Pages 536-545 in "Decision Sciences: tools for today" Proceedings of 12th National ASOR Conference. Eds. D.J. Sutton, C.E.M. Pearce and E.A. Cousins, ASOR, Adelaide.

~~10 cop~~ 5 copies please

The mathematics of designing a network of protected areas for conservation

Hugh Possingham, Jemery Day, Maureen Goldfinch and Franz Salzborn.*

June 21, 1993

Abstract

Australia has committed itself to setting up a representative, comprehensive and adequate system of reserves to preserve biodiversity. One method for designing an efficient reserve system is to minimise the area in reserves while still achieving certain conservation goals, like representing every species. Traditionally scientists have used sub-optimal iterative methods to try to find the most efficient reserve system. We show how two simple rules can simplify small problems so that the most efficient solution can be found. For larger problems integer linear programming packages can be used to find the optimal reserve system for this static problem. We illustrate these techniques with examples. In reality we have to wait until sites become available. Before they can be acquired some of the sites may be lost through land degradation or vegetation clearance. We use stochastic dynamic programming to solve a control problem that allows for this sort of uncertainty. The method is illustrated with an example.

1 Introduction

Biodiversity is the diversity of life and it encompasses the diversity of ecosystems, species, and diversity within species (genetic diversity). Australia, like many other nations, has committed itself to preserving as much of the biodiversity of this country as possible (Biological Diversity Advisory Committee, 1992). A key element of that commitment is the acquisition of a system of protected areas that adequately represents the biodiversity of the country (Commonwealth

*Applied Mathematics Department, University of Adelaide, South Australia, 5005, Australia.

of Australia, 1992a, 1992b). In this paper we consider the problem of acquiring sites for conservation to achieve maximal representation of biodiversity for minimum cost. In more specific terms our objective is to acquire a set of sites that represents every species, or as many species as possible, given presence/absence data for those species in the available sites. The problem of representing all vegetation types or any other element of biodiversity is essentially the same.

Traditionally, when faced with a series of sites that could be included in a reserve system, authors rank sites according to a variety of criteria. These methods of conservation evaluation, reviewed by Margules and Usher (1981), have been widely used but have the failing that they do not rank sites in the context of other possible acquisitions. For example we may find that after ranking a number of sites, the top three sites have similar species compositions and their acquisition does not lead to a broad representation of species. Kirkpatrick and Harwood (1983) first recognised this problem and proposed a method of iteratively selecting sites that takes into account previous acquisitions. These iterative methods are widely used and are based on the notion that sites should be rated according to how complimentary they are, in terms of species, to the sites already acquired (Kirkpatrick, 1983; Margules *et al.*, 1988; Pressey and Nicholls, 1989a). These algorithms are often called "minimum set" algorithms.

The reserve selection methods so far developed are static - they result in a list of sites that should be acquired to maximise the representation of species. There is a tacit assumption that we can acquire what sites we need when we want them. In the real world there are events that are likely to stop us from acquiring the optimal set of reserves. For example, some sites may be destroyed before they are acquired, others may be unavailable, while political and economic changes in a country may mean that the constraints of the problem will change. In the second part of the paper I show how stochastic dynamic programming can be used to approach the reserve selection problem and take into account these uncertainties. The method, interesting conclusions, and problems with the method are illustrated with a specific example.

2 Static reserve design

In this section we consider the problem of reserving the minimum number of sites to represent every species once. Pressey (1993) has advocated that this version of the reserve design problem is important because it most efficiently achieves the goal of representing every species. Although such a minimal goal is not necessarily best for nature conservation it specifys a baseline reserve system that can be obtained for minimal cost (assuming for simplicity each site has equal cost).

Assume that there are m sites to select from and n species that occur in at least one site. Let A be an m by n matrix whose elements, a_{ij}, are

$$a_{ij} = \begin{cases} 1 & \text{if species } j \text{ exists in site } i \\ 0 & \text{otherwise,} \end{cases}$$

for $i = 1, \ldots, m$ and $j = 1, \ldots, n$. Let the control variables that determine whether or not we acquire a site be x_i, where

$$x_i = \begin{cases} 1 & \text{if we reserve site i} \\ 0 & \text{otherwise,} \end{cases}$$

for $i = 1, \ldots, m$. Then, if the cost of each site is the same, the minimum set problem is

$$\begin{array}{ll} \text{minimise} & \sum_i x_i \\ \text{subject to} & \left. \begin{array}{l} \sum_i a_{ij} x_i \geq 1 \\ x_i \in \{0, 1\} \end{array} \right\} \text{for } j = 1, \ldots, n. \end{array}$$

This integer linear program can be solved using a variety of commercially available packages and the problem can be extended to include complexities such as sites with different costs, constraints that require each species to be represented more than once, and so forth (Cocks and Baird, 1989). The common attribute of all these problems is that the control variables can only be in one of two states and we should use packages specifically designed for this sort of problem.

Before considering a specific example we describe two simple rules that can be used to simplify the problem.

Rule 1 Remove every site (row) that contains a set of species that is a subset of, or is equivalent to, the set of species in another site.

Rule 2 Remove every species (column) that occurs in a set of sites that is a superset of, or is equivalent to, the set of sites that another species is in.

Once the problem is reduced in this way, the solution may be self evident or an integer linear programming package can be applied to the reduced problem.

2.1 Examples

It is interesting to show how iterative solution methods can fail and our simple rules can be used to solve small problems by hand. Consider the example presented in Table 1. Here there are 10 species and 10 sites. The optimal solution to this problem is to select sites 4,6 and 9. The optimal solution can be found using our two reduction rules.

	Species										
Site	A	B	C	D	E	F	G	H	I	J	Total
1	1	1	1	1	1	1	0	0	0	1	7
2	1	1	1	1	1	0	0	0	0	1	6
3	1	1	1	1	0	1	0	1	0	0	6
4	1	1	1	1	0	0	0	0	1	0	5
5	1	1	1	0	0	1	0	1	0	0	5
6	0	0	1	0	1	0	0	1	0	1	4
7	1	0	1	1	0	0	0	0	1	0	4
8	1	1	0	0	0	0	0	1	0	0	3
9	0	0	0	0	0	1	1	0	0	0	2
10	0	1	0	0	0	0	0	0	0	0	1

Table 1: Species presence/absence data used in the static problem example. The row numbers refer to the sites and the column letters are different species.

Two commonly used iterative methods are:

1. Reserve the site that adds most species (is most complimentary) to the reserve system already acquired and repeat.
2. Reserve all sites that contain a unique representation of a species, then follow the first iterative method.

In both cases there is the possibility that there are two or more sites that add an equal number of species to the existing reserve system. The way in which we choose between such sites leads to a variety of different iterative algorithms.

By applying iterative method one to our example we would reserve site 1 first (most species) and then regardless of how we proceed we require a total of four sites. Using iterative method two we acquire site 9 first and then either site 1 or 2. Whichever we choose we will again only attain all the species by choosing four sites in total. This shows that the iterative methods fail on even small problems, and yet without using a mathematical programming package we can solve many small problems optimally.

Pressey and Possingham (in prep.) consider a much larger problem with 1885 sites and 248 species, actually landform types (Pressey and Nicholls, 1989b). Thirteen different iterative methods were attempted and the results ranged from

58 to 79 sites. Using a public domain linear programming package LP_SOLVE (Michel Berkelaar, Eindhoven University of Technology, Dept of Electrical Engineering, Design Automation Section, PO Box 513, NL-5600 MB Eindhoven) we found the optimal solution is 57 sites after about 10 days computation on a SUN IPX workstation. Using the reduction methods described above the problem can be reduced to 251 sites and 119 species. Nineteen of the sites must be reserved because after reduction they contain a unique representation of a species. Our reductions simplified the problem considerably. We are currently working on integrating the reduction rules with our own branch and bound method to give a fast and flexible tool for solving this sort of problem.

3 The dynamic reserve design problem

Solving the static problem we identified the smallest number of sites that reserves every species. Practically we could wait until each of these sites becomes available and then acquire them. This strategy will fail if there is the possibility that sites will be destroyed (eg. cleared for agriculture or otherwise degraded to a point at which they are no longer suitable). In the previous example sites 4, 6 and 9 provided a very neat match. If site 6 were destroyed then there are twelve equally optimal solutions achieving full representation with four sites. The solution to the problem has changed significantly. In this section we use stochastic dynamic programming to determine whether or not to acquire a reserve if it becomes available.

3.1 Formulation of the problem

As before let A denote the site by species presence/absence matrix. Without any constraints we would buy every site, so assume that we can only acquire a finite number of sites, C. Let $x_i(t)$ be the state of each site in the system at time t such that

$$x_i(t) = \begin{cases} 2 & \text{if site } i \text{ is destroyed} \\ 1 & \text{if site } i \text{ is in the reserve system} \\ 0 & \text{otherwise.} \end{cases}$$

For convenience we define an additional state variable that tells us whether or not a site is available to be acquired

$$y_i(t) = \begin{cases} 0 & \text{if site } i \text{ is destroyed} \\ 0 & \text{if site } i \text{ is in the reserve system} \\ 1 & \text{if site } i \text{ could become available.} \end{cases}$$

Assume that we can estimate, for each site, a probability that it becomes available at time t, $s_i(t)$, and a probability that it is destroyed, $d_i(t)$. We choose a time step so that $d_i(t)$ and $s_i(t)$ are so small that the probability of two events occurring in the same time step can be ignored. Initially we assume that there is a final time at which we are given a score equal to the num

ber of species in the reserve system. Let this score be V(x(T)) where $\mathbf{x}(T)$ is the vector representing the state of the system at the final time, henceforth referred to as the reserve system. Let $J_t(\mathbf{x})$ be the expected terminal value given we currently have reserve system $\mathbf{x}$ (so $J_t(\mathbf{x}) = V(\mathbf{x}(T))$. Then at any time t, if site i becomes available (only

possible if $y_i(t) = 0$) we choose to acquire that site if

$$|\mathbf{x}| < C$$

and

$$J_{t+1}(\mathbf{x} + \mathbf{i}) > J_{t+1}(\mathbf{x})$$

where i is the ith unit vector. For each time step we can associate an optimal decision vector $\mathbf{b}(t)$ where $b_i(t) = 1$ if we should buy site i at time t. Given we have the terminal values the optimal solution is found by backstepping in the normal way using the following formula for updating $J_t(\mathbf{x})$

$$\begin{aligned} J_t(\mathbf{x}) &= J_{t+1}(\mathbf{x})[1 - \sum_i s_i(t)b_i(t)y_i(t) - \sum_i d_i(t)y_i(t)] + \\ & \sum_i J_{t+1}(\mathbf{x} + \mathbf{i})s_i(t)b_i(t)y_i(t) + \sum_i J_{t+1}(\mathbf{x} + 2\mathbf{i})d_i(t)y_i(t). \end{aligned}$$

By back-stepping far enough it is possible to find a long-term strategy, that is a strategy that does not assume an end to the process. We wrote a program to solve this problem numerically and applied it to the data in Table 2. Henceforth we assume that $\mathbf{d}$, $\mathbf{b}$ and $\mathbf{s}$ are time independent.

3.2 Results

Consider the example in Table 2 where the maximum number of reserves in the system is three, $C = 3$. The optimal solution to the static problem is to acquire sites 2,3 and 8. This reserve system protects all 15 species, there are 14 combinations of three sites that reserve 14 species. Let $s_i = 0.02$ for all i. We will vary $\mathbf{d}$ and explore its impact on both the long-term and short-term opt

imal strategy.

								Species								
Site	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	Total
1	1	1	1	1	1	1	1	0	1	1	0	0	0	0	0	9
2	1	1	0	1	1	0	1	0	1	0	1	1	0	1	0	9
3	1	1	1	1	1	1	0	0	0	0	0	1	1	0	0	8
4	1	1	1	0	1	0	1	1	0	0	0	0	1	0	0	7
5	1	1	1	1	0	1	0	0	0	0	0	1	0	1	0	7
6	1	0	0	0	0	0	1	1	1	1	1	0	1	0	0	7
7	0	0	1	0	0	1	0	1	1	0	1	0	0	0	0	5
8	0	0	0	1	0	0	0	1	0	1	0	0	0	0	1	4

Table 2: The time before termination at which each site becomes attractive for acquisition.

If **d** = 0 then the best long-term strategy is to wait for the perfect set of reserves: 2, 3 and 8. As we increase the possibility of losing sites, more of the reserves become attractive (see Table 3). This reflects our intuition: as our ability to acquire any site decreases, we become less fussy about the reserves we should acquire (and our chance of getting the perfect set decreases)

				Site				
Loss probability	1	2	3	4	5	6	7	8
d = 0.000	-	yes	yes	-	-	-	-	yes-
d = 0.010	-	yes	yes	-	-	yes	-	yes
d = 0.012	yes	yes	yes	-	yes	yes	-	yes
d = 0.015	yes	yes	yes	-	yes	yes	-	yes
d = 0.020	yes	yes	yes	yes	yes	yes	-	yes
d = 0.050	yes	yes	yes	yes	yes	yes	yes	yes

Table 3: Whether or not a reserve should be acquired if no reserves have yet been acquired where $C = 3$, $s = 0.02$ and d is fixed for all sites.

It is interesting to note that, if we rank the sites not in the perfect set according to how often they occur in the systems that protect 14 species, the ranking is: 6,5,1,4 then 7, which is the same order in which sites are included in the best long-term strategy as the risk of site destruction increases, Table 3.

An alternative to assuming sites can be destroyed is to consider a time at which reserve acquisition ceases and no more sites can be acquired. Under these circumstances, sites 2,3 and 8 will be acquired at any time before termination, site 6 should be acquired when there is 76 or less time steps to go, site 5 - 58 time steps, site 1 - 51 time steps, site 4 - 39 time steps and site 7 - 34 time steps. These other sites become suitable because, as we approach the termination the chance of acquiring the perfect set decreases. Again, the ranking of sites in terms of how often they occur in combinations that reserve 14 species is the same as their ranking in terms of how far from termination they become suitable.

Finally we consider the possibility that the total number of sites in the reserve system is 1,2, 3 or 4. In each case the time before the end of the process at which a site becomes suitable is tabulated in Table 4. If the entry is infinity - this indicates that the site is in the long-term optimal set (we would always buy it). Inspection of Table 4 shows how the suitability of a site varies enormously as C varies. For example, site 8 is only of high value if $C = 3$ or 4, while site 1 is only in the optimal long term strategy set if $C = 1$.

	Site							
C	1	2	3	4	5	6	7	8
1	∞	∞	33	18	18	18	9	6
2	47	59	∞	28	45	∞	23	22
3	51	∞	∞	39	58	76	34	∞
4	62	∞	93	66	∞	∞	48	∞

Table 4: The time before termination at which each site becomes attractive for acquisition

4 Discussion

Conservation biology is an emerging discipline in science. Until recently a great deal was idealistic. However faced with a rapid decline in biological diversity there is a move to integrate the scientific ideas into a decision making and economic framework. This paper represents one example of this transition.

We have briefly reviewed existing techniques for solving the problem of choosing the smallest set of sites that represents every species. A method of reducing this sort of problem was introduced. The two reduction rules will often completely solve small problems (eg. 20 sites by 20 species) and can considerably reduce the size of large problems. The primary purpose of this paper is to present a new reserve design problem that incorporates the uncertainties and dynamics of a reserve acquisition process. We formulate this new problem and solve an example using stochastic dynamic programming methods. The example serves to illustrate how the stochastic dynamic programming solutions accord with our intuition about intuitively reasonable strategies for reserve design in an uncertain world.

The examples and problems we explore are only indicative. There are many assumptions that could be relaxed and complexities included. Further work on large and more realistic problems is needed. Although we may strive continually for better algorithms and mathematical techniques it is worth remembering that there will always be a lot of uncertainty about the parameters and data in the problems. For the static problem there may be errors in the presence/absence matrix. For the dynamic problem there will be unavoidable uncertainty about the likelihood sites become available for sale, the possibility sites are destroyed and the constraints under which we are working. Consequently robust heuristic solutions will play a significant role.

The greatest problem with the stochastic dynamic programming solution is an inability to deal with large state spaces as it is currently formulated. One mechanism of dealing with this problem is to develop a simpler heuristic method. For example, in Section 3.2 we found that the set of sites that should be acquired given a certain level of risk reflects the number of times that a site is represented in good (not necessarily the best) combinations of reserves. We are attempting other methods to deal with this problem. We will use the exact stochastic dynamic programming solutions to test heuristic methods that could be applied to large data sets.

Acknowledgments

We are grateful to Bob Pressey for his assistance, access to data, and inspiration.

References

Biological Diversity Advisory Committee (1992). *A national strategy for the conservation of Australia's biological diversity.* AGPS, Canberra.

Cocks, K.D. and Baird, T.A. (1989). Using mathematical programming to address the multiple reserve selection problem: An example from the Eyre Peninsula, South Australia. *Biological Conservation* **49**, 113-130.

Commonwealth of Australia (1992a). *National Forest Policy Statement: A new focus for Australia's forests.* Advance Press, Perth.

Commonwealth of Australia (1992a). *Inquiry into the role of protected areas in the maintenance of biodiversity.* Report of the House of Representatives Standing Committee on the Environment, Recreation and the Arts. AGPS, Canberra.

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

Kirkpatrick, J.B. and Harwood, C.E. (1983). Conservation of Tasmanian macrophytic wetland vegetation. *Proceedings of the Royal Society of Tasmania* **117**, 5-20.

Margules, C.R. and Usher, M.B. (1981). Criteria used in assessing wildlife conservation potential: a review. *Biological Conservation* **24**, 115-128.

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

Pressey, R.L. (1993). Opportunism in acquiring land for reserves - why it is a bad idea. *National Parks Journal (NSW)* **36**, 19-22.

Pressey, R.L. and Nicholls, A.O. (1989a). Efficiency in conservation evaluation: scoring v. iterative approaches. *Biological Conservation* **50**, 199-218.

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

Pressey, R.L. and Possingham, H.P. (in prep.). Effectiveness of alternative heuristic algorithms for identifying indicative minimum requirements for conservation reserves.