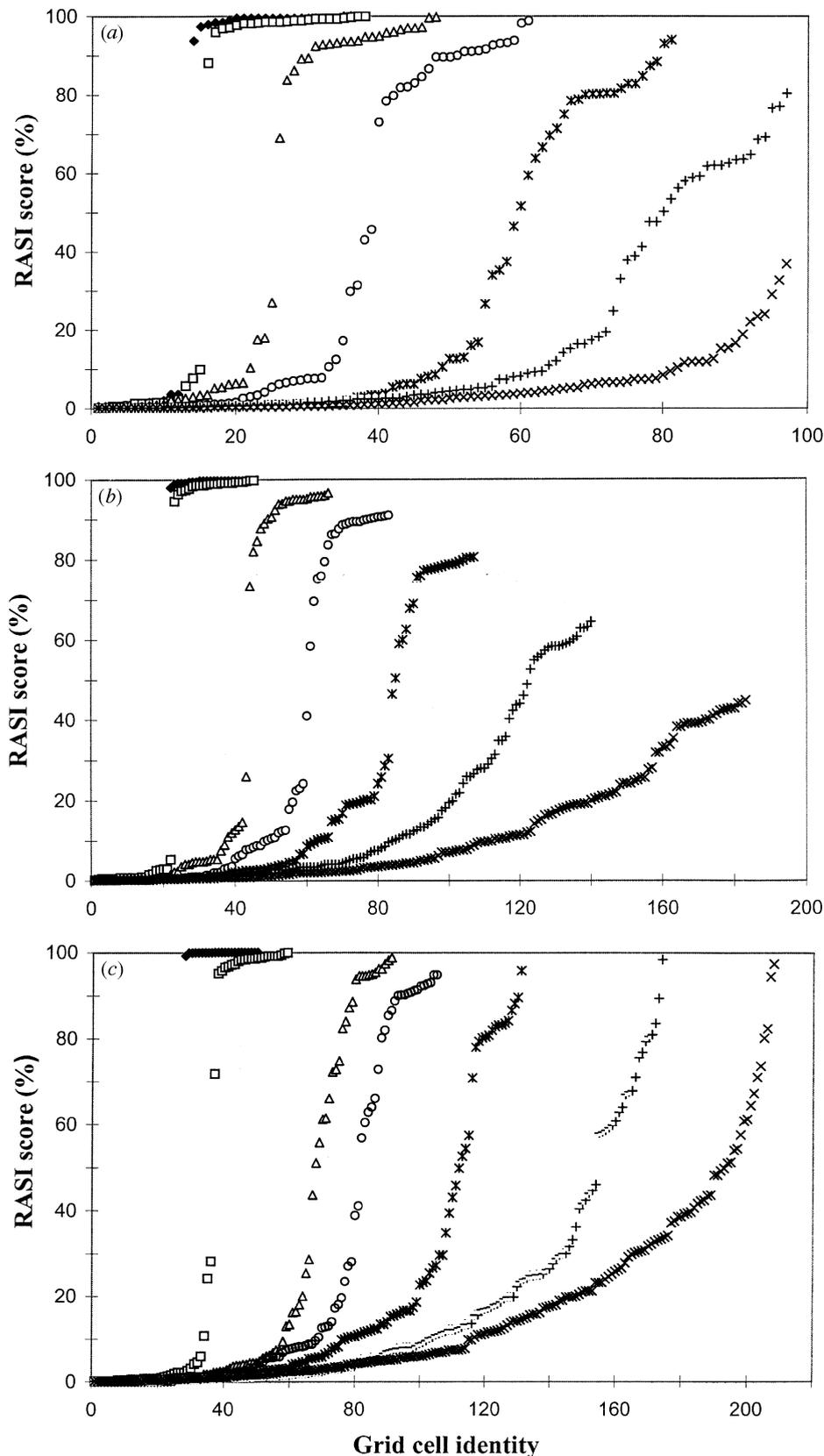
Figure 1. Scatterplots of the grid cell identities selected and their % RASI scores at seven levels of database perturbation. Deletion was for the random exclusion of <1% (filled diamonds), 1% (open squares), 5% (open triangles), 10% (open circles), 20% (asterisks), 40% (plus signs) and 60% (crosses) of (*a*) species, (*b*) grid cells and (*c*) data records.

sensitivity of priority area selection algorithms to systematic data set deletions are as follows.

(a) *Efficiency*

Efficiency refers to the principle of representing greatest biological diversity in the smallest cumulative area (Pressey & Nicholls 1989). There are three dominant efficiency-related trends that emerged from systematic perturbation of the input data set.

1. The removal of species from the input data set results in increased efficiency of priority area selection procedures (23–13 grid cells; figure 2). This is not unexpected,
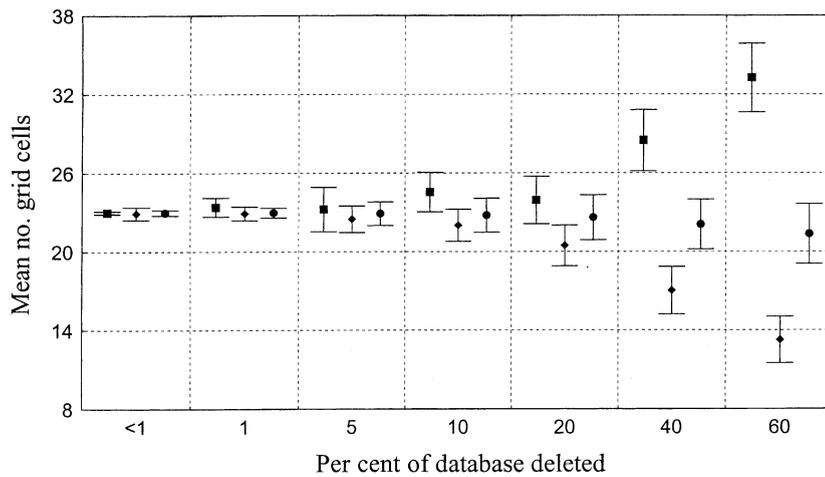
Figure 2. Plot of the mean number of grid cells (±s.d.) selected as representative protected area networks at different levels of database deletion for grid cells (filled circles), species (filled diamonds) and data records (filled squares).
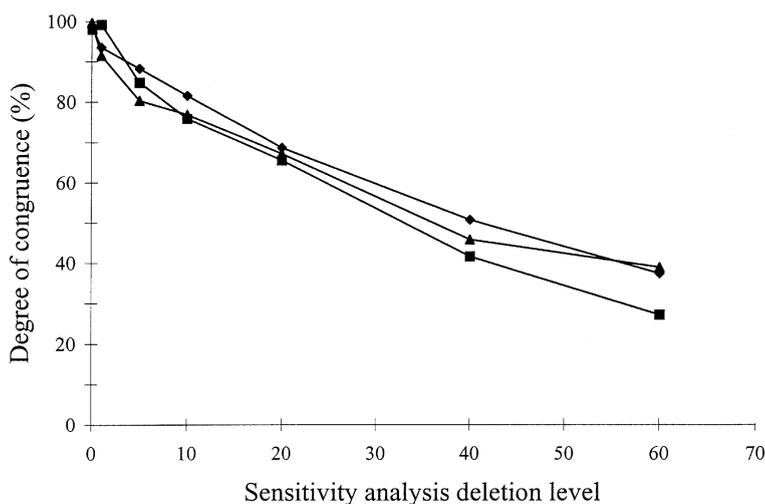

Figure 3. Degree of spatial congruence (in terms of % grid cells shared) among multiple networks selected at different deletion levels for grid cells (filled squares), species (filled diamonds) and data records (filled triangles).

because species deletions result in decreased taxonomic diversity, which requires the representation of less 'features' in the priority area network. Thus, logically, representation efficiency increases with a decrease in the number of 'features' to be represented. The corollary is that as the diversity of features increases, efficient land-use options become constrained (figure 2). This outcome will also be affected by the degree of geographic 'clumping', or scaling ratios (extent and grain (Wiens 1989)), contained in the data set.

2. The exclusion of grid cells (i.e. decreasing survey extent) has a negligible effect on the efficiency of priority area selection procedures. Such exclusions can be expected to have a variable effect on the number of species remaining in the database. This is due to the fact that a grid cell may represent only a fraction of the distribution range of a widespread species, but also the entire distribution range of a restricted-range species (Gaston 1994). Thus, three possible results may flow from decreasing the survey extent. First, the random removal of data for entire grid cells could lead to decreased taxonomic diversity in the data set by removing the only grid cell in which a species had been recorded. This effect will be pronounced if the data set contains many restricted-range species and would increase algorithm efficiency. Second, grid cell deletions could reduce the 'extent of occurrence' of species (Gaston

1994). Consequently, the number of restricted-range species in the database will increase and lead to decreased efficiency of species representation. This effect will, however, be diminished by geographic distributions of species that coincide. Thus, decreased survey extent may either increase or decrease the efficiency of priority area selection procedures depending on the characteristics of specific data sets. The third, and most likely, result of a decrease in survey extent is a combination of the above, i.e. some species will be deleted from the input database whereas others will display reduced 'extent of occurrences'. Collectively, these opposing processes may lead to negligible shifts in the efficiency of priority area selection procedures. This probably explains why the deletion of up to 60% of grid cells in the Transvaal region had limited impact on the efficiency with which mammalian taxa were represented (23–21 cells; figure 2).

3. Data record exclusions result in decreased efficiency of priority area selection procedures. Survey intensity is estimated by the number of data points within the region under consideration, and the sensitivity of priority area selection algorithms to changes in survey intensity can be assessed by removing data records, at random, from the input data set. Increased mammalian record deletions from the Transvaal region results in more grid cells being selected by the algorithm to achieve full representation
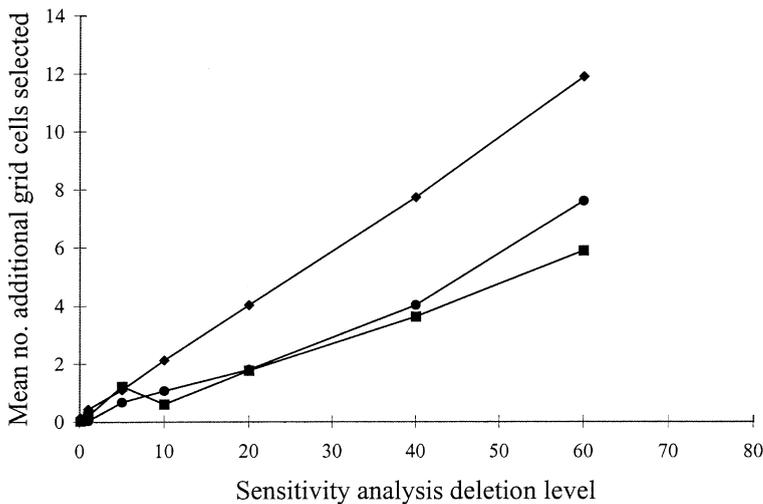
Figure 4. Mean number of additional grid cells required to complete species representation after pre-selection of grids selected in sensitivity analysis deletions for grid cells (filled circles), species (filled diamonds) and data records (filled squares).

(from 23 sites using the full data set to 33 when 60% of the original data records have been deleted; figure 2). This is a result of reduced species congruence, as a function of decreased extent of occurrence of many species, which results in less efficient application of the principle of complementarity.

(b) *Frequently selected grid cells* (*RASI scores*)

In response to systematic data set deletions, influential or frequently selected grid cells (with high RASI scores) become less important (figure 1) and this is accompanied by an increase in the variability of grid cell selection among successive algorithm runs (table 2 and figure 1). Decreased survey intensity (data record exclusion) has the largest impact on the selected grid cell variability; nevertheless, three grid cells are selected with RASI frequencies >80% (figure 1c) when the data record density is reduced by 60%. This is in sharp contrast with the low maximum RASI scores (<50%; figure 1a,b) obtained from algorithms run on databases reduced by 60% in terms of species and grid cells. This suggests that increased random selection of all grid cells occurs when biodiversity data sets suffer from limited survey extent and taxonomic knowledge (figure 1a,b). Limited survey intensity, on the other hand, results in some frequently selected grid cells (figure 1c), although it also results in a substantial increase in the mean number of grid cells required for reaching representation targets (figure 2).

(c) *Congruence and implementation*

Spatial congruence among multiple networks selected at the same deletion level shows a steady decline in the degree of network overlap with increasing data reduction (figure 3). Once more than 15% of the database is deleted, the level of spatial congruence drops below 70%, indicating that priority area selection at lower levels of database completeness becomes increasingly random and inconsistent, and is likely to show little congruence with ideal regional networks. This pattern holds irrespective of whether reductions in survey extent, survey intensity and degree of taxonomic knowledge are performed (figure 3).

The consequence of implementing a priority area network, based on incomplete data sets, either in terms of taxonomic knowledge, survey extent or intensity, on future network performances is illustrated in figure 4. Priority area selection using data sets restricted in terms of survey intensity will provide the most favourable 'stepping stone' towards achieving efficient representation at a future stage. Data sets deficient on this score require the least number of additional grid cells to be added to the implemented network at a future date (figure 4). Furthermore, the recognition of limited taxonomic knowledge in a regional biodiversity data set will impact most severely on network efficiency as improved taxonomic inputs significantly restrict additions to priority area networks through complementarity (figure 4).

(d) *Future survey effort and strategies*

The above trends have significant implications for survey strategy and design, particularly in completely unsurveyed regions. Although the gradsect survey method (Gillison & Brewer 1985; Austin & Heyligers 1991; Wessels *et al.* 1998) has been widely recommended as an efficient and effective survey strategy, a cost-related trade-off (Burbidge 1991) between numbers of areas surveyed and intensity of survey in each area will have to be reached. In other words, should a region be surveyed superficially (few indicator taxa) but extensively, or would more intensive survey efforts (many taxa) at fewer localities be more rewarding?

Results presented here suggest that the detection of taxonomic diversity is the most important and least compromisable component contained in regional biodiversity databases. Although low initial taxonomic detection leads to more efficient application of the principle of complementarity (figure 2), it also results in the greatest instability in selected grid cell identities (i.e. has the least number of high scoring RASI grid cells at high levels of sensitivity analysis deletion; figure 1). Consequently, it becomes more difficult to reach an efficient solution at a later stage once additional data become available (figure 4). In this respect, broadening the extent of surveys, whether in real terms or through using suitable interpolation techniques (Austin & Margules 1986; Margules & Austin 1994; Margules & Redhead 1995), should increase the amount of biodiversity (systematic/taxonomic diversity) detected as well as simultaneously increasing the extent of occurrence of species.

The increase in systematic diversity obtained from increasing the survey extent can be expected to be largest where large numbers of restricted-range species exist and where these ranges have limited overlap in space.

Limited survey intensity is easier to deal with as a shortcoming in biodiversity databases than limited taxonomic knowledge. Although this initially results in less efficient representation targets (figure 2), it is easier to reach an efficient solution once more data are available (figure 4), and there is some stability in selecting regionally important grid cells, even at low levels of survey intensity (figure 1). In terms of survey intensity, this should be conducted at some 80% effort. This recommendation flows from figure 2, where the effect of data record deletions on the mean number of selected grid cells is limited up to a 20% data deletion level. However, this effect is substantial once deletion levels of 40–60% are reached (figure 2). At high record deletion levels, the priority area selection procedure becomes sensitive to limited species distribution data as congruence decreases to below 55% (figure 3). In contrast, it appears that the number of input features (species) in these databases was not affected to a significant degree by data record deletions, as this would manifest itself in increased efficiency of representation. In some instances, however, the fine-scale delineation of species distribution data obtained through intensive sampling may not justify the extra inputs. In particular, this relationship will depend on the taxon/biodiversity units under consideration as well as on the geographic nestedness of features. In addition, the 'discovery' of cryptic species will result in the contraction of many ranges of extant species (Siegfried & Brooke 1994). This can be expected to impact on priority area networks, most likely resulting in a substantial reduction in efficiency when there is a greater than 20% increase in taxonomic detection (figure 2).

In conclusion, it is clear that irrespective of the methods used for priority area determination, there is an urgent need for the development of explicit systematic survey procedures (Haila & Margules 1996) that are both efficient and cost-effective (Burbidge 1991). However, over and above the development of adequate survey procedures, some consideration should be given to the relative spread of survey effort, i.e. should limited conservation inventory moneys be directed towards increasing survey extent, intensity or the detection of taxonomic diversity? From this sensitivity analysis conducted on mammalian taxa, it is clear that a subtle balancing act may be required. Although the biodiversity gains made by increasing survey extent and/or intensity are likely to be taxon- or region-specific, this study highlights the importance of conducting surveys of sufficient taxonomic depth while pursuing a mixed strategy that emphasizes both effort and extent. In addition, by conducting similar sensitivity analyses on different taxa or in different regions, alternative taxon-specific or region-specific survey strategies may emerge.

REFERENCES

Austin, M. P. & Heyligers, P. C. 1991 New approaches to vegetation survey design: gradsect sampling. In *Nature conservation: cost effective biological surveys and data analysis* (ed. C. R. Margules & M. P. Austin), pp. 31–36. Australia: CSIRO.

Austin, M. P. & Margules, C. R. 1986 Assessing representativeness. In *Wildlife conservation evaluation* (ed. M. B. Usher), pp. 45–67. London: Chapman & Hall.

Balmford, A., Green, M. J. B. & Murray, M. G. 1996 Using higher-taxon richness as a surrogate for species richness. I. Regional tests. *Proc. R. Soc. Lond.* B **263**, 1267–1274.

Belbin, L. 1993 Environmental representativeness: regional partitioning and reserve selection. *Biol. Conserv.* **66**, 223–230.

Burbidge, A. A. 1991 Cost constraints on surveys for nature conservation. In *Nature conservation: cost effective biological surveys and data analysis* (ed. C. R. Margules & M. P. Austin), pp. 3–6. Australia: CSIRO.

Camm, J. D., Polasky, S., Solow, A. & Csuti, B. 1996 A note on optimal algorithms for reserve site selection. *Biol. Conserv.* **78**, 353–355.

Church, R. L., Stoms, D. M. & Davis, F. W. 1996 Reserve selection as a maximal covering location problem. *Biol. Conserv.* **76**, 105–112.

Cracraft, J. 1995 The urgency of building global capacity for biodiversity science. *Biodiv. Conserv.* **4**, 463–475.

Davis, F. W., Stoms, D. M., Estes, J. E., Scepan, J. & Scott, J. M. 1990 An information systems approach to the preservation of biological diversity. *Int. J. GIS* **4**, 55–78.

Faith, D. P. & Walker, P. A. 1996 Environmental domains: on the best-possible use of surrogate data for assessing the relative biodiversity of sets of areas. *Biodiv. Conserv.* **5**, 399–416.

Freitag, S. & Van Jaarsveld, A. S. 1995 Towards conserving regional mammalian species diversity: a case study and data critique. *S. Afr. J. Zool.* **30**, 136–144.

Freitag, S. & Van Jaarsveld, A. S. 1997 Relative occupancy, endemism, taxonomic distinctiveness and vulnerability: prioritizing regional conservation actions. *Biodiv. Conserv.* **6**, 211–232.

Freitag, S., Nicholls, A. O. & Van Jaarsveld, A. S. 1996 Nature reserve selection in the Transvaal, South Africa: what data should we be using? *Biodiv. Conserv.* **5**, 685–698.

Freitag, S., Van Jaarsveld, A. S. & Biggs, H. C. 1997 Ranking priority biodiversity areas: an iterative conservation value-based approach. *Biol. Conserv.* **82**, 263–272.

Freitag, S., Hobson, C., Biggs, H. C. & Van Jaarsveld, A. S. 1998 Testing for potential survey bias: the effects of roads, urban areas and nature reserves in a southern African mammal data set. *Anim. Conserv.* **1**, 119–127.

Gaston, K. J. 1994 *Rarity*. London: Chapman & Hall.

Gaston, K. J. 1996 Biodiversity—congruence. *Progr. Phys. Geogr.* **20**, 105–112.

Gillison, A. N. & Brewer, K. R. W. 1985 The use of gradient-directed transects or gradsects in natural resource surveys. *J. Enviro. Mgmt* **20**, 103–127.

Haila, Y. & Margules, C. R. 1996 Survey research in conservation biology. *Ecography* **19**, 323–331.

Lawton, J. H., Prendergast, J. R. & Eversham, B. C. 1994 The numbers and spatial distributions of species: analyses of British data. In *Systematics and conservation evaluation* (ed. P. L. Forey, C. J. Humphries & R. I. Vane-Wright), pp. 177–195. Oxford: Clarendon Press.

Lawton, J. H. (and 12 others) 1998 Biodiversity inventories, indicator taxa and effects of habitat modification in tropical forests. *Nature* **391**, 72–76.

Lodwick, W. A., Monson, W. & Svoboda, L. 1990 Attribute error and sensitivity analysis of map operations in geographical information systems. *Int. J. GIS* **4**, 413–428.

Margules, C. R. & Austin, M. P. 1994 Biological models for monitoring species decline: the construction and use of data bases. *Phil. Trans. R. Soc. Lond.* B **344**, 69–75.

Margules, C. R. & Redhead, T. D. 1995 *BioRap: guidelines for using the BioRap methodology and tools.* Australia: CSIRO/ World Bank.

Margules, C. R., Nicholls, A. O. & Pressey, R. L. 1988 Selecting networks of reserves to maximise biological diversity. *Biol. Conserv.* **43**, 63–76.

Margules, C. R., Cresswell, I. D. & Nicholls, A. O. 1994 A scientific basis for establishing networks of protected areas. In *Systematics and conservation evaluation* (ed. P. L. Forey, C. J. Humphries & R. I. Vane-Wright), pp. 327–350. Oxford: Clarendon Press.

Nicholls, A. O. & Margules, C. R. 1993 An upgraded reserve selection algorithm. *Biol. Conserv.* **64**, 165–169.

Noss, R. F. 1990 Indicators for monitoring biodiversity. *Conserv. Biol.* **4**, 355–364.

Pressey, R. L. 1990 Reserve selection in New South Wales: where to from here? *Aust. Zool.* **26**, 70–75.

Pressey, R. L. 1997 Priority conservation areas: towards an operational definition for regional assessments. In *Selection and delimitation of parks and protected areas* (ed. J. Pigram & R. Sundell), pp. 337–357. Armidale: University of New England Press.

Pressey, R. L. & Nicholls, A. O. 1989 Efficiency in conservation evaluation: scoring versus iterative processes. *Biol. Conserv.* **50**, 199–218.

Pressey, R. L., Humphries, C. J., Margules, C. R., Vane-Wright, R. I. & Williams, P. H. 1993 Beyond opportunism: key principles for systematic reserve selection. *Trends Ecol. Evol.* **8**, 124–128.

Pressey, R. L., Possingham, H. P. & Margules, C. R. 1996 Optimality in reserve selection algorithms: when does it matter and how much? *Biol. Conserv.* **76**, 259–267.

Pressey, R. L., Possingham, H. P. & Day, J. R. 1997 Effectiveness of alternative heuristic algorithms for identifying indicative minimum requirements for conservation reserves. *Biol. Conserv.* **80**, 207–219.

Siegfried, W. R. & Brooke, R. K. 1994 Santa Rosalia's blessing: cryptic and under-recognized species in southern Africa. *S. Afr. J. Sci.* **90**, 57–58.

Stoms, D. M., Davis, F. W. & Cogan, C. B. 1992 Sensitivity of wildlife habitat models to uncertainties in GIS data. *Photogr. Eng. Rem. Sens.* **58**, 843–850.

Underhill, L. G. 1994 Optimal and sub-optimal reserve selection algorithms. *Biol. Conserv.* **70**, 85–87.

Van Jaarsveld, A. S. (and 10 others) 1998 Biodiversity assessment and conservation strategies. *Science* **279**, 2106–2108.

Wessels, K. J., Grimbeek, J. D., Van der Linde, M. J. & Van Jaarsveld, A. S. 1998 An evaluation of the gradsect biological survey method. *Biodiv. Conserv.* (In the press.)

Wiens, J. 1989 Spatial scaling in ecology. *Funct. Ecol.* **3**, 385–397.

Williams, P., Gibbons, D., Margules, C., Rebelo, A., Humphries, C. & Pressey, R. 1996 A comparison of richness hotspots, rarity hotspots, and complementary areas for conserving diversity of British birds. *Conserv. Biol.* **10**, 155–174.