

Integer Programming Methods for Reserve Selection and Design

Robert G. Haight and Stephanie A. Snyder

4.1 Introduction

How many nature reserves should there be? Where should they be located? Which places have highest priority for protection? Conservation biologists, economists, and operations researchers have been developing quantitative methods to address these questions since the 1980s. The first formulations (Kirkpatrick 1983; Margules et al. 1988; Chapters 3 and 19) focused on species protection and efficiency – select the minimum number of reserves from a list of candidate sites to represent all species – and were solved using iterative heuristics, which guarantee only an approximation of the optimal solution (Chapter 5). Researchers later formulated the problem as a 0–1 linear integer programming (IP) model and found mathematically proven optimal solutions using the branch and bound (B&B) method available in commercial optimization software (Saetersdal et al. 1993; Underhill 1994). The main advantage of formulating a reserve selection problem as an IP model is the availability of solution methods such as B&B that guarantee finding the optimal solution. Researchers quickly recognized this advantage and have formulated reserve selection problems with a remarkable array of objectives and constraints.

We review IP methods and formulations applied to reserve selection and design. Because IP is a branch of linear programming (LP), we begin with a discussion LP models, assumptions, solution methods, and linearization techniques (Section 4.2). Then, we use a fundamental reserve selection model called the maximum species covering problem (MSCP) to discuss IP formulations and solution

methods (Section 4.3). In Section 4.4, we discuss computational aspects of solving the MSCP. In Sections 4.5–4.8, we present a range of IP extensions of the MSCP (Table 4.1). We conclude with a discussion of the limitations of IP models and directions for future work.

Throughout the chapter, we attach specific meaning to the terms ‘site’, ‘reserve’, and ‘reserve system’. Following Williams et al. (2005a), a site is a selection unit – a piece of land that may be selected for protection. A site is usually undeveloped open space belonging to one or more cover types, including forest, grassland, pasture, or cropland. A reserve is a single site or a contiguous cluster of sites that has been selected for protection. A reserve system or network is a set of multiple, spatially separated reserves. We also distinguish reserve selection from reserve design models. Reserve selection models identify sites to protect to maximize some measure of biological diversity (e.g. species richness) subject to a budget constraint with no consideration given to the spatial attributes of the reserve system. Reserve design problems use spatial attributes of the reserve system as objectives.

Excellent reviews of IP models for reserve selection and design are available. Rodrigues and Gaston (2002) provide a comprehensive list of published studies that use IP formulations of reserve selection problems. ReVelle et al. (2002) discuss the relationship between IP reserve selection models and their counterparts in the facility location literature. Williams et al. (2005a) discuss spatial attributes of reserve systems and how they can be incorporated as objectives in IP reserve design models.

Table 4.1 Reserve selection (A) and design problems (B) that have been formulated and solved using integer programming methods. Reserve selection models identify sites to protect to maximize some measure of biological diversity (e.g. species richness) subject to a budget constraint with no consideration given to the spatial attributes of the reserve system. Reserve design problems use spatial attributes of the reserve system as objectives

Problem	Objective	Reference
A. Non-spatial reserve selection problems		
Maximum species covering	Maximize number of species protected for a given budget	Church et al. (1996)
Bi-criteria reserve selection	Maximize number of species protected and some other conservation objective	Church et al. (2000); Ruliffson et al. (2003)
Reserve selection with uncertain species presence	Maximize expected number of species protected for a given budget	Camm et al. (2002); Arthur et al. (2004)
Dynamic reserve selection with uncertain site availability	Maximize expected number of species protected at end of horizon	Haight et al. (2005)
B. Spatial reserve design problems		
Reserve proximity	Minimize sum of pairwise distances between reserves	Önal and Briers (2002)
Reserve connectivity	Maximize number of adjacent reserves	Nalle et al. (2002)
Reserve compactness	Minimize boundary length of reserves	Fischer and Church (2003)
Reserve core and buffer areas	Maximize core area protected	Williams and ReVelle (1998)

4.2 Linear programming

Linear programming is a mathematical modelling and optimization technique invented in the 1940s (see Dantzig 1963). LP problems involve finding values of decision variables to optimize a linear objective function subject to linear equality and inequality constraints. The related problem of IP requires some or all of the variables to take integer values. LP and IP methods have proved valuable for modelling problems in planning, routing, scheduling, assignment, and design in a wide range of industries. An excellent guide to LP and IP problems, algorithms, textbooks and software is available online (Fourer 2000).

4.2.1 Land allocation problem

LP models have been used for decades in forest management to allocate land to mutually exclusive uses such as timber production and wildlife habitat (see Hof and Bevers 1998 for review). We describe a simple version of this problem to illustrate the assumptions of LP models. Suppose a planner has forest sites that can be used for timber production or wildlife habitat and must allocate a proportion of

each site to each use. The planner knows the abundance of each species in each site and the minimum total abundance desired for each species across sites. To put a cost on habitat protection, the planner uses the revenue of foregone timber production. The land allocation problem is an LP model with the following notation:

i, I = index and set of sites,

j, J = index and set of species,

a_{ij} = abundance of species j in site i ,

c_i = cost of allocating a unit of site i to wildlife habitat,

T_j = minimum desired abundance of species j ,

x_i = proportion of site i allocated to wildlife habitat,

z = objective function value.

The model is formulated as follows:

$$\text{Minimize } z = \sum_{i \in I} c_i x_i \quad (4.1)$$

$$\sum_{i \in I} a_{ij} x_i \geq T_j \quad \text{for all } j \in J \quad (4.2)$$

$$0 \leq x_i \leq 1 \quad \text{for all } i \in I \quad (4.3)$$

The objective is to minimize the cost of allocating land to habitat (Equation 4.1) subject to a set of constraints requiring a minimum abundance of each

species across sites (Equation 4.2) and restrictions on the decision variables (Equation 4.3).

This land allocation problem illustrates three properties of linear programmes: proportionality, additivity, and continuity. The proportionality assumption states that if it costs c_i units to protect all of site i (i.e. $x_i=1$), then it costs $c_i x_i$ to protect a proportion x_i of site i . Likewise, if a_{ij} units of species j are produced by protecting all of site i , then $a_{ij} x_i$ units of species j are produced by protecting a proportion x_i of site i . The proportionality assumption implies constant returns to scale (e.g. the unit cost of protecting a proportion of site i , $c_i x_i$, does not depend on the proportion of site i protected, x_i). The additivity assumption states that the total contribution of all variables equals the sum of the individual variable contributions regardless of the values of the variables. The additivity assumption implies that the objective function is separable in the variables: $z(x_1, \dots, x_n)$ is separable in variables x_1, \dots, x_n if z can be written as a sum of n functions each of which involves only one variable in the model (i.e. $z = [z_1(x_1) + \dots + z_n(x_n)]$, where $z_i(x_i)$ is the contribution of the variable x_i to the objective function). The third assumption is continuity of the variables: each variable can take on all real values in its allowed range. In the optimization problem above, the decision variable for the proportion of each site protected, x_i , can take on any real value between 0 and 1 (Equation 4.3).

4.2.2 Solving LP problems

The importance of LP problems derives from the existence of general purpose (independent of the problem being solved) and computationally effective (able to solve large problems) solution algorithms. LP problems having tens or hundreds of thousands of continuous variables are regularly solved on Pentium-based personal computers or Unix workstations. Two families of solution algorithms are in wide use: the simplex algorithm introduced in the 1940s (Dantzig 1963) and interior point methods introduced in the 1980s (Kamarkar 1984). Both visit a progressively improving series of trial solutions until a solution is reached that satisfies the mathematical conditions for optimality.

LP software comes in two related but different kinds of packages. The first, algorithmic codes, finds and lists optimal solutions to specified LP problems. The second, modelling systems, helps people solve LP problems by taking a description of an LP problem in a straightforward, logical format, converting the model to a form required by the algorithmic code, and displaying the results of the optimal solution. Modelling systems include programming languages that allow users to specify models in concise algebraic statements (e.g. Generalized Algebraic Modelling System, GAMS) and spreadsheets in which models are represented as systems of linear equations (e.g. Microsoft Excel). Most modelling systems support a variety of algorithmic codes, while the more popular codes can be used with many different modelling systems. The Institute for Operations Research and the Management Sciences (INFORMS) regularly publishes surveys of commercial modelling systems and algorithmic codes, including both LP and IP solvers (Fourer 2007).

4.2.3 Linear approximations of non-linear optimization problems

While LP problems are easy to solve, many real-life problems are better expressed with non-linear equations and inequalities that violate the proportionality and additivity assumptions. It may be possible to approximate the non-linear parts of the problem with linear approximations that satisfy the additivity and proportionality assumptions and then use LP to solve the approximation. For example, suppose the unit cost of site protection c in the land allocation model above does not vary by site so the objective function (Equation 4.1) is $z = c(x_1 + \dots + x_n)$ where n is the number of sites. Further, suppose unit cost c increases as a linear function of the total amount of protected land, $c = b(x_1 + \dots + x_n)$, where b is the slope of the unit cost function. Then, the objective function is $z = b(x_1 + \dots + x_n)(x_1 + \dots + x_n)$, which violates both additivity and proportionality assumptions. We can address the additivity violation by defining a new variable $y = x_1 + \dots + x_n$ which is a linear function of the decision variables, and writing the objective $z = by^2$. While this objective is

separable in the variables (it has only one variable!), it still violates the proportionality assumption. The violation of the proportionality assumption can be addressed with a piecewise linear approximation. First, divide the total amount of protected land, y , into a set of cost classes, K , ordered from lowest cost to highest cost, where c_k is the unit cost of protecting an amount of land in class k and $c_1 < c_2 < c_k$. Each c_k is a linear approximation of the slope of the total cost curve by^2 in the k th interval of y . Further, define y_k as a decision variable for the total amount of land selected for protection in cost class k and d_k be the upper bound on the amount available in class k , so that $0 < y_k < d_k$. Then, we can formulate the following linear approximation of the non-linear land allocation problem:

$$\text{Minimize } z = \sum_{k \in K} c_k y_k \quad (4.4)$$

$$\sum_{k \in K} y_k = \sum_{i \in I} x_i \quad (4.5)$$

$$\sum_{i \in I} a_{ij} x_i \geq T_j \quad \text{for all } j \in J \quad (4.6)$$

$$0 \leq x_i \leq 1 \quad \text{for all } i \in I \quad (4.7)$$

$$0 \leq y_k \leq d_k \quad \text{for all } k \in K \quad (4.8)$$

The non-linear objective function $z=by^2$ is replaced by a piecewise linear approximation (Equation 4.4), and Equations 4.5 and 4.8 define conditions for the decision variables y_k . If the model selects any land for protection, the model will select land with the lowest unit cost first. As a result, for any k , if $y_k > 0$, then $y_t = d_t$ for all $t < k$ and Equation 4.4 is piecewise linear. The major disadvantage of formulating a linear approximation of a non-linear problem is the considerable increase in the number of variables and constraints. Nevertheless, linear approximations of non-linear site selection models have been formulated with thousands of decision variables and successfully solved using LP software (e.g. Hof and Raphael 1997; Arthur et al. 2004).

4.3 Integer programming

Integer programming methods are used to solve linear optimization problems in which one or

more of the variables are restricted to be integers. Zero-one or binary IP problems restrict their integer variables to be 0 or 1. Many reserve selection and design problems are formulated as binary IP problems because site selection decisions are binary and the logic of the conservation objective can be modelled with binary variables. To illustrate a 0-1 IP problem, we present the maximum species covering problem, one of the first of many reserve selection and design problems formulated in the conservation biology literature.

4.3.1 Maximum species covering problem

Suppose a planner has identified a set of potential reserve sites and the cost of protecting each site. The planner only knows whether or not each species is present in each site and does not have information on its abundance. Further, the planner must decide whether or not to protect each site in total and cannot protect a portion of the site. This restriction fits the common situation in which sites are indivisible ownerships. The MSCP identifies the sites to protect to maximize the number of species represented – where a species is represented if it is present in at least one protected site – subject to a budget constraint (Church et al. 1996). The MSCP is analogous to the maximum covering location problem in the location science literature (Church and ReVelle 1974; ReVelle et al. 2002), and it provides sets of sites that efficiently achieve conservation goals and trade-offs between conservation goals and trade-offs between conservation goals. The MSCP is a binary IP problem with the following notation:

- i, I = index and set of potential reserve sites,
- j, J = index and set of species in need of protection,
- a_{ij} = 0-1 parameter: 1 if species j is present in site i , 0 otherwise,
- B = upper bound on budget,
- c_i = cost of protecting site i ,
- x_i = 0-1 variable: 1 if site i is selected for protection, 0 otherwise,
- y_j = 0-1 variable: 1 if species j is represented in at least one protected site, 0 otherwise,
- z = objective function value.

The model is formulated as follows:

$$\text{Maximize } z = \sum_{j \in J} y_j \quad (4.9)$$

$$\sum_{i \in I} a_{ij} x_i \geq y_j \quad \text{for all } j \in J \quad (4.10)$$

$$\sum_{i \in I} c_i x_i \leq B \quad (4.11)$$

$$x_i, y_j \in \{0, 1\} \quad \text{for all } i \in I \text{ and } j \in J \quad (4.12)$$

The objective (Equation 4.9) is to maximize the number of species that are represented or covered in the set of selected sites. Equation 4.10 enforces the logic of covering: a species is covered ($y_j=1$) if at least one site that contains the species is selected for protection. Equation 4.11 is the budget constraint. Equation 4.12 describes the integer restrictions on the variables.

Note that the same IP model structure could include species abundance data rather than presence-absence data. In this case, the abundance parameter a_{ij} would be added to the objective function (Equation 4.9) to maximize the total abundance of species contained in the selected set of sites:

$$\text{Maximize } z = \sum_{i \in I} \sum_{j \in J} a_{ij} y_j$$

Note also that the MSCP problem is written in terms of maximizing species coverage subject to a budget constraint. This budget-constrained formulation fits a common situation in which resources for site protection are limited and the decision-maker wants to allocate resources to optimize a conservation objective. Solving the problem for a given budget B allows the determination of an efficient set of sites, where efficiency means that there are no other sets of sites that provide a higher level of species coverage and stay within the budget. Solving the problem with increasing budgets allows construction of a cost curve, which shows the cost of increasing the number of species covered. Box 4.1 describes an application of the MSCP in Lake County, Illinois, USA, where planners want to determine the impact of budget restrictions on efficient sets of sites for species protection in the face of urban development.

4.3.2 Solving IP problems

The integer requirements on the decision variables make IP problems difficult to solve. In contrast to LP problems, there are no general-purpose and

Box 4.1

We describe an application of the maximal species covering problem (Equations 4.9–4.12) in a case study in the Lake County portion of the Fox River watershed north-west of the city of Chicago (Figure 4.1). In response to rapid population growth and conversion of open space to housing and commercial development, Lake County planners want to acquire land to protect rare animals and plants and provide equitable access to recreation. To help planners identify cost-effective sets of sites, we solved an MSCP and analysed the cost of increasing the number of species covered.

The analysis is conducted using data for 31 privately owned open-space sites (see Haight et al. 2005 for details). The sites vary in size from 1 to 313 ha, with median of 29 ha. Each site is described by a list of rare plants and animals present. Collectively, 27 rare species occur in the

31 sites, and species richness of individual sites varies from 1 to 9 species. We used areas of sites as proxies for site costs and assumed that the planner has an area budget for selecting sites.

We determined the optimal sets of sites to protect for area budgets ranging from 1 to 618 ha and plot the cost curve in Figure 4.2. Each point represents a cost-effective set of sites for a given budget. The slope of the cost curve is the marginal cost of species protection – the area required to protect an additional species. Marginal cost is small (4 ha/species) as coverage increases from 5 to 20 species, moderate (34 ha/species) in the range of 20 to 25 species, and large (195 ha/species) for levels of species coverage greater than 25.

As the budget increases, the optimal set of sites is not always found by adding another site to the previously selected set. For example, to increase

continues

Box 4.1 continued

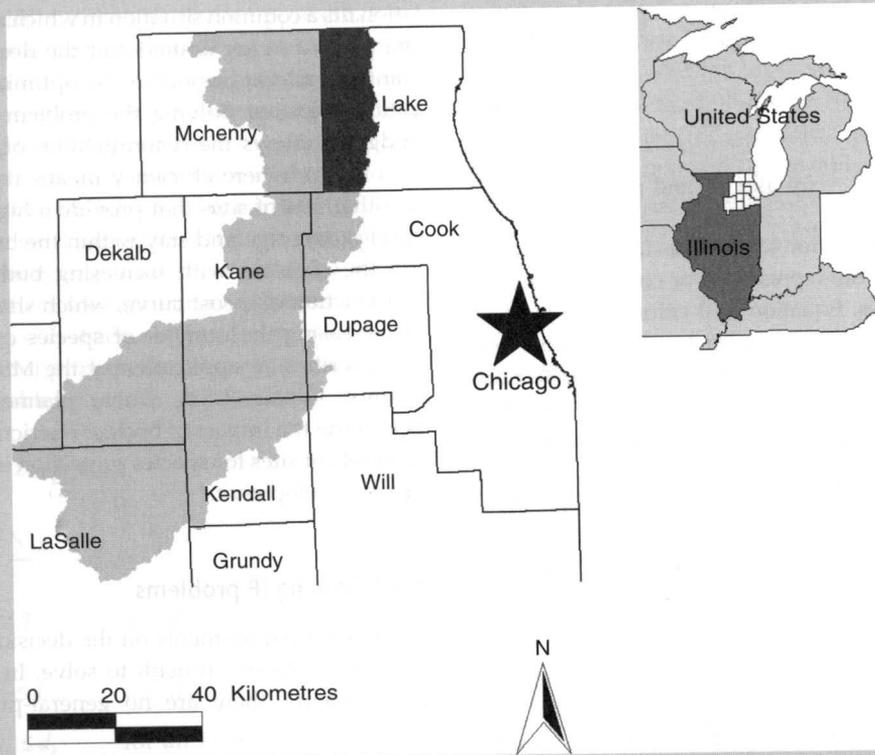


Figure 4.1 Fox River watershed (shaded grey) in counties of north-eastern Illinois, USA. The study area (shaded black) is the north-eastern portion of the watershed located in Lake County, Illinois, USA.

coverage from 20 to 22 species, one site can be added to the list of protected sites (Table 4.2). However, increasing species coverage above 22 species involves dropping one site and adding

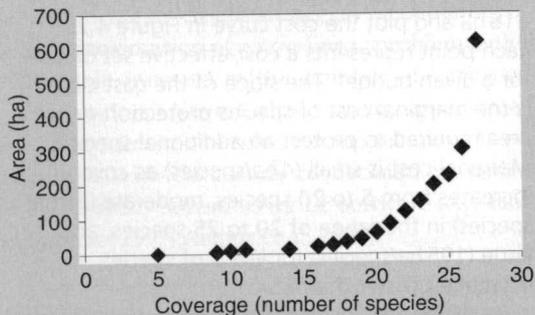


Figure 4.2 Cost curve showing area protected versus number of species covered for the site selection options in the Fox River watershed of Lake County, Illinois, USA.

Table 4.2 Optimal sets of sites selected for protection under increasing area budgets in the Fox River watershed of Lake County, Illinois, USA

Objective values	Site numbers selected for protection												
Species	Area (ha)	3	7	8	14	15	17	18	20	21	22	23	30
20	59		X	X	X		X	X		X			X
22	123		X	X	X	X		X	X		X		X
25	228		X	X	X		X	X	X	X			X
27	618		X	X	X		X	X		X	X	X	X

up to four others. There is consistency in sites selected for protection. Six sites are selected whenever the budget is greater than 50ha. These sites are small (<16ha), have higher numbers of species per hectare, and contain endemics.

computationally effective algorithms for solving IP problems. Solution methods for IP problems can be categorized as 'optimal' and 'heuristic'. An optimal algorithm is one that mathematically guarantees finding the optimal solution while a heuristic is an algorithm that should find a feasible solution which, in objective function terms, is close to the optimal solution. The choice between using an optimal or heuristic algorithm depends in part on whether the problem size is beyond the computational limit of the optimal algorithm and whether the analyst wants to spend the effort needed to solve the problem optimally.

The most effective optimal algorithm for IP problems is LP-based branch and bound (B&B). The algorithm proceeds by solving a sequence of LP problems using either efficient simplex method or interior-point method until a mathematically proven optimal solution to the IP is found. The B&B algorithm begins by solving a 'relaxed' LP version of the problem, ignoring the binary integer restrictions on the decision variables. If this relaxed solution is all-integer, then it is the optimal solution and the algorithm terminates. If the relaxed solution contains one or more binary variables with fractional values, the problem is iteratively branched into sub-problems. Branching is done by picking one of the binary variables with fractional value in the preceding solution. Two new sub-problems are created by setting this variable to 1 and 0, respectively, and the remaining binary variables are allowed to have continuous values between 0 and 1. In this manner, each sub-problem is solved as a linear programme. This sequence of creating sub-problems and solving linear programmes is continued until either an integer solution is found for each sub-problem branch or the relaxation causes the sub-problem to become infeasible. An exhaustive search of all of the sub-problems can be prohibitive because 2^n problems must be solved when there are n binary variables. To reduce the number of sub-problems, the B&B algorithm employs a 'bounding' feature, which computes a bound on the optimum solution at each step. If the solution to a sub-problem is worse than the best integer solution previously found, then that sub-problem is discarded, thereby reducing the search space. Most commercial

modelling packages contain B&B algorithms (Fourer 2007).

4.4 Computational aspects of IP models

The computer time required to solve any particular IP problem is hard to predict. Problems with a hundred variables can be challenging, while others with tens of thousands of variables solve readily. Examples of the MSCP with hundreds or thousands of 0–1 decision variables have been solved using B&B software in one to several minutes (Church et al. 1996; Rodrigues and Gaston 2002; Snyder et al. 2004a); however, reserve design problems with spatial objectives may take much longer to solve. In this section, we offer tips to explore when solving the MSCP with B&B.

4.4.1 Formulation tips

Camm et al. (1996) recommend carefully pre-processing MSCP datasets to reduce problem size prior to attempting solution with B&B. Specifically, they suggest removing any eligible site from consideration if it contains all the same species or a subset of the species of another site that costs less. Similarly, if two species are found in the same set of eligible sites, then only one of the species needs to be included.

As recognized by Camm et al. (1996) and Church et al. (1996), the species coverage variables, y_j , need not be explicitly defined as binary decision variables. The structure of the MSCP and the binary site selection variables, x_i , will automatically produce integer values for the coverage variables when solved via B&B. As a result, the coverage variables can be defined as non-negative variables with an upper bound of 1. This reduction in the number of integer variables may provide computational savings because the solution difficulty of integer programmes is often significantly influenced by the number of integer variables.

A mathematically equivalent specification of the MSCP can be formulated based upon a 'minimal uncovering' model, which may have some computational advantage (Church and ReVelle 1974;

Church et al. 1996; Arthur et al. 1997; Snyder et al. 2004a). In the uncovering formulation, each decision variable for species coverage is 0 if the species is represented in at least one site and 1 if the species is unrepresented. The objective of this new formulation is to minimize the number of unrepresented species. As a result, the uncoverage variables can be declared non-negative (rather than integer) variables. This lack of a required upper bound may make the minimal uncovering version of the MSCP more computationally efficient than the equivalent maximal covering version.

4.4.2 Solver settings – optimality gap

The B&B algorithm provides, for each feasible solution obtained, a provable upper bound on its distance from optimality. Each time a sub-problem is solved, the algorithm calculates the absolute and percent deviation between the objective function value of the current best feasible integer solution and the current upper bound on the optimal objective function value. Based on this deviation, the analyst can specify an 'optimality gap' that functions as a termination criterion (McDill and Braze 2001; Önal 2003; Snyder et al. 2004a). For example, to ensure that the optimal solution is found, the analyst would set the optimality gap equal to 0. Then, all sub-problem branches will be explored as long as a non-0 difference exists between the best current solution and the current upper bound. When the problem takes too long to solve to optimality, an analyst can set the optimality gap to a non-0 value (e.g. 1%), which reduces problem size and hastens algorithm termination. However, allowing a non-0 optimality gap also allows for the possibility of a suboptimal solution. Thus, one must be judicious in setting the optimality gap because better solutions could be missed by allowing the algorithm to terminate before reaching optimality.

An important feature of the MSCP allows the absolute optimality gap to be just less than 1 while still guaranteeing that the optimal solution will be found (Önal 2003). Because the objective function of the MSCP can only take integer values (e.g. number of species covered), no improvement in the objective function value is possible when the absolute

optimality gap is below 1. Then, the current integer solution is optimal and the algorithm can terminate. This is an important feature of the MSCP that should be exploited when using B&B. If the objective function includes non-binary coefficients on the decision variables (e.g. weights for species importance or rarity), then optimality gaps less than one no longer guarantee optimal solutions.

4.5 Bi-criteria reserve selection problem

While the MSCP has traditionally focused on the single objective of maximizing species representation, additional objectives can be considered in bi-criteria formulations, including spatial attributes of the reserve systems (e.g. Rothley 1999; Cerdeira et al. 2005; Önal and Briers 2005, 2006; Alagador and Cerdeira 2007), total habitat area (Snyder et al. 2004a), habitat quality (Church et al. 2000), and public access or proximity (Ruliffson et al. 2003; Haight et al. 2005). Cohon (1978) provides a comprehensive discussion of multi-objective programming theories and methods. While solution of a single objective problem yields an optimal solution, solution of a multi-objective model yields a set of solutions that are termed non-dominated or non-inferior. A non-inferior solution is one in which no other feasible solution exists to the problem that will lead to an increase in the value of one objective without simultaneously causing a degradation in the value of another. The choice of a preferred solution from the set of non-inferior solutions depends on the preferences of a decision-maker.

To illustrate a bi-criteria problem, we extend the MSCP to handle a second objective of maximizing the number of people with access to reserves (Ruliffson et al. 2003). We assume that people in a town have access if one or more reserves are within a required distance of the town. The data for each town include its population and a list of sites that are within the required distance. In addition to the notation of the MSCP, the bi-criteria site selection model has the following notation:

k, K = index and set of towns,

Q_1 = number of species represented in the protected sites,

Q_2 = number of people with access to protected sites,

r_k = number of people in town k ,

M_j = set of sites that contain species j ,

N_k = set of sites that are within the required distance of town k ,

z_k = 0–1 variable: 1 if town k has at least one protected site within the required distance, 0 otherwise.

The model is formulated as follows:

$$\text{Maximize } Q_1 = \sum_{j \in J} y_j \quad (4.13)$$

$$\text{Maximize } Q_2 = \sum_{k \in K} r_k z_k \quad (4.14)$$

$$\sum_{i \in N_k} x_i \geq z_k \quad \text{for all } k \in K \quad (4.15)$$

$$\sum_{i \in M_j} x_i \geq y_j \quad \text{for all } j \in J \quad (4.16)$$

$$\sum_{i \in I} c_i x_i \leq B \quad (4.17)$$

$$x_i, y_j, z_k \in \{0, 1\} \quad \text{for all } i \in I, j \in J, k \in K \quad (4.18)$$

The problem has two objective functions: maximize the number of species represented in protected sites (Equation 4.13) and maximize the number of people with access to protected sites (Equation 4.14). Public access is the number of towns with access weighted by population size, r_k . Equation 4.15 is the condition under which town k has access (i.e. $z_k = 1$): at least one site that is within the required distance of town k must be selected for protection. Equations 4.16–4.18 define species coverage, the budget constraint, and the integer restrictions on the variables.

One approach to solving a bi-criteria optimization problem is the weighting method, which creates a single objective function as a weighted sum of the two objectives. In our case, the problem is to maximize the weighted sum, $wQ_1 + (1-w)Q_2$, where $0 \leq w \leq 1$. The value of the weight, w , is systematically varied between 0 and 1, and the problem re-solved many times to produce an estimate of the non-inferior set of solutions. The weight w represents the decision-maker's position on the relative importance of the two objectives.

The constraint method is another approach to solving bi-criteria problems in which one of the objectives is transformed into a constraint. For example, the bi-criteria problem above could be

solved by optimizing the species' representation objective (Q1) subject to a constraint requiring the total number of people with access to protected sites to be greater than, or equal to, some specified threshold:

$$\sum_{k \in K} r_k z_k \geq T \quad (4.19)$$

The value of the parameter T would then be systematically varied and the problem resolved to yield an estimate of the non-inferior set of solutions.

While both solution methods are effective means of transforming a bi-criteria problem into a problem with a single objective function and generating an estimate of the non-inferior set of solutions, there are some potential computational differences. As ReVelle (1993) suggests and Snyder et al. (2004a) illustrate, an MSCP model with a constraint in which the coefficients are not 0 or 1 (e.g. Equation 4.19) is not likely to be integer-friendly (i.e. a structure that is amenable to integer solutions) nor solve quickly to optimality. In such cases, including the constraint as an objective in a bi-criteria optimization formulation using the weighting method might be computationally more efficient. A computational issue known as gap points, however, can arise when the weighting method is applied to integer models (Cohon 1978). Gap points are non-inferior solutions to a multi-objective integer model that cannot be found using the weighting method because they are located within the interior of the convex hull of the trade-off curve formed by the non-gap solutions. These solution points can be found, however, through the use of the constraint method.

For these computational reasons, analysts use both the weighting method and constraint method. The weighting method is used to quickly generate an estimate of the non-inferior set of solutions. Then, if an analyst wanted to hone in on a particular segment of the curve, the constraint method is used to explore small ranges of the curve that the weighting method might miss as a result of gap points. Box 4.2 illustrates the bi-criteria site selection problem (Equations 4.13–4.18) with results from an application to Lake County, Illinois, USA.

Box 4.2

We used the bi-criteria site selection model (Equations 4.13–4.18) to analyse how the optimal set of protected sites in Lake County, Illinois, USA, varies as we trade off species representation and public access under different budgets. There are 34 towns in western Lake County. Based on the 2000 U.S. Census, the towns collectively held 222,000 people, and individual towns were home to from 1,000 to 30,000 people. We assume that people in a town have access to a site if the site is within 3.2 km of the town, and we know the sites that are within the required distance of each town. All towns have at least one site within 3.2 km.

We used the multi-objective weighting method to solve the bi-criteria problem. We computed optimal sets of sites for problems in which the objective function weight is decreased from 1.0 to 0.0 in increments of 0.05 subject to area budgets of 81 and 200 ha. The curves showing the trade-offs between species representation and public access have concave shapes in which species representation drops as public access increases (Figure 4.3). The points on each curve represent non-dominated sets of sites and their performance with respect to the two objectives for a given budget. For each non-dominated set of sites, improvement in one objective cannot be achieved without simultaneously causing degradation in the value of the other objective. As a result, the points on each trade-off curve represent a frontier beyond which no better solutions can be found.

Among the non-dominated solutions for a given budget, the best depends on the decision-maker's preference for the two objectives. If species representation is most important and the budget is 81 ha, the choice is alternative A, in which species representation is 20 (74 % of the maximum representation without a budget constraint) and public access is 73,000 people (33 % of the maximum accessibility). The dashed horizontal line between the y-axis and point A indicates that several solutions exist with the same species representation as alternative A, but with less public access. The highest level of public access (point D, 91,000 people) is obtained with a 35 % reduction in species representation. The dashed vertical line from point D to the x-axis indicates that several solutions exist with the same level of public access as alternative D but with less species representation. Increasing the budget from 81 ha to 200 ha shifts the trade-off curve up and to the right while reducing the trade-off between objectives.

To complement the trade-off curves, we look at the site selection decisions and identify core sites, which are sites selected for protection regardless of the weights given to the objective functions. With a budget of 81 ha, three core sites are protected in all four solutions (Table 4.3). With a budget of 200 ha, there are four additional core sites. The core sites are typically small (<30 ha) and have relatively large numbers of species and people with access.

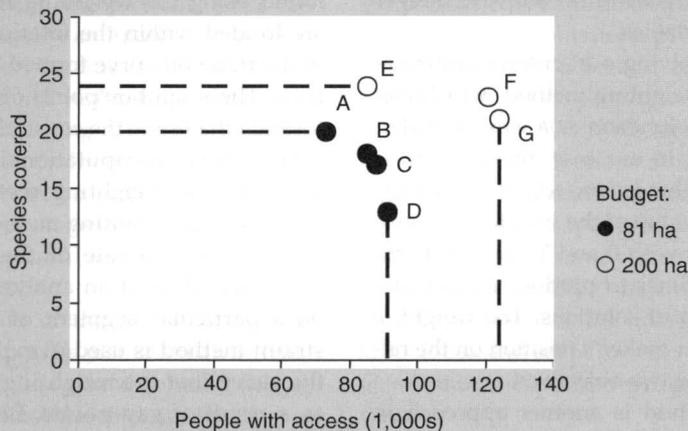


Figure 4.3 Trade-offs between open-space protection objectives of maximizing species coverage and maximizing public access under different area budgets.

Table 4.3 Objective function values and sites selected for protection for non-dominated solutions with area budgets of 81 ha (solutions A, B, C, D) and 200 ha (solutions E, F, G) in the Fox River watershed of Lake County, Illinois, USA

Solution	Objective value		Site numbers selected for protection											Area Protected (ha)	
	Species	People (1000s)	3	7	8	14	15	16	17	18	19	20	21		30
A	20	73		X		X			X	X		X	X	X	73
B	18	85		X	X	X			X	X		X			75
C	17	88		X	X			X	X	X		X	X	X	80
D	13	91		X	X					X	X	X	X		81
E	24	85	X	X	X		X		X	X			X	X	198
F	23	121		X	X	X	X	X	X	X		X	X	X	196
G	21	124		X	X		X	X	X	X	X	X	X		193

4.6 Reserve selection with uncertain species presence

In many cases, species presence in each site is not known with certainty and expressed as a probability of occurrence. The MSCP can be extended to handle probabilities of occurrence and maximize the expected number of species covered subject to a budget constraint. Let p_{ij} be the probability that species j is present in site i where the probability of species presence is independent of its occurrence in neighbouring sites. This is an important assumption because it allows us to write the probability that species j is not covered in the sites selected for protection as a product of the absence probabilities over all sites:

$$v_j = \prod_{i \in I} (1 - p_{ij})^{x_i} \quad \text{for all } j \in J \quad (4.20)$$

where x_i is the 0–1 decision variable for whether or not site i is selected for protection. If occurrence probabilities are not independent, then we would have a much more complicated expression for v_j . With the independence assumption, the problem is to select sites to maximize the expected number of species covered subject to a budget constraint:

$$\text{Maximize: } \sum_{j \in J} (1 - v_j) \quad (4.21)$$

$$\sum_{i \in I} c_i x_i \leq B \quad (4.22)$$

$$x_i \in \{0, 1\} \quad \text{for all } i \in I \quad (4.23)$$

A linear approximation of this non-linear problem can be solved using IP methods (Camm et al. 2002),

and the model has been illustrated using probabilistic occurrence data for 403 terrestrial vertebrates in 147 candidate sites in western Oregon, USA (Arthur et al. 2004).

In some situations, the decision-maker may be concerned about the likelihood that a subset of endangered species is represented in the reserves. This concern can be addressed by imposing additional constraints for minimum coverage probabilities for target species (Haight et al. 2000). Letting E be the set of endangered species and h_j be the minimum coverage probability for endangered species j (e.g. 0.95), the minimum threshold coverage constraints are

$$1 - v_j \geq h_j \quad \text{for all } j \in E \quad (4.24)$$

Rearranging and taking the natural logarithm produces an equivalent set of linear constraints:

$$\sum_{i \in I} x_i \ln(1 - p_{ij}) \leq \ln(1 - h_j) \quad \text{for all } j \in E \quad (4.25)$$

Arthur et al. (2004) added these constraints to the maximum expected species covering problem to estimate the trade-offs between total species coverage and the likelihood of endangered species representation.

4.7 Dynamic reserve selection with uncertain site availability

The MSCP assumes that site selections are made all at once and protection takes place rapidly before site degradation or loss. In practice, decisions are

made sequentially with budget restrictions and uncertainties about site degradation and loss. Although dynamic site selection problems and heuristic solution methods are discussed in detail in Chapter 10, we discuss how to formulate and solve a dynamic site selection problem as an IP model.

Snyder et al. (2004b) developed a two-period linear-integer model for sequential site selection in which uncertainty about future site availability is represented with a set of probabilistic scenarios. The two-period problem maximizes the expected number of species covered at the end of the second period subject to an upper bound on the total cost of site protection. The model employs a list of sites, some of which are available for protection in the first period and others which are not. Each site, not protected in the first period, has a probability of remaining undeveloped and being available for protection in the second period. Uncertainty about the development of unprotected sites is represented with a set of development scenarios. Each scenario is one possible development outcome identifying which sites are undeveloped and available for protection in the second period. Associated with each scenario is a probability of occurrence. The model has two sets of 0–1 site selection variables. The first set includes the protection choices for sites in the first period. The model assumes that protection decisions in the second period are made after the decisions in the first period are implemented and the site development scenario is revealed. Thus, the second set of decision variables includes the protection choices for sites in the second period under each development scenario. The two-period problem is readily solvable using IP methods (Snyder et al. 2004b) and provides information about how uncertain site availability affects current site selection decisions (Haight et al. 2005).

4.8 Spatial reserve design problems

One shortcoming of reserve selection models is that they do not consider the spatial distribution of selected sites. As a consequence, MSCP solutions may consist of scattered reserves with little spatial coherence. A scattered distribution of reserves

may not maintain or support long-term persistence of target species (Cabeza and Moilanen 2001) and may increase the difficulty and expense of reserve management. Scattered reserves are particularly troublesome when they are surrounded by a matrix of land uses and cover types that adversely impact species' persistence.

Reserve design attributes such as reserve proximity, connectivity, and shape can be incorporated into IP models for site selection (Williams et al. 2005a). Here, we discuss how these attributes can be formulated as spatial objectives in IP models. Each model includes a spatial objective combined with a species coverage constraint. By varying the level of the constraint, trade-offs between the spatial and coverage objectives can be obtained.

4.8.1 Reserve proximity

When a reserve system consists of disjunct areas of protected habitat, the distance between reserves may influence species' mobility and viability. A reserve system in which the reserves are closer together may be preferred because shorter migration distances facilitate recolonization of areas where a species has become locally extinct and help prevent the loss of genetic diversity because of inbreeding. One way to reduce the distances between reserves is to minimize the sum of distances between all pairs of selected sites. Letting d_{ik} be the distance between sites i and k and u_{ik} be a 0–1 variable for whether or not both sites i and k are selected, the problem can be written:

$$\text{Minimize: } \sum_{i \in I} \sum_{k > i} d_{ik} u_{ik} \quad (4.26)$$

$$u_{ik} \geq x_i + x_k - 1 \quad \text{for all } i, k \in I, k > i \quad (4.27)$$

$$\sum_{i \in M_j} x_i \geq y_j \quad \text{for all } j \in J \quad (4.28)$$

$$\sum_{j \in J} y_j \geq R \quad (4.29)$$

$$x_i, y_j \in \{0, 1\}, u_{ik} \in \{0, 1\} \quad \text{for all } i, k \in I, j \in J \quad (4.30)$$

The objective (Equation 4.26) minimizes the sum of the pairwise distances between selected sites subject to constraints (Equations 4.28 and 4.29) that require at least R species to be represented.

Equation 4.27 enforces the definition of u_{ik} by requiring both $x_i = 1$ and $x_k = 1$ for $u_{ik} = 1$. Önal and Briers (2002) apply a similar formulation to the problem of selecting a subset of 131 pond sites in Oxfordshire, UK, to protect 256 invertebrate species. Variants of this approach include maximizing the inverse pairwise distance between all selected reserves (Rothley 1999), minimizing the maximum intersite distance between the selected reserves (Önal and Briers 2002), constraining the maximum distance between eligible reserves (Malcolm and ReVelle 2002), and minimizing the pairwise distance between reserves and existing core areas (Alagador and Cerdeira 2007; Snyder et al. 2007).

4.8.2 Reserve connectivity

The degree to which separate reserves are structurally or functionally connected is an important attribute contributing to species persistence. Connectivity can be defined as the degree to which separate habitat reserves are accessible from other patches. Structural connectivity calls for strict adjacency of reserves while functional connectivity is achieved if reserves are within a certain distance of each other. Connectivity is a species-specific and landscape-specific function which is dependent upon mobility characteristics of individual species.

In situations where the landscape is subdivided into contiguous polygons representing candidate sites, structural connectivity can be promoted by selecting sites for protection that are adjacent to each other. Letting u_{ik} be a 0–1 variable for whether or not both sites i and k are selected, the objective is to maximize the number of adjacent pairs of selected sites:

$$\text{Maximize: } \sum_{i \in I} \sum_{k \in A_i, k > i} u_{ik} \quad (4.31)$$

$$u_{ik} \geq x_i + x_k - 1 \quad \text{for all } i \in I, k \in A_i, k > i \quad (4.32)$$

$$\sum_{i \in M_j} x_i \geq y_j \quad \text{for all } j \in J \quad (4.33)$$

$$\sum_{j \in J} y_j \geq R \quad (4.34)$$

$$x_i \in \{0, 1\}, u_{ik} \in \{0, 1\} \quad \text{for all } i, k \in I, j \in J \quad (4.35)$$

where the set A_i represents all sites that are adjacent to site i (Williams et al. 2005a) and Equations 4.33

and 4.34 represent the species coverage constraints. Nalle et al. (2002) employ a similar formulation to the problem of selecting a subset of 4,181 sites in Josephine County, Oregon, USA, to protect examples of 13 habitat types.

Another approach to reserve connectivity uses constructs from graph theory and network optimization (e.g. Williams 1998, 2002; Önal and Briers 2005, 2006; Cerdeira et al. 2005). These models enforce rather than promote structural connectivity of the reserve system, although model variations allow for contiguity gaps (e.g. Williams 2002; Önal and Briers 2005). In network terminology, sites are viewed as network nodes. Network arcs are defined for each pair of adjacent nodes or sites. A contiguous reserve system is a network formed by a set of selected sites linked by arcs between those sites. These constructs can be integrated into an IP model for reserve selection (e.g. Williams 2002; Önal and Briers 2005, 2006).

4.8.3 Reserve shape

The shape of reserves may be important for species survival and many authors advocate creating compact reserves that are nearly circular and have low edge/area ratios. Compact reserves are better for edge-intolerant species that prefer large areas of interior habitat. One approach to forming reserves that meet size and shape requirements is to predefine desirable clusters of sites and include the clusters as decision variables in a site selection model (Williams and ReVelle 1998; Rebain and McDill 2003; Marianov et al. 2008). Another approach, the 'core and buffer' method, involves creating an inner, protected reserve area surrounded by a ring of land managed to buffer the core areas from negative impacts of the surrounding landscape (Williams and ReVelle 1998). A third approach involves selecting sites to minimize a measure of compactness of the resulting reserves, where compactness is the total length of the reserve boundaries (Fischer and Church 2003; Önal and Briers 2003). Total boundary length is the difference between the length of the boundaries of all the selected sites and two times the length of the shared boundaries between the selected sites.

Letting b_i be the length of the boundary of site i and sb_{ik} be the length of the shared boundary between sites i and k , the problem of minimizing total boundary length is:

$$\text{Minimize: } \sum_{i \in I} b_i x_i - 2 \sum_{i \in I} \sum_{k \in A_i, k > i} sb_{ik} u_{ik} \quad (4.36)$$

$$u_{ik} \geq x_i + x_k - 1 \quad \text{for all } i \in I, k \in A_i, k > i \quad (4.37)$$

$$\sum_{i \in M_j} x_i \geq y_j \quad \text{for all } j \in J \quad (4.38)$$

$$\sum_{j \in I} y_j \geq R \quad (4.39)$$

$$x_i, y_j, u_{ik} \in \{0, 1\} \quad \text{for all } i, k \in I, j \in J \quad (4.40)$$

In the objective function (Equation 4.36), the boundary length of the reserve system is calculated by adding the boundary lengths of the selected sites and then subtracting twice the length of the boundaries shared by adjacent sites. Fischer and Church (2003) utilized this model to analyse trade-offs between total area and compactness of reserve systems to protect examples of 55 plant community types in northern California forests.

4.9 Applicability and limitations of IP

We have presented a wide range of reserve selection and design problems that have been successfully formulated and solved using IP methods. These problems involve tens, hundreds, or thousands of 0–1 site selection variables and assume that parameters such as site cost and species presence are fixed and independent of the sites selected for protection. While IP is an efficient way to find optimal solutions to problems that satisfy these assumptions, IP models have limitations. Most importantly, any relevant non-linear relationship between the decision variables, the objective function, and the constraints must be transformed into, or approximated by, linear equations. Non-linear relationships arise, for example, when species presence depends on the spatial arrangement of the set of selected sites (Moilanen 2005b), when species representation is valued with non-linear utility functions (Arponen et al. 2005), or when site protection cost depends on the number and location of the selected sites (Tajibaeva et al. 2008). Although some non-linear relationships can be captured with linear

transformations and approximations (e.g. Hof and Raphael 1997; Arthur et al. 2004), more complex and possibly more realistic non-linear models can be analysed using non-linear optimization methods or heuristics that approximate optimal solutions (e.g. Arponen et al. 2005; Moilanen 2005b; Tajibaeva et al. 2008). Another limitation of IP models is problem size. It is difficult to generalize about the limits of problem size because IP solution time is often a function of the structure of the data matrix. However, generally speaking, problem complexity and solution time tend to increase with the number of binary variables. Thus, a reserve selection or design problem with thousands of potential reserve sites (e.g. remote sensing grid cells at a very fine-grain spatial scale) might be difficult to solve to optimality in a reasonable time frame using IP methods.

4.10 Next steps

While a rich body of research has been published on the application of IP to reserve selection and design, more work is needed in a number of areas. Of growing importance is the need to address species' dynamics and persistence in reserve systems. One approach is to incorporate stochastic, non-linear population dynamics directly into site selection models and use heuristic algorithms to search for solutions that come close to maximizing species persistence (e.g. Moilanen and Cabeza 2002; Haight and Travis 2008). Another approach is to approximate population dynamics with a system of linear equations and formulate linear or integer programming models for site selection (e.g. Bevers et al. 1997; Hof et al. 2002). While optimal solutions can be obtained for the linear models, additional work is needed to evaluate under what conditions linear equations are suitable to represent population dynamics.

While site selection models address problems involving site acquisition and protection, planners also face problems involving the allocation of resources to a wider range of conservation activities, including fire management, invasive species control, and reintroduction of extirpated species. Wilson et al. (2007) describe the decision steps involved in a general conservation investment framework, including definition of the conservation objective, threats to achieving the objective, and possible actions for

each site along with their costs and benefits. When the conservation objective function can be expressed as a linear function of the activities, then LP or IP methods can be used to optimally allocate resources over time (e.g. Bevers et al. 1997).

Another area of current research is the development of models that address species-specific reserve design and habitat needs. Moilanen et al. (2005) develop a heuristic for selecting core areas for multiple species based on species-specific habitat connectivity requirements and conservation weights. More research is needed to develop IP models with species-specific habitat requirements and analyse the trade-offs associated with reserve systems that favour certain species over others.

Finally, more research is needed to benchmark and compare the solution performance of heuristic

methods to exact IP optimization techniques in a variety of reserve selection and design problems. While the use of heuristics may allow an analyst to more rapidly reach a solution, little is known or reported about the quality of solutions generated via heuristics. If heuristics are the preferred option for complex reserve selection and design problems, much more needs to be known about how well these solution methods perform.

Acknowledgement

We are grateful to the reviewers, Atte Moilanen and Helen Regan, for thoughtful suggestions and comments that improved earlier drafts. Research associated with this chapter was supported by the U.S. Forest Service Northern Research Station.