available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/biocon

# Heuristic algorithms vs. linear programs for designing efficient conservation reserve networks: Evaluation of solution optimality and processing time

*Robert P.D. Vanderkam*[a,*], *Yolanda F. Wiersma*[b], *Douglas J. King*[c]

[a]Canadian Wildlife Service, Environment Canada, 351 St. Joseph Blvd., Gatineau, QC, Canada K1A 0H3
[b]Department of Biology, Memorial University, St. John's, Newfoundland and Labrador, Canada A1B 3X9
[c]Department of Geography and Environmental Studies, Carleton University, 1125 Colonel By Drive, Ottawa, ON, Canada K1S 5B6

ARTICLE INFO

*Article history:*
Received 27 September 2006
Received in revised form
27 January 2007
Accepted 22 February 2007
Available online 18 April 2007

*Keywords:*
Systematic conservation planning
Reserve design
Integer linear programs
Heuristic algorithms
Optimization
Efficiency

ABSTRACT

Systematic approaches to efficient reserve network design often make use of one of two types of site selection algorithm; linear programs or heuristic algorithms. Unlike with linear programs, heuristic algorithms have been demonstrated to yield suboptimal networks in that more sites are selected in order to meet conservation goals than may be necessary or fewer features are captured than is possible. Although the degree of suboptimality is not known when using heuristics, some researchers have suggested that it is not significant in most cases and that heuristics are preferred since they are more flexible and can yield a solution more quickly. Using eight binary datasets, we demonstrate that suboptimality of numbers of sites selected and biodiversity features protected can occur to various degrees depending on the dataset, the model design, and the type of heuristic applied, and that processing time is not dramatically different between optimal and heuristic algorithms. In choosing an algorithm, the degree of suboptimality may not always be as important to planners as the perception that optimal solvers have feasibility issues, and therefore heuristic algorithms might continue to be a popular tool for conservation planning. We conclude that for many datasets, feasibility of optimal algorithms should not be a concern and that the value of heuristic results can be greatly improved by using optimal algorithms to determine the degree of suboptimality of the results.



## 1. Introduction

Many resources are being expended to mitigate the global reduction in biodiversity, including through the delineation of protected areas (which might include parks, wildlife refuges, or conservation reserves) (Arcese and Sinclair, 1997; Scott et al., 2001; Margules et al., 2002; Williams et al., 2005a). Delineation of these areas (hereafter referred to as reserves) is an exercise in prioritization (Curio, 2002; Kingsland, 2002). Since not all areas can be set aside from development, and different areas have varying conservation and economic values, decisions must be made as to which areas are the most appropriate to conserve. Such decision-making can be facilitated through the use of systematic approaches with clearly defined and consensus driven methodologies based on current scientific practices, such as the principle of complementarity (Pressey et al., 1993). These approaches have been used around the world, and are seen to be advantageous

* *Corresponding author:* Tel.: +1 819 953 1444; fax: +1 819 994 4445.
E-mail addresses: robert.vanderkam@ec.gc.ca (R.P.D. Vanderkam), ywiersma@mun.ca (Y.F. Wiersma), doug_king@carleton.ca (D.J. King).
0006-3207/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.biocon.2007.02.018

because they allow the development of networks of reserves that make use of efficiencies of scale to increase the number of biodiversity features (species, landscape elements, ecological processes, or spatial relationships) that are captured in them (Vane-Wright et al., 1991; Rothley, 1999; Andelman and Willig, 2003; Margules and Pressey, 2000; Possingham et al., 2000; Williams et al., 2005b; Salomon et al., 2006).

The explicit and transparent nature of systematic approaches may provide more justifiable arguments for creating and siting conservation reserves than approaches based on subjective, ad hoc, or purely expert-based decisions (Williams, 1998; Cowling et al., 2003a; Noss, 2003; Bojórquez-Tapie et al., 2004; Pullin et al., 2004). However, there are different aspects of systematic approaches that can be challenged and there is considerable debate as to whether all aspects are equally appropriate or justifiable. Pressey (2002) outlines the history of systematic approaches that use complementarity and how it began with the work of Kirkpatrick (1983). Margules and Pressey (2000) summarize the methodologies that evolved in the two decades following Kirkpatrick's work. Systematic approaches vary, but they commonly adopt a regional, multi-feature perspective (Cowling et al., 2003b) that is flexible, efficient (in terms of minimizing the number of sites required while maximizing the number of features conserved), and transparent (i.e., an explicit methodology in which the steps and decisions can be re-traced) (Possingham et al., 2000).

There are two main types of algorithms used in systematic reserve selection and these are described in more detail below. In brief, optimal algorithms use complex mathematical processes (i.e., linear programs) that output the entire set of sites that satisfy the model exactly (Church et al., 1996; McDonnell et al., 2002). In reserve design, linear programs often use binary (presence/absence) or integer (e.g., area) variables with presence or percent area targets as constraints, and are commonly referred to as integer linear programs (IP or ILP). The second type, known as heuristic algorithms, prioritize potential reserve locations by using iterative procedures to select sites one at a time, sometimes with an additive-only process and sometimes with a give-and-take process, to build a solution set to satisfy stated objectives (Pressey et al., 1997; Margules and Pressey, 2000). Heuristic algorithms can select sites using rules that maximize feature additions (i.e., greedy or richness algorithms) or that maximize specific feature attributes (e.g., rarity algorithms) and can be used to solve non-linear models as well (Williams et al., 2004). This paper does not consider non-linear cases where variables are interdependent, such as with population viability modeling, as some researchers have reported IP unsuitable for that purpose (Williams et al., 2004; Moilanen and Wintle, 2006; but see Haight et al., 2000; Haight et al., 2005).

### 1.1. *Integer programs*

Integer programs (IPs) have been in development for military and business applications since the second world war and have been adapted for reserve selection since the late 1980s (Cocks and Baird, 1989; Kingsland, 2002; ReVelle et al., 2002). In a reserve selection problem, variables such as presence/absence (e.g., species) or extent of biodiversity features (e.g., habitat cover) in a set of candidate sites can be represented in an IP or MIP (mixed integer program). Two types of optimization objectives can be applied to select the optimal number of sites from the set for inclusion in a nominal reserve network. The first, in which the model objective is to find the minimum number or set of sites that can include some number of occurrences of the features or some percentage or total area, is known as the locational set covering problem (LSCP, sensu Church et al., 1996; but also called the Species Set Covering Problem by ReVelle et al., 2002). Detailed descriptions of the LSCP model can be found in Rodrigues et al. (2000), Williams et al. (2004), and others. The second type of objective sets the maximum number of sites a priori as a constraint, and the solution set includes the maximum possible number or area of features that can be included in that set of sites. This is known as the maximal covering location problem (MCLP, sensu Church et al., 1996; but also called the Maximal Covering Species Problem by ReVelle et al., 2002). The two problems (LSCP and MCLP) are closely related, the LSCP being equivalent to the MCLP when the a priori fixed number of sites for an MCLP objective includes all features.

### 1.2. *Heuristic algorithms*

The use of heuristic methods to select representative reserve networks began in the 1980s, incorporating logical iterative methodologies to process datasets. Heuristic algorithms typically use a decision tree procedure, often with a similar goal to the LSCP of representing all species or other surrogate features in at least one location in the region, or to the MCLP of maximizing the number of features found in a predetermined number of sites or area. Such a heuristic sorts the sites according to one or more actual or derived attribute values, such as species richness or presence of rare species, decided on by the researchers or decision makers to represent, or act as a surrogate for, biodiversity. It then selects the site with the highest value for that attribute and places it in the network of notional sites. The algorithm then selects the site with the highest score from those not yet selected, often based on the principle of complementarity—that is, it re-calculates the conservation attribute values after first removing from further calculations those attributes that are included in the notional set (e.g., Saetersdal et al., 2003; Poulin et al., 2006; Oetting et al., 2006). This process reiterates in a step-wise fashion until all conservation features are represented as per the model goals. Most heuristic algorithms include rules for breaking ties as necessary throughout the process. These rules can be based on additional attributes, such as size, condition, or cost of the sites under consideration, or can be based on random selection methods.

### 1.3. *Heuristic vs. optimal algorithms*

The two approaches (heuristic and optimal algorithms) were developed independently of each other; heuristic methods address the same goals as the optimal approaches but use logical rather than mathematical methodologies (Possingham et al., 1993; Underhill, 1994). However, Possingham et al. (1993) and Underhill (1994) both pointed out that heuristic algorithms sometimes resulted in solution sets that contained sites that were redundant or extraneous to the problem. That is, heuristics could not guarantee optimal results. They also noted that heuristic algorithms could not inform the user of the degree of suboptimality in the solution. This is true for the greedy method or more complex non-IP methods such as simulated annealing (Possingham et al., 2000). Optimality, in the sense used by Underhill (1994), and adopted here, refers to the minimum number of sites in a network needed to include all features in at least one location (for the LSCP) or the maximum number of features covered by a given size of network (for the MCLP).

Pressey et al. (1996) wrote that suboptimality should be minimized as much as possible but that optimal solutions were not always practical. IPs were not capable, at that time, of solving problems such as area percentage models in a reasonable time if too many sites or variables were involved, and today, even larger datasets are in common use (e.g., Crossman and Bryan, 2006). However, Pressey et al. (1996) and others (Willis et al., 1996; Rosing et al., 2002) pointed out that a certain degree of suboptimality may not be a drawback in practice because ecological and political factors may also influence final decisions on site locations, even when IPs are used to find optimal solutions. In a comparison of 30 different heuristics, Pressey et al. (1997) found that those that had explicit sets of rules yielded reserve networks that were only 5–10% larger than the network identified by an optimal algorithm. Since then, however, other comparisons between heuristic and optimal algorithms have concluded that greedy heuristics can yield considerably suboptimal solutions (see Table 1 in Rodrigues et al., 2000 for a summary; also see Moore et al., 2003; Fischer and Church, 2005).

This paper addresses this debate by exploring the differences between the two algorithm types in terms of solution suboptimality and the time needed to produce a solution. We compare an IP algorithm to a greedy heuristic algorithm with four variations for both LSCP and MCLP models using eight datasets for each test. The datasets are presence/absence (binary) data that represent different features, extents, and degrees of spatial resolution. We did not test datasets with non-binary variables, such as feature extent per site, but discuss their use. Similar tests have previously been done for the LSCP (e.g., Rodrigues and Gaston, 2002) and the results for the LSCP model here can be analyzed from within the context of that body of knowledge. As far as we know, however, no comprehensive testing has been done using the MCLP. In addition, since there have not been conclusive comparisons of the speed of these algorithms, we measured the time each took to produce an LSCP solution for nine datasets of varying size and features.

## 2. Methods

### 2.1. Data

Eight datasets were used from three geographical regions: British Columbia (BC), Canada; a portion of Ontario (ON), Canada; and the Carrickalinga Creek watershed near Adelaide, Australia (AU). They were acquired from existing reserve selection projects (Lindsay et al., unpublished; Olsen, unpublished; Bryan, unpublished, respectively). The datasets were all 0–1, presence–absence matrices that varied in number of sites, number of features, and the biodiversity surrogate type they were purported to represent (Table 1). We selected datasets of various size in order to test the algorithms of interest and provide some context to previous, similar studies regarding the performance of the two optimizing algorithm types.

### 2.2. Application of optimal algorithms

OPL Studio (Optimization Programming Library Studio, v.3.7 ILOG, Inc., Mountain View, California) "CPLEX MIP (Mixed Integer Program)" (v.9.0) algorithm was used as the optimal algorithm. This is a general purpose algorithm that makes proprietary use of a number of techniques such as branch-and-bound, branch-and-cut (Mitchell, 2002; Fischer and Church, 2005), two families of heuristics, pre-processing, and probing (see Önal, 2003 for general details on commercial optimizing methods) and solves binary or binary-integer problems. LSCP and MCLP models were applied to all datasets and the first complete solution set (many optimal solutions are typically found) was used as an example solution to compare to the heuristic solutions.

### 2.3. Application of heuristic algorithms

C-Plan (Conservation Planning Program; Anonymous, 1999) software was used to implement four variations of a greedy heuristic. Cells were prioritized and selected for the minimum set procedure for both the LSCP and MCLP using four attributes: rarity (inverse of feature frequency; Pressey et al.,

**Table 1 – Description of the eight datasets used in this analysis**

| Region | Biodiversity surrogate features | Number of features in group (F) | Number of grid cells (G) | Array size (F × G) | Cell size | Source |
|---|---|---|---|---|---|---|
| BC | Reptile species | 15 | 1048 | 15,720 | 640 $km^2$ hexagon | Lindsay et al., unpublished |
| BC | Amphibian species | 18 | 1551 | 27,918 | 640 $km^2$ hexagon | Lindsay et al., unpublished |
| BC | Species of concern | 25 | 1589 | 39,725 | 640 $km^2$ hexagon | Lindsay et al., unpublished |
| BC | Species at risk | 59 | 1588 | 93,692 | 640 $km^2$ hexagon | Lindsay et al., unpublished |
| BC | Mammal species | 102 | 1589 | 162,078 | 640 $km^2$ hexagon | Lindsay et al., unpublished |
| BC | Bird species | 274 | 1589 | 435,386 | 640 $km^2$ hexagon | Lindsay et al., unpublished |
| BC | Bird species (modified) | 255 | 15,890 | 4,051,950 | 640 $km^2$ hexagon | Lindsay et al., unpublished |
| ON | Bird species | 213 | 1203 | 256,239 | 10 $km^2$ square | Olsen, unpublished |
| AU | Soil and climate variables | 65 | 89,376 | 6,166,944 | 25 m square | Bryan, unpublished |

1994); richness (the sum of the number of different features present in the site); irreplaceability (Pressey et al., 1994; Ferrier et al., 2000); and summed irreplaceability (Ferrier et al., 2000). Irreplaceability and summed irreplaceability are statistical measures of the likelihood that a site/cell will be required in order to meet the goals of the model. All four prioritizing rules were applied to the LSCP and MCLP for all datasets.

### 2.4. Comparison of optimal and heuristic algorithms

Both algorithms were compared for the LSCP and MCLP problems in terms of the size of the solution set. The solution using the IP algorithm for LSCP is guaranteed to be optimal (the smallest size possible) and this was compared to the heuristic result in order to quantify its suboptimality. For MCLP models, the numbers of features captured using the optimal and heuristic algorithms were compared, with the IP algorithm being guaranteed to give the maximum number of features possible.

To evaluate the time required to achieve a solution to the LSCP using the four algorithms in C-Plan, a stopwatch was used to measure the time from when the prioritization attribute was selected until the minimum solution set was compiled. In OPL, the time to find the solution to the linear objective function and the time to find a solution set for that number of cells were added together, the former measured using a stopwatch and the latter using computer code. Timing was stopped after only one solution set was found in order to match the single minimum set solution found using C-Plan. In order to provide a dataset of intermediate size along with the smaller (BC and ON) and larger (AU) datasets for these speed tests, we modified the BC bird dataset by multiplicatively increasing the size as shown in Table 1. It was used in the speed tests with the other eight datasets. All work was done on a Pentium IV™ computer with dual Xeon 2.66 GHz CPU's, 1 GB RAM and the Windows XP Professional™ operating system.

## 3. Results

### 3.1. LSCP model

The results comparing the number of cells selected by OPL to the number of cells selected using the four functions in C-Plan for the locational set covering problem (LSCP) are shown in Fig. 1. As expected, the results from the optimal algorithm (OPL) were never larger than any of the C-Plan results for the LSCP. In all but one case (i.e., BC species at risk), C-Plan gave results that were closer to optimal when the two statistical prioritization attributes of irreplaceability and summed irreplaceability were used. Prioritizing based on summed irreplaceability also calculated redundancy as an additional sorting criterion and this produced optimal results in all eight

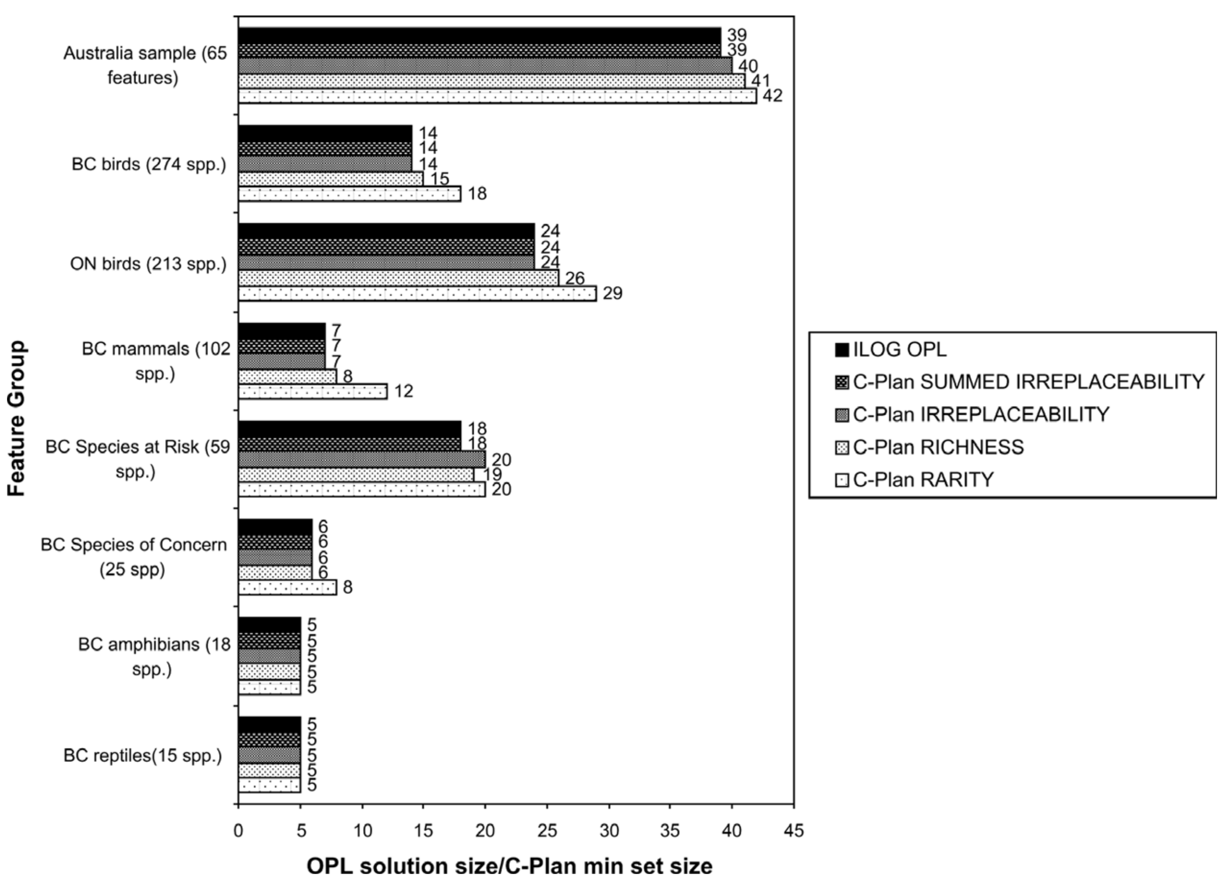


**Fig. 1 – Locational set covering problem (LSCP) solution sizes from an optimal algorithm (OPL) and four different heuristic algorithms (C-Plan).**

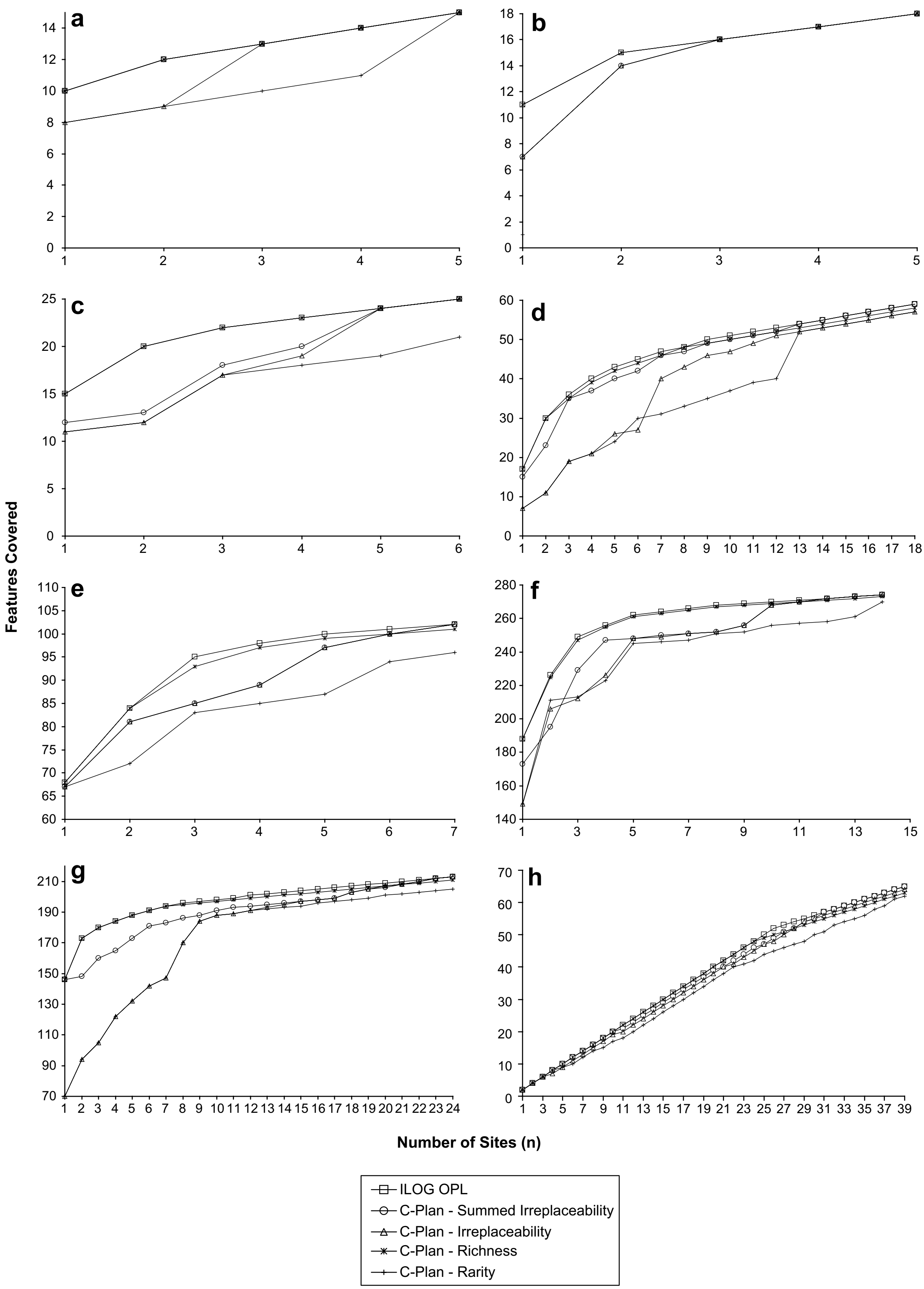


**Fig. 2 – Number of species captured (*y*-axis) under the maximal covering location problem (MCLP) for different numbers of sites (*x*-axis) using an optimal algorithm (OPL: open diamonds, solid lines) and four heuristic (C-Plan) algorithms (summed irreplaceability: open squares, solid lines; irreplaceability: open triangles, solid lines; richness: dark squares, dashed lines; rarity: dark triangles, dashed lines). Note that not all symbols are visible, due to overlaps in data points, and that axis scales vary between datasets. (a) BC reptiles, (b) BC amphibians, (c) BC species of concern, (d) BC species at risk, (e) BC mammals, (f) BC birds, (g) Ontario birds, (h) Australia data.**

datasets, while prioritization using irreplaceability did so for six of the eight. For the algorithms with richness or rarity prioritized, only three and two of eight datasets, respectively, produced the same number of sites as OPL, with the output prioritized on rarity requiring 7–71% more sites than OPL.

### 3.2. *MCLP model*

Results comparing the number of species or other features selected for each possible number of sites using the same algorithms for the maximal covering location problem (MCLP) are shown in Fig. 2. For the MCLP, the suboptimality of the four heuristic solutions for most datasets varied for different numbers of sites ($n$) as well as between heuristic prioritization methods. Put another way, suboptimality, when it exists in an MCLP problem, is usually greater than for the LSCP equivalent for the same heuristic and often increases (becomes more suboptimal) as the value of $n$ diminishes. Many of the prioritizations achieved the optimal number of species as $n$ approached the minimum number of sites required to capture all species (as identified using the LSCP). However, this did not occur for the richness attribute for BC mammals, BC species of concern, and ON birds, nor for the irreplaceability or rarity attributes for BC species at risk.

Results using richness as the sorting attribute were optimal or very close to optimal for every $n$ in every case. Summed irreplaceability was optimal in one case (BC reptiles) and nearly so for every $n$ for the BC species at risk and AU datasets. In other cases, it was quite similar in suboptimality to the irreplaceability results, except for the ON bird dataset, where irreplaceability results were more suboptimal. Rarity results were as suboptimal, or more so, than the other attribute with the most suboptimality, although this sometimes included optimal or near optimal results (e.g., ON birds for $n > 9$, BC species at risk for $n > 13$).

### 3.3. *Solution times*

The times to solve the LSCP for the nine datasets using the optimal algorithm and the four heuristic algorithms are shown in Fig. 3. Comparison of the solution times to solve the LSCP for the nine datasets showed that the time required by the optimal algorithm (OPL) was consistently equivalent to, or less than, the time required by the heuristic algorithms (C-Plan), often being less by as much as 25–50%.

## 4. Discussion

### 4.1. *Implications for suboptimality in reserve design*

Previous studies (e.g., Csuti et al., 1997; Pressey et al., 1997; Fischer and Church, 2005) have found variation in

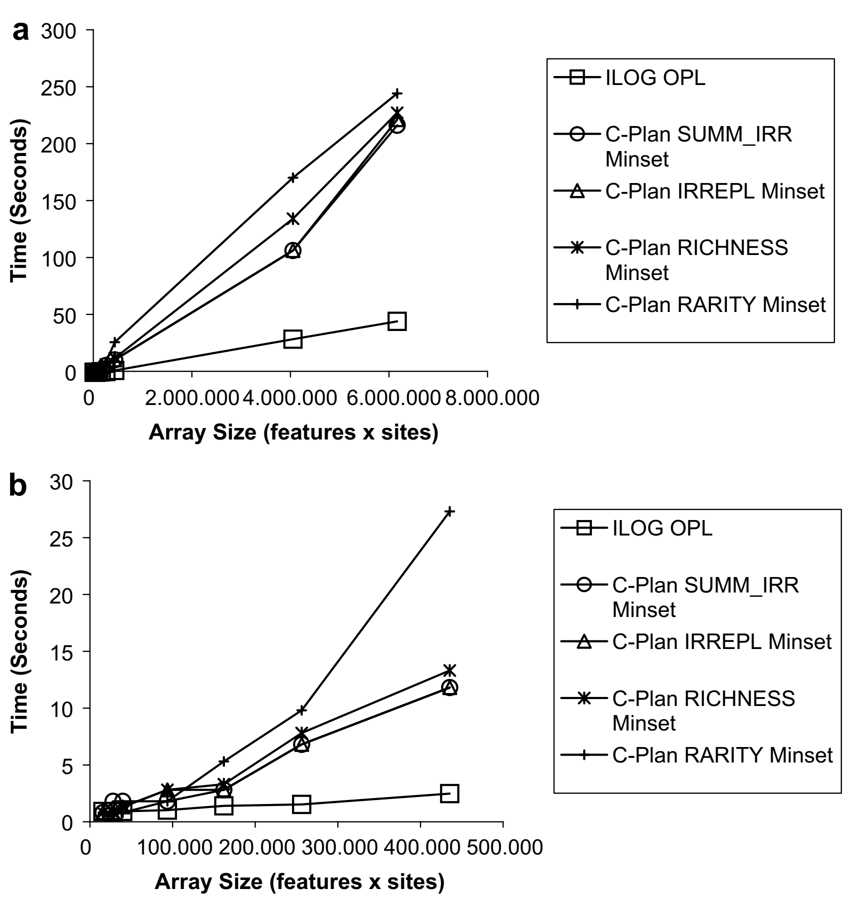


**Fig. 3 – (a) Time required to solve the LSCP for all nine data sets in Table 1 using an optimal (OPL) and four types of heuristic (C-Plan) algorithms. (b) As in Fig. 2a, but showing results for the seven smallest datasets in Table 1.**

suboptimality for the LSCP to be common and suggested that it varies with the design of the algorithm (i.e., the choice of prioritizing attribute and greedy vs. meta-heuristic) and the distribution of the features in the region. Two studies (Pressey et al., 1997; Csuti et al., 1997) used a single dataset each and found that suboptimality varied with differences in heuristic algorithm design. Csuti et al. (1997) found, as we did, that richness based algorithms gave better results than other heuristics for the MCLP. However, our results differ from two previous LSCP-type analyses, which showed that rarity based algorithms gave results closer to optimal (Csuti et al., 1997; Wiersma and Nudds, 2006); our LSCP results show that prioritizing for rarity usually yielded greater suboptimality than when prioritizing for richness.

Sætersdal et al. (1992) used a heuristic algorithm on two datasets and found that having more rare species gave better results, which was substantiated by Pressey et al. (1999) in a simulation study and by Moore et al. (2003) using a variety of datasets. Willis et al. (1996) found no suboptimality in output from two heuristic algorithms using a single dataset. Additionally, Pressey et al.'s (1999) simulation study found that the amount of nestedness or overlap of feature ranges was not related to suboptimality for a number of heuristics. However, none of these studies examined the degree of suboptimality when heuristics were based on statistical prioritization using irreplaceability or summed irreplaceability. In this study, these were generally found to have a lower degree of suboptimality than the richness or rarity based heuristic algorithms for the LSCP and varying degrees of suboptimality for the MCLP.

It is likely that the amount of suboptimality that is acceptable by a planner varies depending on the species, the nature of the conservation problem, and the planner(s). Moore et al. (2003) felt that their results of 0–9% suboptimality were close to optimal because they translated into additions of less than 1% of total area in the region. However, Önal (2003) notes that even a small overall regional percentage of additional land can have very large cost consequences. Rodrigues et al. (2000), assessing previous comparisons, report a range of 0–50% suboptimality, suggesting that results greater than about 25% were unacceptable, although these figures were not translated into percent of total area. Pressey et al. (1996), on the other hand, suggested that results in the range of 5% represented results from a well designed algorithm. In this study, 31 of 32 results (97%) had less than 40% suboptimality and 25 of 32 (78%) had less than 10% suboptimality. These results are good by most standards, although some caution is required when heuristics are used without knowing the level of suboptimality (e.g., McDonnell et al., 2002; Fox and Beckley, 2005; Pain et al., 2005; Wilson et al., 2005; Das et al., 2006; Strange et al., 2006b; Bergle et al., 2007). An understanding of this, and the fact that many heuristic algorithms are not deterministic (i.e., random decision steps can result in different solutions with differing and unknown amounts of suboptimality each time (Pressey et al., 1997; Fischer and Church, 2005)) is important, especially as software packages and training used for complementary reserve selection become more common (Williams et al., 2004) or are used to methodically evaluate other aspects of systematic planning (e.g., Wilson et al., 2005). With this in mind, as suggested by Pressey et al. (1997), a planner using a heuristic algorithm should probably do more than just run a variety of heuristics to increase confidence. Planners could potentially achieve a very high degree of confidence in the output of their algorithm(s) by obtaining optimal results for comparison. The feasibility of this suggestion is also discussed below in the section on speed.

In practice, planning is often constrained, by limits placed on resources, to use of the MCLP model. However, we are aware of only two papers that examined suboptimality for the MCLP model, each with only one dataset (Church et al., 1996; Csuti et al., 1997; Briers, 2002, for a heuristic-only MCLP comparison; Malcolm and ReVelle, 2005, for analysis of an extended MCLP where features are covered more than one time). Thus, our analysis presents a significant contribution to understanding suboptimality in the MCLP IP model. Our results show that suboptimality occurred across all eight datasets tested, but that it tended to decrease as the number of specified sites increased. We also found that the performance of the four prioritization attributes varied across datasets, but that the richness prioritization generally was least suboptimal and often optimal. These results are valuable in a real-world context. Our findings present a strong argument for including as many sites as possible as part of a real-world planning exercise under an MCLP model when using heuristic algorithms, in contrast to many policy documents which advocate for a single protected area to represent a given target region (e.g., Parks Canada, 1997; Yukon Protected Areas Strategy, 1998; Wiersma and Urban, 2005). They also suggest that using a richness-based greedy heuristic as the primary criterion for MCLP-based site selection may be an effective strategy for minimizing suboptimality when limited to the use of heuristics. However, it is still necessary to use optimal algorithms to determine the exact level of suboptimality.

### 4.2. Solution times

The findings of our speed tests contradict many statements in the literature that optimal algorithms take much longer or are infeasible. Heuristic algorithms are still regularly used in IP binary and non-binary data projects because of the perception that optimal algorithms are unable to produce results in a useful time frame, if at all (Fuller et al., 2006; Tsuji and Tsubaki, 2004; Poulin et al., 2006; Strange et al., 2006a). This perception is likely due to the fact that when systematic reserve selection first came into widespread use, limitations in computing power and lack of knowledge about IP processes sometimes prevented quick solutions using optimal algorithms with larger datasets. For example, Pressey et al. (1996) argued that optimal algorithms such as linear programming models were computationally intensive to run, and thus felt that heuristics were preferable since they can be incorporated into real-time decision making with stakeholders, as alternative scenarios can quickly be solved and incorporated. Rodrigues et al. (2000) reiterated this idea, pointing out that the set covering IP models are known to be of a class of problems called NP-hard, where there is always a chance, depending on unknown characteristics of the dataset and the algorithm, that a solution cannot be found except by the lengthy and often infeasible process of complete enumeration. However, since then, a number of researchers have used optimal algorithms on binary and non-binary data without finding processing time to be onerous (Rodrigues and Gaston, 2002; Moore et al., 2003; Fischer and Church, 2005; Haight et al., 2005; Önal and Briers, 2005; Crossman and Bryan, 2006). The results here are in line with those findings. While we did not test datasets with non-binary variables, any differences in speed should only be due to dataset dimensions and not to differences in feasibility, as with heuristics, although this has yet to be tested to any degree for MCLP. For those datasets or algorithms where the solution is not delivered quickly, the bounds can be used to find a non-optimal solution where the suboptimality is known (Rodrigues and Gaston, 2002; Williams et al., 2005b) and which can be less suboptimal than a heuristic's final solution (e.g., Fischer and Church, 2005).

Finding a single optimal or heuristic solution quickly does not ensure that the best network design will be quickly generated. Multiple runs of either type of algorithm may be necessary to find the site network that is most appropriate or acceptable to the situation at hand (e.g., Jackson et al., 2004). Depending on the number of alternative sets required, this can take many hours even when a single solution is delivered in seconds (Vanderkam, 2005). The greedy heuristic methodology is often used interactively, and in such situations the prioritizing attribute used and the importance of site values by different stakeholders could vary at each step. We are unaware of any studies that illustrate this same degree of flexibility using optimal algorithms, but the work done here with the MCLP makes it apparent that it can be done with appropriate software development. That is, using the MCLP with an increment value of 1 for $n$, the solution set can be built using an optimal algorithm one site at a time so that stakeholder input can be considered. Our MCLP results support the value of such an approach since we showed numerous cases of greater suboptimality for heuristics that used smaller values of $n$.

## 5. Conclusion

This study comparing optimal and heuristic algorithms agrees with previous findings that heuristic algorithms sometimes yield suboptimal solutions, and adds that for the MCLP, suboptimality may increase as fewer sites are examined. The degree of suboptimality of heuristic algorithms may be inconsequential when the results of reserve selection models are applied in practice, especially since trade-offs in land selection may result in solutions that deviate from the optimal, but it may be useful to know that certain model parameters, such as the number of sites the solution is constrained to, are likely to result in higher suboptimality.

As well, our findings run counter to the still widely held perception that these problems are intractable when using IP algorithms (Moilanen, 2005a). Across nine different binary datasets, no important difference was found in the time required to find a solution using the two algorithms, indicating a possible advantage of optimal algorithms that could become important with larger datasets. Non-binary variables should be no more constrained by IP algorithms than heuristic algorithms. With increases in computing power, it is relatively easy to make use of both types of algorithm. Optimal algorithms can be used to set a benchmark against which the degree of suboptimality of heuristics, or of land allocation decisions resulting from the more general planning process, can be compared (Rothley, 1999).

Heuristic algorithms have the advantage of being flexible and intuitive and will remain an attractive tool for planners because of the ability to interface the software directly with stakeholders. Using the optimal and heuristic algorithms in concert will give planners and managers a more comprehensive set of tools to delineate reserves, and ultimately allow for better conservation of biodiversity. We see good value in the development of tools and expertise for optimal algorithm use, such as modeling non-linear objectives as IP problems (Haight et al., 2004; Moilanen, 2005b), site connectivity (Orestes Cerdeira et al., 2005), and multiple time period modeling (Snyder et al., 2004).

## Acknowledgements

The proprietary OPL code for the LSCP and MCLP models were provided by B. Bryan of the Commonwealth Scientific and Industrial Research Organization (CSIRO), Australia. Modifications were made to the MCLP code and an interface for both models was written by Richard Arnold of Carleton University. Thanks also to K. Lindsay, L. Olsen, and B. Bryan for supplying data with which to test the algorithms. We are grateful to H. Önal and P. Rubin for helpful discussions. Y. Zharikov and one anonymous reviewer provided helpful comments on an earlier version of the manuscript.

REFERENCES

Andelman, S.J., Willig, M.R., 2003. Present patterns and future prospects for biodiversity in the Western Hemisphere. Ecology Letters 6, 818–824.

Anonymous, 1999. C-Plan: Conservation planning software, user manual, version 2.2. National Parks and Wildlife Service, Armidale, NSW, Australia.

Arcese, P., Sinclair, A.R.E., 1997. The role of protected areas as ecological baselines. Journal of Wildlife Management 61, 587–602.

Bergle, R.A., Oates, J.F., Fotso, R., 2007. Distribution and protected area coverage of endemic taxa in West Africa's Biafran forests and highlands. Biological Conservation 134, 195–208.

Bojórquez-Tapie, L.A., de la Cueva, H., Díaz, S., Melgarejo, D., Alcantar, G., Solares, M.J., Grobet, G., Cruz-Bello, G., 2004. Environmental conflicts and nature reserves: redesigning Sierra San Pedro Mártir National Park, Mexico. Biological Conservation 117, 111–126.

Briers, R.A., 2002. Incorporating connectivity into reserve selection procedures. Biological Conservation 103, 77–83.

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

Cocks, K.D., Baird, I.A., 1989. Using mathematical programming to address the multiple reserve selection problem: an example from the Eyre Peninsula, South Australia. Biological Conservation 49 (2), 113–130.

Cowling, R.M., Pressey, R.L., Sims-Castley, R., le Roux, A., Baard, E., Burgers, C.J., Palmer, G., 2003a. The expert or the algorithm?—comparison of priority conservation areas in the Cape Floristic

Region identified by park managers and reserve selection software. Biological Conservation 112, 147–167.
Cowling, R.M., Pressey, R.L., Rouget, M., Lombard, A.T., 2003b. A conservation plan for a global biodiversity hotspot—the Cape Floristic Region, South Africa. Biological Conservation 112, 191–216.
Crossman, N.D., Bryan, B.A., 2006. Systematic landscape restoration using integer programming. Biological Conservation 128, 369–383.
Csuti, B., Polasky, S., Williams, P.H., Pressey, R.L., Camm, J.D., Kershaw, M., Kiester, A.R., Downs, B., Hamilton, R., Huso, M., Sahr, K., 1997. A comparison of reserve selection algorithms using data on terrestrial vertebrates in Oregon. Biological Conservation 80, 83–97.
Curio, E., 2002. Prioritisation of Philippine Island avifauna for conservation: a new combinatorial measure. Biological Conservation 106, 373–380.
Das, A., Krishnaswamy, J., Bawa, K.S., Kiran, M.C., Srinivas, V., Samba Kumar, N., Ullas Karanth, K., 2006. Prioritisation of conservation areas in the Western Ghats, India. Biological Conservation 133, 16–31.
Ferrier, S., Pressey, R.L., Barrett, T.W., 2000. A new predictor of the irreplaceability of areas for achieving a conservation goal, its application to real-world planning, and a research agenda for further refinement. Biological Conservation 93, 303–325.
Fischer, D.T., Church, R.L., 2005. The SITES reserve selection system: a critical review. Environmental Modeling and Assessment 10, 215–228.
Fox, N.J., Beckley, L.E., 2005. Priority areas for conservation of Western Australia coastal fishes: a comparison of hotspot, biogeographical and complementarity approaches. Biological Conservation 125, 399–410.
Fuller, T., Munguía, M., Mayfield, M., Sánchez-Cordero, V., Sarkar, S., 2006. Incorporating connectivity into conservation planning: a multi-criteria case study from central Mexico. Biological Conservation 133, 131–142.
Haight, R.G., ReVelle, C.S., Snyder, S.A., 2000. An integer optimization approach to a probabilistic reserve site selection problem. Operations Research 48 (5), 697–708.
Haight, R.G., Cypher, B., Kelly, P.A., Phillips, S., Ralls, K., Possingham, H.P., 2004. Optimizing reserve expansion for disjunct populations of San Joaquin kit fox. Biological Conservation 117, 61–72.
Haight, R.G., Snyder, S.A., ReVelle, C.S., 2005. Metropolitan open-space protection with uncertain site availability. Conservation Biology 19 (2), 327–337.
Jackson, S.F., Kershaw, M., Gaston, K.J., 2004. The performance of procedures for selecting conservation areas: waterbirds in the UK. Biological Conservation 118, 261–270.
Kingsland, S.E., 2002. Creating a science of nature reserve design: perspectives from history. Environmental Modeling and Assessment 7, 61–69.
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.
Malcolm, S.A., ReVelle, C., 2005. Representational success: a new paradigm for achieving species protection by reserve site selection. Environmental Modeling and Assessment 10, 341–348.
Margules, C.R., Pressey, R.L., 2000. Systematic conservation planning. Nature 405, 243–253.
Margules, C.R., Pressey, R.L., Williams, P.H., 2002. Representing biodiversity: data and procedures for identifying priority areas for conservation. Journal of Bioscience 27 (Suppl. 2), 309–326.
McDonnell, M.D., Possingham, H.P., Ball, I.R., Cousins, E.A., 2002. Mathematical methods for spatially cohesive reserve design. Environmental Modeling and Assessment 7, 107–114.
Mitchell, J.E., 2002. Branch-and-cut algorithms for combinatorial optimization problems. In: Pardalos, P.M., Resende, M.G.C. (Eds.), Handbook of Applied Optimization. Oxford University Press, New York, pp. 65–77.
Moilanen, A., 2005a. Methods for reserve selection: interior point search. Biological Conservation 124, 485–492.
Moilanen, A., 2005b. Reserve selection using nonlinear species distribution models. The American Naturalist 165 (6), 695–706.
Moilanen, A., Wintle, B.A., 2006. Uncertainty analysis favours selection of spatially aggregated reserve networks. Biological Conservation 129, 427–434.
Moore, J.L., Folkmann, M., Balmford, A., Brooks, T., Burgess, N., Rahbek, C., Williams, P.H., Krarup, J., 2003. Heuristic and optimal solutions for set-covering problems in conservation biology. Ecography 26, 595–601.
Noss, R.F., 2003. A checklist for wildlands network designs. Conservation Biology 17, 1270–1275.
Oetting, J.B., Knight, A.L., Knight, G.R., 2006. Systematic reserve design as a dynamic process: F-TRAC and the Florida Forever program. Biological Conservation 128, 37–46.
Önal, H., 2003. First-best, second-best, and heuristic solutions in conservation reserve site selection. Biological Conservation 115, 55–62.
Önal, H., Briers, R.A., 2005. Designing a conservation reserve network with minimal fragmentation: a linear integer programming approach. Environmental Modeling and Assessment 10, 193–202.
Orestes Cerdeira, J., Gaston, K.J., Pinto, L.S., 2005. Connectivity in priority area selection for conservation. Environmental Modeling and Assessment 10, 183–192.
Pain, D.J., Fishpool, L., Byaruhanga, A., Arinaitwe, J., Balmford, A., 2005. Biodiversity representation in Uganda's forest IBAs. Biological Conservation 125, 133–138.
Parks Canada, 1997. National Parks System Plan. Parks Canada, Ottawa, ON. 106 pp.
Possingham, H., Day, J., Goldfinch, M., Salzborn, F., 1993. The mathematics of designing a network of protected areas for conservation. In: Sutton, D., Cousins, E., Pierce, C. (Eds.), Decision Sciences, Tools for Today. Proceedings of the 12th Australian Operations Research Conference, Australian Society for Operations Research, Adelaide, Australia, pp. 536–545.
Possingham, H., Ball, I., Andelman, S., 2000. Mathematical methods for identifying representative reserve networks. In: Ferson, S., Burgman, M. (Eds.), Quantitative Methods for Conservation Biology. Springer-Verlag, New York, pp. 291–305.
Poulin, M., Bélisle, M., Cabeza, M., 2006. Within-site habitat configuration in reserve design: a case study with a peatland bird. Biological Conservation 128, 55–66.
Pressey, R.L., 2002. Classics in physical geography revisited. Progress in Physical Geography 26, 434–441.
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 in Ecology and Evolution 8, 124–128.
Pressey, R.L., Johnson, I.R., Wilson, P.D., 1994. Shades of irreplaceability: towards a measure of the contribution of sites to a reservation goal. Biodiversity and Conservation 3, 242–262.
Pressey, R.L., Possingham, H.P., Margules, C.R., 1996. Optimality in reserve selection algorithms: when does it matter and how much? Biological Conservation 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. Biological Conservation 80, 207–219.
Pressey, R.L., Possingham, H.P., Logan, V.S., Day, J.R., Williams, P.H., 1999. Effects of data characteristics on the results of reserve selection algorithms. Journal of Biogeography 26, 179–191.

Pullin, A.S., Knight, T.M., Stone, D.A., Charman, K., 2004. Do conservation managers use scientific evidence to support their decision-making? Biological Conservation 119, 245–252.

ReVelle, C.S., Williams, J.C., Boland, J.J., 2002. Counterpart models in facility location science and reserve selection science. Environmental Modeling and Assessment 7, 71–80.

Rodrigues, A.S.L., Gaston, K.J., 2002. Optimisation in reserve selection procedures—why not? Biological Conservation 107, 123–129.

Rodrigues, A.S., Cerdeira, J.O., Gaston, K.J., 2000. Flexibility, efficiency, and accountability: adapting reserve selection algorithms to more complex conservation problems. Ecography 23, 565–574.

Rosing, K.E., ReVelle, C.S., Williams, J.C., 2002. Maximizing species representation under limited resources: a new and efficient heuristic. Environmental Modeling and Assessment 7, 91–98.

Rothley, K.D., 1999. Designing bioreserve networks to satisfy multiple, conflicting demands. Ecological Applications 9, 741–750.

Sætersdal, M., Line, J.M., Birks, H.J.B., 1992. How to maximize biological diversity in nature reserve selection: vascular plants and breeding birds in deciduous woodlands, Western Norway. Biological Conservation 66, 131–138.

Sætersdal, M., Gjerrde, I., Blom, H.H., Ihlen, P.G., Myrseth, E.W., Pommeresche, R., Skartveit, J., Solhøy, T., Aas, O., 2003. Vascular plants as a surrogate species group in complementary site selection for bryophytes, macrolichens, spiders, carabids, staphylinids, snails, and wood living polypore fungi in a northern forest. Biological Conservation 115, 21–31.

Salomon, A.K., Ruesink, J.L., DeWreede, R.E., 2006. Population viability, ecological processes and biodiversity: valuing sites for reserve selection. Biological Conservation 128, 79–92.

Scott, J.M., Davis, F.W., McGhie, R.G., Wright, R.G., Groves, C., Estes, J., 2001. Nature reserves: do they capture the full range of America's biological diversity? Ecological Applications 11, 999–1007.

Snyder, S.A., Haight, R.G., ReVelle, C.S., 2004. A scenario optimization model for dynamic reserve site selection. Environmental Modeling and Assessment 9, 179–187.

Strange, N., Thorsen, B.J., Bladt, J., 2006a. Optimal reserve selection in a dynamic world. Biological Conservation 131, 33–41.

Strange, N., Rahbek, C., Jepsen, J.K., Lund, M.P., 2006b. Using farmland prices to evaluate cost-efficiency of national versus regional reserve selection in Denmark. Biological Conservation 128, 455–466.

Tsuji, N., Tsubaki, Y., 2004. Three new algorithms to calculate the irreplaceability index for presence/absence data. Biological Conservation 119, 487–494.

Underhill, L.G., 1994. Optimal and suboptimal reserve selection algorithms. Biological Conservation 70, 85–87.

Vanderkam, R.P.D., 2005. A comparison of heuristic and optimal reserve selection algorithms used in systematic conservation planning. MA thesis. Department of Geography and Environmental Studies, Carleton University, Ottawa, Ontario, Canada.

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

Wiersma, Y.F., Urban, D.L., 2005. Beta diversity and nature reserve system design in the Yukon, Canada. Conservation Biology 19 (4), 1262–1272.

Wiersma, Y.F., Nudds, T.D., 2006. Conservation targets for viable species assemblages: data independent targets are not appropriate. Biodiversity and Conservation 15, 4555–4567.

Williams, J.C., ReVelle, C.S., Levin, S.A., 2004. Using mathematical optimization models to design nature reserves. Frontiers in Ecology and the Environment 2, 98–105.

Williams, J.C., ReVelle, C.S., Levin, S.A., 2005a. Special issue on eco-informatics: modeling biological conservation decisions—Introduction to the special issue. Environmental Modeling and Assessment 10, 161–162.

Williams, J.C., ReVelle, C.S., Levin, S.A., 2005b. Spatial attributes and reserve design models: a review. Environmental Modeling and Assessment 10, 163–181.

Williams, P.H., 1998. Key sites for conservation: area-selection methods for biodiversity. In: Mace, G.M., Balmford, A., Ginsberg, J.R. (Eds.), Conservation in a Changing World: Integrating Processes into Priorities for Action. Symposia of the Zoological Society of London no. 72, Cambridge University Press, pp. 211–249.

Willis, C.K., Lombard, A.T., Cowling, R.M., Heydenrych, B.J., Burgers, C.J., 1996. Reserve systems for limestone endemic flora of the Cape Lowland Fynbos: iterative versus linear programming. Biological Conservation 77, 53–62.

Wilson, K.A., Westphal, M.I., Possingham, H.P., Elith, J., 2005. Sensitivity of conservation planning to different approaches to using predicted species distribution data. Biological Conservation 122, 99–112.

Yukon Protected Areas Strategy, 1998. http://www.nrtee-trnee.ca/eng/programs/Current_programs/Nature/Natural-Heritage/Documents/YPAS-Case-Study_E.pdf. Last accessed 26 January 2007.