Biological Conservation, **80** (1997) 207–219
© 1997 Published by Elsevier Science Ltd
Printed in Great Britain. All rights reserved
0006-3207/97 $17.00+.00

ELSEVIER

PII: S0006-3207(96)00045-6

EFFECTIVENESS OF ALTERNATIVE HEURISTIC ALGORITHMS FOR IDENTIFYING INDICATIVE MINIMUM REQUIREMENTS FOR CONSERVATION RESERVES

R. L. Pressey,[a]* H. P. Possingham[1b] & J. R. Day[2b]

[a]*New South Wales National Parks and Wildlife Service, PO Box 402, Armidale, NSW 2350, Australia*
[b]*Department of Applied Mathematics, University of Adelaide, Adelaide SA 5005, Australia*

(Received 17 August 1995; accepted 4 December 1995)

Abstract

We compared the results of 30 heuristic reserve selection algorithms on the same large data set. Twelve of the algorithms were for presence–absence representation goals, designed to find a set of sites to represent all the land types in the study region at least once. Eighteen algorithms were intended to represent a minimum percentage of the total area of each land type. We varied the rules of the algorithms systematically to find the influence of individual rules or sequences of rules on efficiency of representation. Rankings of the algorithms according to relative numbers or areas of selected sites needed to achieve a specified representation target varied between the full data set and a subset and so appear to be data-dependent. We also ran optimizing algorithms to indicate the degree of suboptimality of the heuristics. For the presence–absence problems, the optimizing algorithms had the advantage of guaranteeing an optimal solution but had much longer running times than the heuristics. They showed that the solutions from good heuristics were 5–10% larger than optimal. The optimizing algorithms failed to solve the proportional area problems, although heuristics solved them quickly. Both heuristics and optimizing algorithms have important roles to play in conservation planning. The choice of method will depend on the size of data sets, the representation goal, the required time for analysis, and the importance of a guaranteed optimal solution.

INTRODUCTION

A recent development in systematic reserve selection is the use of iterative algorithms that can identify minimum or near-minimum solutions, in terms of the number or area of sites, to the problem of representing all the targeted natural features in a region. They have been used most commonly in Australia (Kirkpatrick, 1983; Kirkpatrick & Harwood, 1983; Margules & Nicholls, 1987; Margules *et al.*, 1988; Margules, 1989; Pressey & Nicholls, 1989*a,b*, 1991; Pressey *et al.*, 1990; Kirkpatrick & Brown, 1991, 1994; Lewis *et al.*, 1991; Bedward *et al.*, 1992; Nicholls & Margules, 1993; Pressey & Tully, 1994; Pressey & Logan, 1995). The same approach is being increasingly applied elsewhere, for example in South Africa (Rebelo & Siegfried, 1990, 1992; Rebelo, 1994; Freitag *et al.*, 1996; Lombard *et al.*, 1995; Willis *et al.*, 1996), the United States (Ryti, 1992; Church *et al.*, 1996; Csuti *et al.*, 1996), Norway (Saetersdal *et al.*, 1993), and for world-wide or regional assessments of conservation priority by British scientists (Ackery & Vane-Wright, 1984; Vane-Wright *et al.*, 1991; Mickleburgh *et al.*, 1992; Vane-Wright & Rahardja, 1993; Williams *et al.*, 1991, 1992, 1996*a,b*; Kershaw *et al.*, 1994, 1995). Although not formalized as algorithms, the site selection methods trialled by Thomas and Mallorie (1985) included approaches that maximized the total number of species and the total number of restricted species in a subset of sites.

The heuristic algorithms used in these studies proceed stepwise, adding sites at each step that contain features most complementary to those in the sites already 'reserved'. Two main approaches have been used. Richness algorithms (e.g. Kirkpatrick, 1983) start with the site having the greatest number of unreserved features and then add sites one at a time according to which contains the most remaining unreserved features. Rarity algorithms (e.g. Margules *et al.*, 1988) begin with sites containing unique features and add sites progressively according to which contains the rarest unrepresented feature. Mixes of these two approaches are also possible. For example, Rebelo and Siegfried (1992) used an algorithm that progressively selected sites with the highest total rarity scores for all unrepresented species in each site, effectively combining rarity and richness.

[1]Present address: Department of Environmental Science, University of Adelaide, Roseworthy Campus, Roseworthy, SA 5371 Australia.
[2]Present address: Centre for Industrial and Applied Mathematics, School of Mathematics, University of South Australia, The Levels, SA 5095 Australia.
*Bob Pressey, Tel.: +61 67 737128; Fax: +61 67 722424; email: bpressey@ozemail.com.au

A basic need of both types of algorithm is the resolution of ties, for example when 10 sites each contain the rarest unrepresented feature. This can be done with a random choice, by taking the first on the list, or by calling a procedure or series of procedures to distinguish the sites on the basis of one or more criteria expected to lead to an efficient solution or to achieve some other goal. The 'best' rules, those that will most closely approach a minimum solution, cannot necessarily be identified *a priori* and the algorithms are heuristic in the sense defined by Parker (1984), i.e. 'operating in a way that is to some extent, unpredictable in advance, generally because... logical decisions... are made on the basis of quantities computed during the course of the algorithm'.

The features used as reservation targets in these analyses include species (e.g. Kirkpatrick, 1983; Rebelo & Siegfried, 1990; Ryti, 1992; Kershaw *et al.*, 1994), vegetation types (Margules & Nicholls, 1987) and landscapes (Pressey & Nicholls, 1989*a*). Elaborations on the basic goal of sampling each feature at least once include targets based on the probability of features occurring in sites (Margules & Nicholls, 1987), multiple representations (Margules *et al.*, 1988), quantitative representation targets such as a minimum percentage area of each land type (Pressey & Tully, 1994), combining species richness with measures of phylogenetic distance (Vane-Wright *et al.*, 1991), and taking into account the proximity of sites (Nicholls & Margules, 1993).

Networks of sites identified by heuristic algorithms represent the natural features of a region more efficiently, in terms of number or total area of sites, than those derived from more *ad hoc* decisions (Pressey & Tully, 1994) or those from systematic scoring approaches which do not operate iteratively (Pressey & Nicholls, 1989*b*; Lombard *et al.*, 1995). This efficiency gives them two important advantages as tools in conservation planning. First, they can make planners aware of the approximate mi nimum costs of conservation goals and therefore the feasibility of those goals. Second, the analyses can be used to compare reservation scenarios, for example by estimating the minimum reserve requirements if certain sites are made mandatory or unavailable for conservation, if sites with certain characteristics are given preference for selection, or if different data bases are used for the same region. These roles are all the more important because the algorithms are not difficult to program and can run quickly, even for very large data sets, on small computers.

The utility of heuristic algorithms lies in this ability to quickly identify indicative minimum reserve requirements. They do not, however, support all the decisions needed in conservation planning. They are limited in their ability to explore alternative configurations for reserve networks and, for this purpose, should be linked to more flexible systems (e.g. Williams *et al.*, 1991; Bedward *et al.*, 1992) or to geographic information systems (Pressey *et al.*, 1995). Unless applied repeatedly with different sets of sites as starting points (as by Rebelo & Siegfried, 1992) they also give no information on the potential 'value' or contribution of unselected sites to reservation goals. It can therefore be dangerous to assume that the selected sites are necessarily more valuable for nature conservation than unselected sites. The difference might simply be an artifact of the algorithm rules. Other approaches are being developed to indicate the potential contribution of all the sites in a region to a reservation goal (Pressey *et al.*, 1994).

Heuristic reserve selection algorithms are therefore potentially valuable indicative tools but provide only some of the needs of conservation planners, unless complemented by other analyses. Recently, even this indicative role has been questioned. Underhill (1994) criticized the use of heuristic algorithms in conservation planning because they could not guarantee optimal solutions to representation problems. The term 'optimal' is used in this paper to describe a solution to a reserve planning problem that consists of a minimum number or total area of sites, depending on how the reservation goal is framed. Lack of optimality (or 'suboptimality') has been demonstrated for real-world data sets in several comparisons of heuristics with branch-and-bound optimizing algorithms (Cocks & Baird, 1989; Saetersdal *et al.*, 1993; Pressey *et al.*, 1996; Csuti *et al.*, 1996), although heuristics can find optimal solutions at least in terms of number of sites (Saetersdal *et al.*, 1993; Willis *et al.*, 1996; Church *et al.*, 1996). Despite the lack of guaranteed optimality, heuristics still have some compelling advantages over optimizing algorithms for many problems (Pressey *et al.*, 1996). Running times for heuristics on large data sets are in the order of seconds or minutes rather than hours or days for optimizing algorithms. This makes them very suitable as components of real-time interactive systems. Furthermore, optimizing algorithms can fail to find the best solution, or any solution, to problems such as representing a percentage area of each land type, even after days or weeks of processing. Such problems are easily solved by heuristics, even if suboptimally. In the development of analytical tools for conservation planning, heuristics and optimizing algorithms are best seen as complementary rather than competing approaches (Pressey *et al.*, 1996).

Considering the potential importance of heuristic algorithms for conservation planning and their use for more than a decade, there have been few comparisons of different algorithms on the same data sets. The comparisons have been of two types. First, several studies have compared alternative analyses in terms of the efficiency of representation of natural features. Pressey and Nicholls (1989*b*) compared the efficiencies of a rarity and a richness algorithm on two data sets. Rebelo and Siegfried (1992) did the same type of comparison, with two different algorithms, on a data set from South Africa. Kershaw *et al.* (1994, 1995) used one data set to compare five very different algorithms according to efficiency, sequence of selected sites, and the species composition of notional reserves. The second type of comparison has involved considerations other than basic representation of features. One Australian and three South African studies have compared rarity algorithms with and without adjacency rules to improve reserve design (Nicholls & Margules, 1993; Freitag *et al.*, 1996; Lombard *et al.*, 1995; Willis *et al.*, 1996. In addition, Lewis *et al.* (1991) identified reserve requirements in Tasmania in relation to a variety of criteria relating to reserve design. As far as we know, this study and that of Nicholls and Margules (1993) are the only comparative analyses involving quantitative representation targets. In both cases, the targets were a minimum percentage area of each land type.

Previous comparisons of heuristic algorithms have therefore not involved systematic alteration of rules within the same type of algorithm and have mostly dealt with qualitative representation targets requiring one or more occurrences of each feature. Only one of the comparative studies (Willis *et al.*, 1996) has tested alternative heuristics against an optimizing algorithm. In developing efficient heuristics for use in New South Wales, we wanted to know four things: (1) the relative contribution to efficient representation made by alternative rules within the same type of heuristic algorithm; (2) the relative efficiencies of alternative algorithms for quantitative, as well as qualitative, representation targets; (3) efficiencies relative to optimizing algorithms; and (4) details of the operation of heuristic algorithms as a basis for understanding the advantages and disadvantages of alternative rules. This paper reports on the findings for 30 heuristic algorithms, 12 for qualitative representation targets (at least one occurrence of each land type) and 18 for quantitative targets (a minimum percentage area of the total extent of each land type). We begin by describing the data set and the objectives of the algorithms. We then outline the formulation of the optimizing algorithm and the rules of the heuristics and compare their results on a large regional data set. Finally, we discuss the implications of the results, including the strengths and limitations of the alternative analyses.

METHODS

Data set

We used a regional data set from the Western Division of New South Wales, a semi-arid region dominated by leasehold grazing land and occupying about 325,000 km^2 or 40% of the State. The natural features to be represented in notional reserves are land systems, recurring patterns of landform, soil and vegetation (Mabbutt, 1968) which have been mapped at a scale of 1:250,000 by the Soil Conservation Service of New South Wales (Walker, 1991). The 248 land systems in the Western Division are well-established as a way of defining natural environments. Like any landscape classification, they have limitations as a sole basis for adequately representing all the species of the region in reserves (Pressey, 1994) and might, in time, be superseded by alternative classifications. For the analyses in this comparative paper, the value of the land system mapping is that it provides a consistent data base over a large area.

The 'sites' assessed and selected by the analyses approximate pastoral holdings. No cadastral base was available to us in digital form so we used simple rectangles which are close to the average size and shape of the actual pastoral holdings on each of the 25 1:250,000 map sheets. The size of holdings varies between map sheets mainly in relation to rainfall and major soil types. The total number of sites defined in this way was 1885. The simulated holdings would not be suitable for a conservation planning process intended to identify specific areas for acquisition but are adequate for this indicative exercise. They should also give a reasonably accurate indication of the total number and area of sites needed to achieve particular representation targets in the region. They will, however, reduce the variation in total area of selected sites between algorithms and between repeated runs of the same algorithm because of the smaller local variation in the size of the simulated, compared to actual, holdings.

Representation targets

We applied heuristic and optimizing algorithms to four representation problems:

(1) the minimum number of sites needed to represent at least one occurrence of each feature;
(2) the minimum total area of sites needed to represent at least one occurrence of each feature;
(3) the minimum number of sites needed to represent at least 5% of the total regional extent of each feature; and
(4) the minimum total area of sites needed to represent at least 5% of the total regional extent of each feature.

Problems 2 and 4 are important for data sets in which sites differ in size. Studies that use sites of varying area such as forest fragments (Saetersdal *et al.*, 1993), wetlands (Margules *et al.*, 1988) or parcels of land tenure (Pressey & Nicholls, 1989*a*) might be more interested in minimizing the total area of sites required to achieve a representation goal. In studies where sites are rectangles or hexagons of equal areas (e.g. Church *et al.*, 1996; Willis *et al.*, 1996; Csuti *et al.*, 1996), the goal of minimizing number and area are equivalent.

Heuristic algorithms

We used two types of rarity algorithms. Presence–absence algorithms find a small set of sites to represent every feature at least once, regardless of the area of features occurring in the notional reserves. Proportional area algorithms select sites to represent at least a nominated percentage area of each feature. The sequences of rules for the 12 presence–absence algorithms and 18 proportional area algorithms are listed below. All algorithms have a primary rule, applied first, to identify the site(s) with unique features. For each subsequent selection, a secondary rule is applied. In most cases, this identifies the sites with the next rarest under-represented feature(s) in the region. Ties for the secondary rule are resolved by one or more additional rules, each being called only if more than one site has been selected by the previous rule. After each iteration (for each new selected site) the level of representation of each feature in the data set is updated so, in the next iteration, the rules apply only to unselected sites and under-represented features.

Sequences of rules for presence–absence algorithms (see Table 1 for definition of rules):

PA1 unique/next rarest/random
PA2 unique/next rarest/richest/random
PA3 unique/next rarest/smallest/random
PA4 unique/next rarest/average rarity/random
PA5 unique/next rarest/total rarity/random
PA6 unique/next rarest/nextnext rarest/random
PA7 unique/next rarest/pc richest/random
PA8 unique/next rarest/richest per area/random
PA9 unique/next rarest/richest/nextnext rarest/random
PA10 unique/next rarest/richest/smallest/random
PA11 unique/next rarest/richest/total rarity/random
PA12 unique/next rarest/richest/nextnext rarest/total rarity/random

Sequences of rules for proportional area algorithms (see Table 1 for definition of rules):

PR1 unique/next rarest/random
PR2 unique/next rarest/maxcontrib/random
PR3 unique/maxcontrib/random
PR4 unique/next rarest/maxcontrib/smallest/random
PR5 unique/next rarest/weighted maxcontrib/random
PR6 unique/weighted maxcontrib/random
PR7 unique/next rarest/weighted maxcontrib/smallest/random
PR8 unique/next rarest/max rarcontrib/random
PR9 unique/next rarest/max rarcontrib/mostcontrib/random
PR10 unique/next rarest/max rarcontrib/weighted maxcontrib/random
PR11 unique/next rarest/max rarcontrib/weighted propcontrib/random
PR12 unique/next rarest/weighted propcontrib/random
PR13 unique/weighted propcontrib/random
PR14 unique/next rarest/weighted propcontrib/smallest/random
PR15 unique/next rarest/max rarcontrib/weighted propcontrib/max pccontrib/random
PR16 unique/next rarest/max pccontrib/random
PR17 unique/max pccontrib/random
PR18 unique/next rarest/max rarcontrib/max pccontrib/random

One source of inefficiency in iterative algorithms is that later selections can incidentally represent features that have already been represented in previous sites,

Table 1. Definitions of rules for heuristic algorithms

Rule	Selection of site(s) with the following characteristics
average rarity	highest average rarity fraction (100/frequency in the data set) of all under-represented features
maxcontrib	highest sum of contributions to full representation (contribution = area of each feature that would narrow the gap between the target area and the currently represented area)
max pccontrib	highest sum of contributions (as in maxcontrib) expressed as percentages of site area
max rarcontrib	highest sum of contributions (as in maxcontrib) for under-represented feature(s) with highest rarity fractions (100/frequency in data set)
mostcontrib	highest number of under-represented features that would be fully represented with the notional reservation of the sites
next rarest	highest rarity fraction (100/frequency in the data set) of any under-represented feature
nextnext rarest	second highest rarity fraction (100/frequency in the data set) of any under-represented features
pc richest	highest number of under-represented features as a percentage of the total number of features in the site
random	randomly chosen position in the list of unselected sites
richest	highest number of under-represented features
richest per area	highest number of under-represented features per unit area
smallest	smallest area
total rarity	highest sum of rarity fractions (100/frequency in the data set) for all under-represented features
unique	unique feature(s)
weighted maxcontrib	highest sum of contributions (as in maxcontrib) weighted by the rarity fraction (100/frequency in the data set) of each feature
weighted propcontrib	highest sum of weighted contributions (as in weighted maxcontrib) but with contributions expressed as a percentage of the remaining area of each feature still to be represented

making one or more of the earlier selections redundant. We therefore included a procedure in all algorithms to check for redundant sites. This took each initially selected site in turn and tested whether its removal caused the representation of any feature to fall below target. If no features fell below target, the site was identified as redundant and removed from the initial list of selections. Each algorithm gave results for number and total area of selections with and without redundant sites.

We ran all algorithms 100 times on the same data set to give a range of results from the random selections.

Optimizing algorithms

For each of the four representation problems we attempted to find an optimal solution by applying two packages: LP_SOLVE, available as freeware from Michel Berkelaar, Eindhoven University of Technology, Department of Electrical Engineering, Design Automation Section, PO Box 513, NL-5600 MB Eindhoven, The Netherlands, and LINGO, a major commercial package. Both packages use the 'branch-and-bound' method to find optimal solutions (Lawler & Wood, 1966).

We formulated the branch-and-bound algorithms to find a set of sites of minimum total cost that achieves adequate representation of each feature. Assume that there are m sites and n land systems (in this case, $m = 1885$ and $n = 248$). Let A be an m by n matrix whose elements, a_{ij}, are a measure of the representation of feature j in each site i. For each feature j, the user sets a minimum level of representation, r_j. Each site can have a unique cost c_i. The general integer linear programming problem is then (Possingham *et al.*, 1993): minimize $\sum c_i x_i$ subject to $\sum a_{ij} x_i \geq r_j$ for $j = 1...n$ (this is the minimum representation constraint for each feature). $x_i = 0$ or 1 for $i = 1...m$ where x_i are the control variables such that $x_i = 1$ if site i is in the reserve system $x_i = 0$ if site i is not in the reserve system for $i = 1...m$.

The four representation problems are specific cases of this general formulation.

Problem 1

In this case, the occurrence and not the area of a feature in a site is of interest, so $a_{ij} = 1$ if feature j is in site i and $a_{ij} = 0$ otherwise, for $i = 1...m$, $j = 1...n$. The minimum acceptable representation of each land system is 1, so $r_j = 1$ for all j and the cost of every site is the same, so $c_i = 1$ for all i. This is a special case of an integer linear programming problem known as a 'set-covering' problem.

Problem 2

The data matrix, A, is the same as for problem 1. The occurrences of features and not their areas in particular sites is of interest. Similarly, only a single representation of each feature is required, so $r_j = 1$ for all j. However, for this problem the cost of a site is its area, so c_i = area of site i and the objective is to minimize the total area of the selected sites.

Problem 3

As with problem 1, the cost of each site is 1 but in this case the data matrix, A, contains the area of each feature in each site: a_{ij} = area of feature j in site i for all i and j. The minimum adequate representation in this case is not 1 but 5% of the total area occupied by land system j in the region.

Problem 4

This is the most complicated problem. The minimum adequate representation is 5% of the total area of each land system, as in problem 3, and the cost of each site is its area, as in problem 2.

We ran all four problems on SUN IPX SPARC workstations.

RESULTS AND DISCUSSION

Redundant sites in sets selected by heuristic algorithms

Elimination of redundant sites reduced numbers and areas selected for most presence–absence and proportional area algorithms (Tables 2 and 4). The size of the reduction was proportional to the initial results including redundant sites. Pearson correlation coefficients for average percentage reduction against average initial results are $r = 0{\cdot}969$ ($P < 0{\cdot}001$) for presence–absence algorithms on site numbers, $r = 0{\cdot}973$ ($P < 0{\cdot}001$) for presence–absence algorithms on total site area, $r = 0{\cdot}987$ ($P < 0{\cdot}001$) for proportional area algorithms on site numbers, and $r = 0{\cdot}948$ ($P < 0{\cdot}001$) for proportional area algorithms on total site area.

Elimination of redundant sites did not result in maximum efficiency of any heuristic algorithms. Even after elimination of redundant sites, there was a wide range of average and minimum results for number and total area of selected sites for both presence–absence and proportional area algorithms (Tables 2 and 4). For example, PA1 selected a minimum of 69 sites to represent each feature after redundant sites had been eliminated. The much smaller result for PA2 (57 sites) indicates that PA1 and other algorithms selected sets of sites that were only loosely complementary in the features they contained, i.e. while no sites could be eliminated without some feature becoming unrepresented, the sites duplicated one another in the features they contained to a larger extent than in the smaller set selected by PA2.

All further comparisons of algorithms are based on the results of heuristics after elimination of redundant sites.

Comparisons of presence–absence algorithms

The heuristic algorithms differed widely in average and minimum results without redundant sites (Table 2, Figs 1 and 2). The smallest number and smallest total area of sites needed to represent every feature were found by different algorithms. PA2, selecting sites with the highest increment of unrepresented features at each step, was

Table 2. Summary of results from 100 runs of each presence–absence algorithm: average numbers and total areas (km²) of selected sites with and without redundant sites; and minimum numbers and areas without redundant sites; bracketed numbers are percentage reductions after taking out redundant sites; OP, solution from branch-and-bound algorithm

	Av. no. (+red.)	Av. no. (−red.)	Min no. (−red.)	Av. area (+red.)	Av. area (−red.)	Min area (−red.)
PAI	82·30	75·47 (8·30)	69	18906·15	17427·65 (7·82)	15974·50
PA2	59·70	59·57 (0·22)	57	14487·14	14456·56 (0·21)	13620·25
PA3	89·06	81·00 (9·05)	81	17284·92	16093·75 (6·89)	16093·75
PA4	116·98	86·94 (25·68)	79	25479·04	19415·98 (23·80)	16958·25
PA5	70·94	68·37 (3·62)	66	16853·00	16345·56 (3·01)	15590·25
PA6	63·60	62·33 (2·00)	59	15084·94	14831·89 (1·68)	13962·25
PA7	71·78	71·68 (0·14)	67	16177·08	16171·08 (0·04)	15198·75
PA8	62·00	61·00 (1·61)	61	13380·58	13360·83 (0·15)	13359·75
PA9	60·41	60·25 (0·26)	59	14530·13	14488·70 (0·29)	13839·00
PA10	59·24	59·24 (0·00)	59	13811·32	13811·32 (0·00)	13784·50
PA11	60·00	59·65 (0·58)	58	14326·82	14241·53 (0·60)	13718·00
PA12	60·00	59·98 (0·03)	59	14480·47	14478·39 (0·01)	14200·50
OP			54 12710·00[a]			12084·50 60[b]

[a]Smallest total area corresponding to minimum number of sites.
[b]Smallest number corresponding to minimum total area of sites.

the best algorithm for site number. PA8, selecting on number of unrepresented features per unit area, was the best for total site area. Overall, the ranking of algorithms from highest to lowest average results for site number was significantly but imperfectly correlated with the ranking on average results for total site area (Kendall rank correlation coefficient $T = 0{\cdot}697$, $P < 0{\cdot}01$).

Elaboration of algorithms by adding further rules did not always lead to smaller results. Although most rules added to the basic rarity algorithm (PA1) reduced the numbers and total areas selected, size (PA3) did not reduce required numbers of sites (Table 2, Fig. 1). Average rarity (PA4) did not reduce either numbers or total area (Table 2, Figs 1 and 2). PA3 selected larger numbers of sites that were smaller and had fewer features on average than those selected by PA1. PA4, in averaging the rarity of all features in each site, tended to select sites with few relatively rare features and avoid sites with large numbers of commoner features that reduced average rarity values. Adding further rules to PA2 in PA9-12 led to a slight reduction in average number of sites in PA10 and in average total area of sites in PA10 and PA11 (Table 2, Fig. 1 and 2) but it did not reduce or even match the overall minimum number or total area of sites selected. This indicates that random selection, focused on a set of sites with high potential contributions to representing new features as in PA2, can be more effective in finding minimum

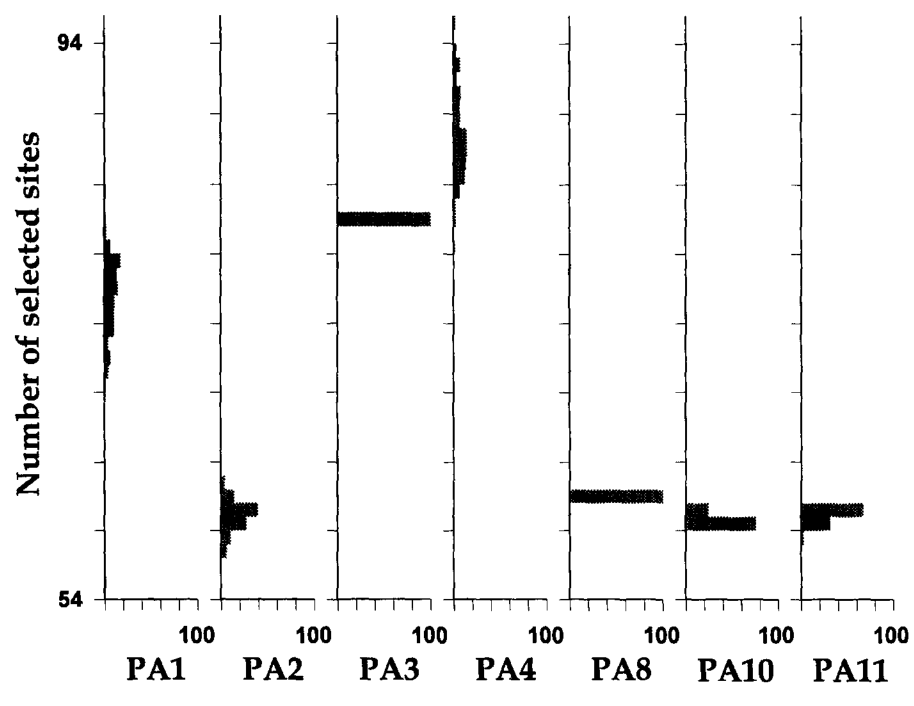

Fig. 1. Ranges of numbers of selected sites (without redundant sites) from 100 runs of some of the presence–absence algorithms; the base of the Y-axis is the result from the optimizing algorithm.

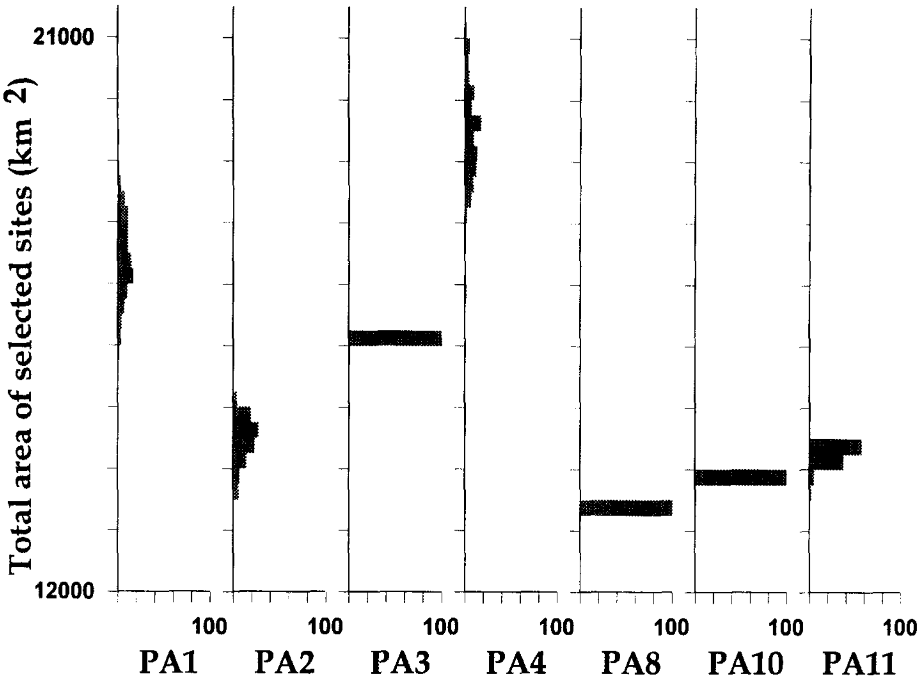

Fig. 2. Ranges of total areas of selected sites (without redundant sites) from 100 runs of some of the presence–absence algorithms; the base of the Y-axis is the result from the optimizing algorithm.

requirements, at least over many repeated runs, than additional rules. However, the lower average results for PA10 and PA11 compared to PA2 indicate that these elaborations would be slightly more likely to find a small result for one or a few runs.

The value of repeated random selections depends strongly on the sequence of preceding rules. Most algorithms produced smaller results than PA1, the basic algorithm on which all others were based, and which made all choices randomly. The likelihood of randomly making optimal or near-optimal selections for each of about 75 iterations on average in the case of PA1 is extremely small. Very many runs of the algorithm, many orders of magnitude more than 100, would probably be necessary to find a result for number or total area of sites comparable to the best of the presence–absence algorithms.

The algorithms also differed widely in the ranges of results from 100 runs (Figs 1 and 2). Small ranges of results tended to come from algorithms with rules that relied on real numbers with many possible values such as site area (PA3, PA10) or number of features per unit area (PA8). There were relatively few ties for these rules so subsequent random selections to resolve ties were rarely necessary. The 100 runs of the algorithms therefore selected sets of sites with very similar compositions and, consequently, very similar numbers of sites and total site areas. Larger ranges of results tended to come from algorithms that selected sites on the basis of integers (PA2) or real numbers with fewer possible values (PA4). There were more ties for these rules and more random selections needed to resolve them. As a result, the sets of selected sites differed in composition more often between the 100 runs of each algorithm. Some sequences of random selections were more effective than others in efficiently representing all features, leading to differences between runs in numbers and total areas of sites.

Supporting this interpretation is a correlation between the range of results produced by each algorithm and the average percentage of random choices over 100 runs of each algorithm for site number (Pearson correlation coefficient $r = 0{\cdot}766$, $P < 0{\cdot}01$) and total site area ($r = 0{\cdot}802$, $P < 0{\cdot}01$). Algorithms leading to very few random selections also produced relatively few unique sets of selected sites (Table 3). Those with few unique sets (PA3, PA8 and PA10) had small ranges for number and total area of sites (Figs 1 and 2). At least two other factors are not accounted for by these relationships: random choices following some rules led to greater ranges of results than random choices following other rules; and the differences between unique sets of sites in terms of number and total area were greater for some algorithms than others.

Comparisons of proportional area algorithms

As for presence–absence algorithms, different proportional area algorithms found the smallest number of sites (PR11) and the smallest total area of sites (PR18) to represent 5% of the extent of each feature (Table 4, Figs 3 and 4) and the overall rankings of algorithms according to average number and average area were significantly but imperfectly matched (Kendall $T = 0{\cdot}704$, $P < 0{\cdot}001$).

All but one of the elaborations of the most basic proportional area algorithm (PR1) improved the results for number of sites and all other algorithms improved

Table 3. Average percentage of random selections over 100 runs and number of unique sets of selected sites from 100 runs for each heuristic algorithm

Algorithm	Average percentage random choices over 100 runs	Number of unique selected sets[a]
PA1	100·00	100
PA2	70·14	100
PA3	7·95	76
PA4	61·96	100
PA5	45·29	100
PA6	70·83	100
PA7	54·89	100
PA8	4·92	38
PA9	52·16	100
PA10	5·56	46
PA11	40·68	100
PA12	38·98	100
PR1	100·00	100
PR2	11·70	100
PR3	25·29	100
PR4	3·12	16
PR5	10·00	100
PR6	19·53	100
PR7	0·77	2
PR8	92·68	100
PR9	67·88	100
PR10	6·40	99
PR11	8·94	100
PR12	8·80	100
PR13	8·00	100
PR14	1·60	6
PR15	2·44	12
PR16	10·15	99
PR17	44·80	100
PR18	0·78	2

[a]A unique set differs in composition from all other sets in at least one selected site; selection order is not considered.

the total area of selected sites (Table 4, Figs 3 and 4). PR17 increased the number of sites needed to represent all features relative to PR1 (Fig. 3) because it tended to select small sites with proportionally high contributions to representing features but needed to select more of these to achieve the representation target.

Some of the rules for the proportional area algorithms were more complex than those for presence–absence algorithms. Those relating to the 'contribution' of sites calculated the extent to which reservation of the site would narrow the gap between target and currently represented areas of each feature and totalled this for all features in the site (Table 1). We tested several variations on this basic idea. Weighted proportional contribution (PR12-14) generally produced smaller results for number and total area of selected sites than maximum contribution (PR2-4), weighted maximum contribution (PR5-7), or maximum percentage contribution (PR16-17) (Table 4). These four variations usually produced similar or smaller results for site number and total site area if they followed a 'next rarest' rule (PR2/PR3, PR5/PR6, PR12/PR13, PR16/PR17). Another preliminary rule that invariably increased the effectiveness of the contribution rules was "max rarcontrib": selecting the site that made the largest contribution to full representation of the next rarest under-represented feature(s). Contribution rules always selected smaller site numbers and total site areas when they followed this rule (PR1/PR8, PR5/PR10, PR12/PR11, and especially PR16/PR18). Maximum percentage contribution became the most effective rule for minimizing total site area when it followed 'max rarcontrib'.

As for presence–absence algorithms, there was wide variation in the ranges of results for both number and total area of selected sites (Figs 3 and 4). The factors behind this variation appear to be the same as discussed for presence–absence algorithms. There were significant

Table 4. Summary of results from 100 runs of each proportional area algorithm: average numbers and total areas (km^2) of selected sites with and without redundant sites; and minimum numbers and areas without redundant sites

	Av.no.(+red.)	Av.no.(−red.)	Min no.(−red.)	Av. area(+red.)	Av. area (−red.)	Min area (−red.)
PRI	200·68	157·02 (21·76)	150	41063·87	32667·08 (20·45)	30756·25
PR2	140·61	135·11 (3·91)	133	30992·42	29546·06 (4·67)	29086·50
PR3	140·11	132·65 (5·32)	129	31627·75	29797·94 (5·79)	28800·00
PR4	140·79	136·37 (3·14)	136	30820·47	29567·26 (4·07)	29515·50
PR5	131·00	128·88 (1·62)	126	29029·31	28258·49 (2·66)	27346·75
PR6	129·00	127·80 (0·93)	125	29220·16	28945·60 (0·94)	28166·25
PR7	131·00	130·00 (0·76)	130	28806·00	28347·50 (1·59)	28347·50
PR8	154·29	145·65 (5·60)	139	32969·01	31063·71 (5·78)	29062·50
PR9	134·28	132·81 (1·09)	128	28874·80	28611·61 (0·91)	27456·00
PR10	126·00	125·92 (0·06)	125	27571·37	27545·26 (0·09)	27099·50
PR11	124·00	123·90 (0·08)	123	26791·76	26782·54 (0·03)	26549·00
PR12	126·00	125·80 (0·16)	125	27884·37	27843·68 (0·15)	27503·25
PR13	126·00	125·92 (0·06)	125	27878·18	27837·71 (0·15)	27356·50
PR14	126·00	126·00 (0·00)	126	27640·00	27640·00 (0·00)	27640·00
PR15	124·00	124·00 (0·00)	124	26413·00	26413·00 (0·00)	26413·00
PR16	158·70	147·30 (7·18)	145	31651·83	30154·22 (4·73)	29773·25
PR17	208·85	164·55 (21·21)	157	32396·08	29566·99 (8·73)	28075·50
PR18	130·00	126·00 (3·08)	126	27094·50	25887·50 (4·45)	25887·50

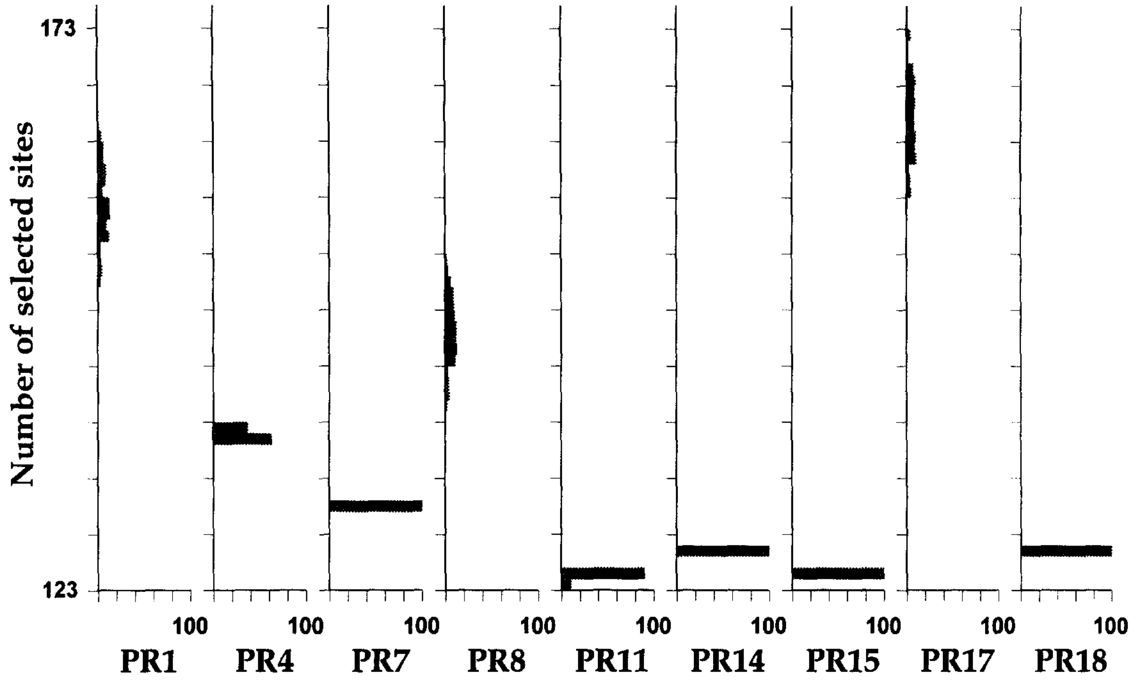

Fig. 3. Ranges of numbers of selected sites (without redundant sites) from 100 runs of some of the proportional area algorithms.

correlations between the ranges of results from algorithms and the average percentage of random choices made by algorithms for site number (Pearson $r = 0{\cdot}872$, $P < 0{\cdot}001$) and total site area ($r = 0{\cdot}947$, $P < 0{\cdot}001$). Algorithms that produced relatively few random choices (PR4, PR7, PR14, PR15 and PR18) also produced relatively few unique sets of sites (Table 3). These were also the algorithms with small ranges of results for number and total area (Figs 3 and 4).

Relative efficiencies and running times of heuristic and optimizing algorithms

The problems were too large to run with the version of LINGO to which we had access and no results could be produced with this package.

The best results for the two presence–absence problems from LP_SOLVE were 54 sites for minimum number and 12,084·5 km_2 for minimum total area. As for the heuristic algorithms, the set of sites that minimized

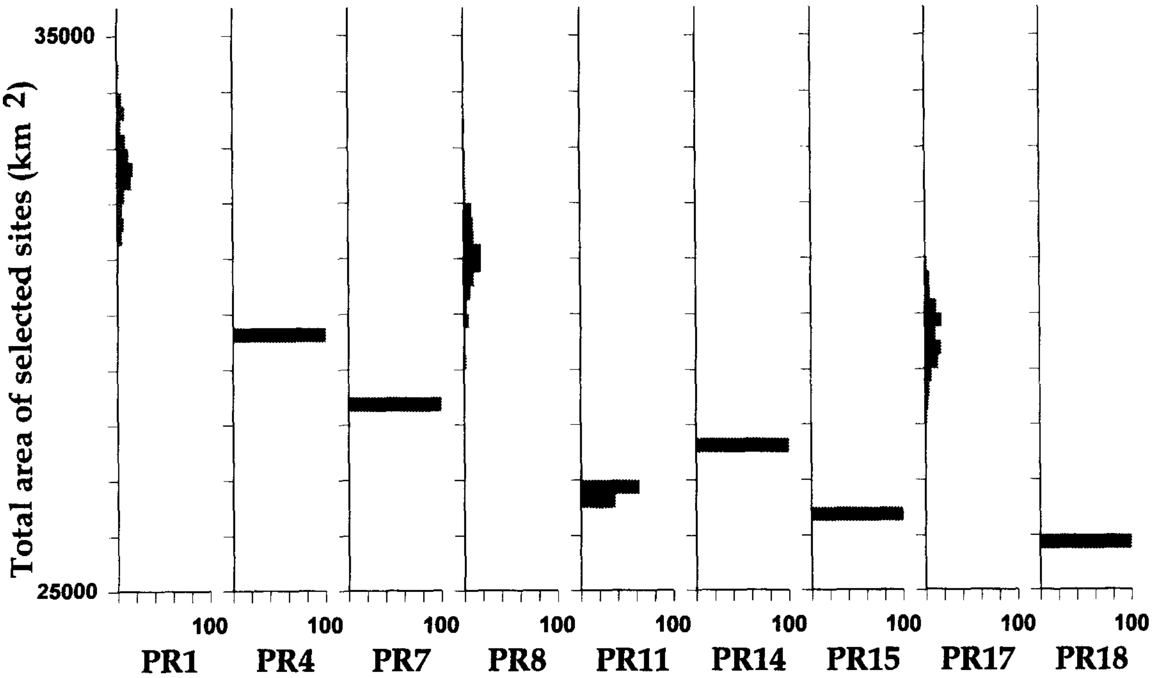

Fig. 4. Ranges of total areas of selected sites (without redundant sites) from 100 runs of some of the proportional area algorithms.

number was different from that needed to minimize total area (Table 2). The best heuristic result for overall minimum site number to represent each feature once, from PA2, was 5·6% larger than the LP_SOLVE result. The best heuristic result for average site number, from PA10, was 9·7% larger than that from LP_SOLVE. For total site area, the best presence–absence algorithm (PA8) produced a minimum and average result 10·6% larger than optimal.

For the presence–absence problems, heuristic and optimizing algorithms differed considerably in running time. For problem 1, the minimum number of sites to represent at least one occurrence of each feature, LP_SOLVE (version 1·5) could not guarantee an optimal solution but did produce a solution of 54 sites after about a day of running on the SPARC station. After about a month of running, in conjunction with other tasks on the machine, no guaranteed optimum was found. More recent work with LP_SOLVE (version 2·0) since this manuscript was first submitted has confirmed that 54 sites is the optimal result. To run to completion, the package took about 76 h on a HP 9000/735 machine which is approximately 10 times faster than the SPARC stations previously used. The optimal result for problem 2, minimum total area of sites to represent at least one occurrence of each feature, was found by LP_SOLVE (version 1·5) after about 14 h of CPU time on a SPARC station. LP_SOLVE was much more effective for problem 2 because the cost function requiring total area of sites to be minimized was much more selective and enabled a clear choice to be made between sites with similar features but different areas. The presence–absence heuristics varied in running time but averaged about 5 min per run on a 486 33Mhz machine. All the heuristic programs were prototypes and have since been optimized for speed to run in seconds on the same problems.

Both optimizing algorithms failed to find optimal solutions to the problem of representing at least 5% of the area of each feature, even after weeks of running on the SPARC stations. LP_SOLVE (version 1·5) produced suboptimal results, larger than those from some heuristics, but did not run to completion. Reducing the data set from 1885 sites to 400 sites did not make the problem tractable for the optimizing algorithms. Subsequent work with LP_SOLVE (version 2·0) has provided optimal results for the 5% problems, but only for data sets of 40 sites. Analysis of larger data sets is still problematic. The proportional area heuristics took up to 10 min per run on a 486 33Mhz machine. As for the presence–absence programs, the prototypes have now been optimized for speed and run on the same problems in 20–30 s.

The size of the four problems (1885 sites × 248 features) is large and close to the limit of what most branch-and-bound packages can deal with. Problems 3 and 4, with proportional area targets, are far too large for the two optimizing packages that we tried, even with a reduced data set of 400 sites. Alternative optimizing packages might perform better with these problems.

GENERAL DISCUSSION

The large differences in results between the heuristic algorithms are due to the relative effectiveness of their rules in 'packing' the required occurrences of features into the relatively inflexible framework of property boundaries. This is not a peculiarity of our data set. The same challenge would be posed if we had used natural units such as catchments or wetland boundaries, arbitrary units like grid cells or hexagons, or other imposed units like vegetation remnants in a largely cleared landscape.

The results are guidelines for writing efficient heuristic algorithms for indicative analyses in conservation planning. They have determined the choice of algorithms in comparisons of reserve requirements for alternative reservation scenarios (Pressey & Tully, 1994), alternative data bases (Pressey & Logan, 1995), and other analyses yet to be published. These guidelines are all the more important because of the failure of the optimizing algorithms to solve the proportional area problems. This is a significant limitation of optimization because quantitative representation targets will often be more important than qualitative targets. One occurrence of each land system in the Western Division is an unrealistic reservation target. Some land systems would be represented in notional reserves by very large areas and others by very small and inadequate areas. A proportional area target ensures that each land system is adequately represented. The same argument applies to representation of species. One, two or even five occurrences of each species in a region could lead to inadequate representation of many species. More realistic targets would be based on indices such as proportions of total populations or core distributional ranges where this information is available.

The results reported here on the relative efficiencies of different algorithms are likely to depend to some extent on the particular data set used. Data-dependence is indicated by the imperfect correlation between rankings of proportional area algorithms from best to worst for the full data set and a reduced data set of 400 sites (for minimum site number, Kendall's rank correlation $T=0{\cdot}614$, $P<0{\cdot}001$; for minimum total area of sites, $T=0{\cdot}11$, $P>0{\cdot}5$). The relative rankings of presence-absence algorithms in this study, according to number of selected sites, also differed somewhat from the rankings of the same algorithms applied to a species data set from Oregon (Csuti *et al.*, 1996). Pressey and Nicholls (1989*b*) found that the relative efficiencies of a rarity algorithm and a richness algorithm varied between data sets. More tests of the algorithms compared in this study, as well as other algorithms, are therefore desirable on a variety of data sets. This line of work might lead to an understanding of how particular data structures influence the relative performance of alternative algorithms.

Including procedures in heuristics for eliminating redundant selections usually makes them more efficient but does not lead to optimality. All the presence–absence heuristics tested in this study are suboptimal, the best ones 5·6% larger than optimal for number of sites and 10·6% larger for total area of sites. The degree of suboptimality of the proportional area algorithms is unknown.

Some recent discussions of optimizing algorithms and heuristics (Underhill, 1994; Church *et al.*, 1996) have emphasized the issue of efficiency of representation. In this respect, optimizing algorithms would always be preferable to heuristics. The choice of algorithms can be different, however, in actual planning exercises, using large data sets and attempting to solve more complex problems than one or more occurrences of each feature. In a previous paper, we have discussed the need to look beyond efficiency at a range of other issues relevant to conservation planning on the ground (Pressey *et al.*, 1996) and only three main points need to be made here. First, the value of reserve selection algorithms is primarily indicative rather than prescriptive. Useful indicative results, say in comparing the area costs of alternative reservation scenarios, can be provided without the guarantee of optimality. If more comparative precision is needed than heuristics can offer, and if running time is not a constraint, then an optimizing analysis is a better choice. Second, running times are so much faster for heuristics that they can be incorporated into real-time interactive systems (e.g. Pressey *et al.*, 1995), even for large data sets. For smaller data sets, recent applications of branch-and-bound algorithms have demonstrated running times of seconds (Church *et al.*, 1996; Csuti *et al.*, 1996), which certainly makes them suitable for interactive use. Third, a more compelling advantage of heuristics is that they can actually find an answer to proportional area problems for large data sets and, furthermore, can do this quickly enough to be used interactively. The same difference might apply to some complex problems that combine representation targets with goals relating to the suitability for protection and configuration of individual reserves.

There is also much potential for the development of heuristic approaches other than the stepwise rarity and richness analyses that have been commonly applied in conservation planning. We have been exploring other heuristic methods including genetic algorithms and simulated annealing and the latter approach has been the most successful to date. Simulated annealing can generate a range of efficient solutions to reserve selection problems and will improve on these if left to run for longer. On the same data set used in this paper, simulated annealing found a solution to problem 1 equal to that from the branch and bound package, although it was less successful for problem 2.

With improved programs and faster processing, the scope for using optimizing algorithms in conservation planning will continue to increase and the relative importance of heuristics will decline. For the time being, the speed and versatility of heuristics and the many questions that they can answer about alternative approaches to conservation planning justify their further refinement on a range of problems.

There are, of course, many questions that heuristics or any other reserve selection algorithms, used alone, cannot answer. We emphasize again that the primary role of selection algorithms is indicative, to give conservation planners a rapid and accurate picture of the potential costs and difficulties of achieving one or more reservation goals. This role can make them important tools in policy analysis and in shaping agreed, achievable goals. Real-world conservation planning, which marks out the final boundaries of feasible reserves and establishes their formal protection, is something else. It might have goals derived from the use of selection algorithms but it requires more than the algorithms can provide on their own. Final decisions will ideally be based on comparisons of alternative sites and alternative networks and will often involve a range of agencies and interest groups. The interactive systems that are being developed for systematic real-world planning can have selection algorithms as components (e.g. Williams *et al.*, 1991; Bedward *et al.*, 1992) but also allow the flexibility needed for prescriptive planning. These systems and other similar ones (e.g. Pressey *et al.*, 1995) can provide information on alternative sites in addition to that used by the selection algorithms. This extra information, which might include land use history, the occurrence of species of interest, and the attitudes of current owners, can be influential in deciding between alternatives for reservation. There is also much potential for interactive systems to be used to allocate a range of coordinated management approaches when outright reservation is neither practical nor appropriate.

ACKNOWLEDGEMENTS

The work reported here was supported by funding from the Murray-Darling Basin Commission through the Natural Resources Management Strategy. Michel Berkelaar gave us a great deal of help in using his package, LP_SOLVE, and provided time on a powerful machine for the most recent analyses. Ian Ball and Alister Smith also helped in running the optimizing algorithms and Graeme Dandy, Maureen Goldfinch and Franz Salzborn offered further technical advice on these packages. We thank Sharon Tully for compiling the data base used in the analyses and for running many of the heuristic programs and organizing the results. Peter Bowen also assisted with running heuristic programs and compiling results. Vicki Logan ran the statistical analyses and prepared the figures.

REFERENCES

Ackery, P. R. and Vane-Wright, R. I. (1984). *Milkweed butterflies* British Museum (Natural History), London.

Bedward, M., Pressey, R. L. and Keith, D. A. (1992). A new approach for selecting fully representative reserve networks: addressing efficiency, reserve design and land suitability with an iterative analysis. *Biol. Conserv.*, **62**, 115–125.

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

Cocks, K. D. and Baird, I. A. (1989). Using mathematical programming to address the multiple reserve selection problem: an example from the Eyre Peninsula, South Australia. *Biol. Conserv.*, **49**, 113–130.

Csuti, B., Camm, J. D., Downs, B., Huso, M., Hamilton, R., Kershaw, M., Kiester, A. R., Polasky, S., Pressey, R. L., Sahr, K. and Williams, P. H. (1996). A comparison of reserve selection algorithms using data on terrestrial vertebrates in Oregon. *Biol. Conserv.*, **80**, 83–97.

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

Kershaw, M., Williams, P. H. and Mace, G. M. (1994). Conservation of Afrotropical antelopes: consequences and efficiency of using different site selection methods and diversity criteria. *Biodiv. Conserv.*, **3**, 354–372.

Kershaw, M., Mace, G. M. and Williams, P. H. (1995). Threatened status, rarity and diversity as alternative selection measures for protected areas: a test using Afrotropical antelopes. *Conserv. Biol.*, **9**, 324–334

Kirkpatrick, J. B. (1983). An iterative method for establishing priorities for the selection of nature reserves: an example from Tasmania. *Biol. Conserv.*, **25**, 127–134.

Kirkpatrick, J. B. and Brown, M. J. (1991). *Reservation analysis of Tasmanian forests.* Resource Assessment Commission Forest and Timber Inquiry Consultancy Series Number **FTC91/16**, RAC, Canberra.

Kirkpatrick, J. B. and Brown, M. J. (1994). A comparison of direct and environmental domain approaches to planning reservation of forest higher plant communities and species in Tasmania. *Conserv. Biol.*, **8**, 217–224.

Kirkpatrick, J. B. and Harwood, C. E. (1983). Conservation of Tasmanian macrophytic wetland vegetation. *Proc. R. Soc. Tas.*, **117**, 5–20.

Lawler, E. L. and Wood, D. E. (1966). Branch-and-bound methods: a survey. *Op. Res.*, **14**, 699–719.

Lewis, A., Stein, J. L., Stein, J. A., Nix, H. A., Mackey, B. G. and Bowyer, J. K. (1991). *An assessment of regional conservation adequacy: Tasmania.* Resource Assessment Commission Forest and Timber Inquiry Consultancy Series Number **FTC91/17**, RAC, Canberra.

Lombard, A. T., Nicholls, A. O. and August, P. V. (1995). Where should nature reserves be located in South Africa? A snake's perspective. *Conserv. Biol.*, **9**, 363–372.

Mabbutt, J. A. (1968). Review of concepts of land classification. *In Land evaluation*, ed. G. A. Stewart. MacMillan, Melbourne, pp. 11–28.

Margules, C. R. (1989). Selecting nature reserves in the South Australian mallee. In *Mediterranean landscapes in Australia: mallee ecosystems and their management*, ed. J. C. Noble and R. A. Bradstock. CSIRO, Melbourne, pp. 398–405.

Margules, C. R. and Nicholls, A. O. (1987). Assessing the conservation value of remnant islands: mallee patches on the western Eyre Peninsula, South Australia. In *Nature conservation: the role of remnants of native vegetation*, ed. D. A. Saunders, G. W. Arnold, A. A. Burbidge and A. J. M. Hopkins. Surrey Beatty and Sons, Sydney, pp. 89–102.

Margules, C. R., Nicholls, A. O. and Pressey, R. L. (1988). Selecting networks of reserves to maximise biological diversity. *Biol. Conserv.*, **43**, 663–676.

Mickleburgh, S. P., Hutson, A. M. and Racey, P. A. (1992). *Old World fruit bats: an action plan for their survival.* International Union for the Conservation of Nature and Natural Resources, Gland.

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

Parker, S. P. (1984). *McGraw-Hill dictionary of scientific and technical terms*, 3rd ed. McGraw-Hill, New York.

Possingham, H., Day, J., Goldfinch, M. and Salzborn, F. (1993). The mathematics of designing a network of protected areas for conservation. In *Proceedings Australian Operations Research Conference, 12th* ed. D. Sutton, E. Cousins and C. Pearce. University of Adelaide, Adelaide, pp. 536–545.

Pressey, R. L. (1994). Land classifications are necessary for conservation planning but what do they tell us about fauna? In *Future of the fauna of western New South Wales*, ed. D. Lunney, S. Hand, P. Reed and D. Butcher. Royal Zoological Society of NSW, Sydney, pp. 31–41.

Pressey, R. L. and Logan, V. S. (1995). Reserve coverage and requirements in relation to partitioning and generalization of land classes: analyses for western New South Wales. *Conserv. Biol.*, **9**, 1506–1517.

Pressey, R. L. and Nicholls, A. O. (1989*a*). Application of a numerical algorithm to the selection of reserves in semi-arid New South Wales. *Biol. Conserv.*, **50**, 263–278.

Pressey, R. L. and Nicholls, A. O. (1989*b*). Efficiency in conservation evaluation: scoring vs iterative approaches. *Biol. Conserv.*, **50**, 199–218.

Pressey, R. L. and Nicholls, A. O. (1991). Reserve selection in the Western Division of New South Wales: development of a new procedure based on land system mapping. In Nature conservation: cost effective biological surveys and data analysis, ed. C. R. Margules and M. P. Austin. CSIRO, Melbourne, pp. 98–105.

Pressey, R. L. and Tully, S. L. (1994). The cost of *ad hoc* reservation: a case study in the Western Division of New South Wales. *Aust. J. Ecol.*, **19**, 375–384.

Pressey, R. L., Bedward, M. and Nicholls, A. O. (1990). Reserve selection in mallee lands. In *The mallee lands: a conservation perspective*, ed. J. C. Noble, P. J. Joss and G. K. Jones. CSIRO, Melbourne, pp. 167–178.

Pressey, R. L., Johnson, I. R. and Wilson, P. D. (1994). Shades of irreplaceability: towards a measure of the contribution of sites to a reservation goal. *Biodiv. Conserv.* **3**, 242–262.

Pressey, R. L., Ferrier, S., Hutchinson, C. D., Sivertsen, D. P. and Manion, G. (1995). Planning for negotiation: using an interactive geographic information system to explore alternative protected area networks. In *Nature conservation: the role of networks*, ed. D. A. Saunders, J. L. Craig and E. M. Mattiske. Surrey Beatty and Sons, Sydney, pp. 23–33.

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

Rebelo, A. G. (1994). Iterative selection procedures: centres of endemism and optimal placement of reserves. In *Botanical diversity in southern Africa*, ed. B. J. Huntley National Botanical Institute, Pretoria, pp. 231–257.

Rebelo, A. G. and Siegfried, W. R. (1990). Protection of Fynbos vegetation: ideal and real world options. *Biol. Conserv.*, **54**, 15–31.

Rebelo, A. G. and Siegfried, W. R. (1992). Where should nature reserves be located in the Cape Floristic Region, South Africa? — models for the spatial configuration of a reserve network aimed at maximising the protection of floral diversity. *Conserv. Biol.*, **6**, 243–252.

Ryti, R. T. (1992). Effect of the focal taxon on the selection of nature reserves. *Ecol. Appl*, **2**, 404–410.

Saetersdal, M., Line, J. M. and Birks, H. J. B. (1993). How to maximize biological diversity in nature reserve selection: vascular plants and breeding birds in deciduous woodlands, western Norway. *Biol. Conserv.*, **66**, 131–138.

Thomas, C. D. and Mallorie, H. C. (1985). Rarity, species richness and conservation: butterflies of the Atlas Mountains in Morocco. *Biol. Conserv*, **33**, 95–117.

Underhill, L. G. (1994). Optimal and suboptimal reserve selection algorithms. *Biol. Conserv.*, **70**, 85–87.

Vane-Wright, R. I. and Rahardja, D. P. (1993). An evaluation of the diversity of subspecies, species and genera of Hesperiidae within The Philippines, using the WORLDMAP computer program. *Zool. Verh.*, **288**, 116–121.

Vane-Wright, R. I., Humphries, C. J. and Williams, P. H. (1991). What to protect? — systematics and the agony of choice. *Biol. Conserv.*, **55**, 235–254.

Walker, P. J. (1991). *Land systems of western New South Wales* Soil Conserv. Serv. NSW Tech. Rep., No. 25 SCS, Sydney.

Williams, P. H., Humphries, C. J. and Vane-Wright, R. I. (1991). Measuring biodiversity: taxonomic relatedness for conservation priorities. *Aust. Syst. Bot.*, **4**, 665–679.

Williams, P. H., Vane-Wright, R. I. and Humphries, C. J. (1992). Measuring biodiversity for choosing conservation areas. In *Hymenoptera and biodiversity*, ed. J. LaSalle and I. D. Gauld. CABI, Wallingford, UK, pp. 309–328.

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

Williams, P. H., Prance, G. T., Humphries, C. J. and Edwards, K. S. (1996*b*). Priority-areas analysis and the Manaus-90 areas for conserving the diversity of Neotropical plants (families Proteaceae, Dichapetalaceae, Lecythidaceae, Caryocaraceae and Chrysobalanaceae). *Biol. J. Linn. Soc.*, **58**, 125–157.

Willis, C. K., Lombard, A. T., Cowling, R. M., Heydenrych, B. J. and Burgers, C. J. (1996). Reserve systems for limestone endemic flora of the Cape lowland fynbos: iterative versus linear programming techniques. *Biol. Conserv.*, **77**, 53–62.