

Application of the Maximal Covering Location Problem to Habitat Reserve Site Selection: A Review

International Regional Science Review

2016, Vol. 39(1) 28-47

© The Author(s) 2014

Reprints and permission:

sagepub.com/journalsPermissions.nav

DOI: 10.1177/0160017614551276

irx.sagepub.com



Stephanie A. Snyder¹ and Robert G. Haight¹

Abstract

The Maximal Covering Location Problem (MCLP) is a classic model from the location science literature which has found wide application. One important application is to a fundamental problem in conservation biology, the Maximum Covering Species Problem (MCSP), which identifies land parcels to protect to maximize the number of species represented in the selected sites. We trace the evolution of the MCSP from the MCLP, review extensions, and offer suggestions for new lines of research related to the MCSP.

Keywords

maximal covering species problem, MCLP, biological conservation, p-median, binary integer programming

Maximal Covering Location Problem (MCLP)

The MCLP is a classic optimization model from the location science literature (Church and ReVelle 1974). Location models, in general, seek to site facilities on

¹ USDA Forest Service, Northern Research Station, St. Paul, MN, USA

Corresponding Author:

Stephanie A. Snyder, USDA Forest Service, Northern Research Station, 1992 Folwell Avenue, St. Paul, MN 55108, USA.

Email: stephaniesnyder@fs.fed.us

the landscape, often on a network of nodes and arcs, to meet public service demands. In the MCLP, the objective is to locate a fixed number of facilities on a network to maximize the population-weighted number of demand nodes that are covered or serviced within a specified distance or time standard.

The MCLP formula has been utilized in a diverse set of application areas, including the optimal location of emergency response facilities, services, and vehicles (Li et al. 2011); communication networks (Lee and Murray 2010); retail stores (Plastria and Vanhaverbeke 2007); and security monitoring sensors (Murray et al. 2007). See Chung (1986) for an early review article on the MCLP. It is the application of the MCLP to a fundamental problem in conservation biology, however, which is the focus of this review article.

Maximal Covering Species Problem (MCSP)

A long-held strategy for preserving biodiversity on the landscape is to develop a system of protected land reserves that contain key species and/or ecological features (Pimm and Lawton 1998). Quantitative decision models have emerged as useful tools for assisting planners and conservation organizations in determining where to assemble sets of sites on the landscape as protected reserves (Moilanen, Wilson, and Possingham 2009).

Twenty-two years after the publication of the MCLP formulation, Church, Stoms, and Davis (1996) and Camm et al. (1996) recognized that the MCLP could be adapted to the question of where to optimally locate sites on a landscape to designate as protected habitat reserves to conserve species of interest. Prior to this, applications of the MCLP had largely focused on the siting of facilities to meet demand for public sector services such as emergency response (Schilling et al. 1980). Thus, the extension of the MCLP to an ecological application was a significant new direction for, and interpretation of, this classic facility location model.

Church, Stoms, and Davis (1996) and Camm et al. (1996) were the first to mathematically state the MCSP, with Church, Stoms, and Davis (1996) the first to report solving an application of it. This extension of the MCLP can be made if species are treated as demand nodes and potential reserve sites as the facility location nodes (see Figure 1). Mathematically, the structure of the MCSP formula is nearly identical to the MCLP. However, the interpretation of the coefficients and decision variables is quite different, as defined subsequently:

$$\text{Maximize } \sum_{i \in I} y_i. \quad (1)$$

Subject to:

$$\sum_{j \in N_i} x_j \geq y_i \quad \forall i \in I. \quad (2)$$

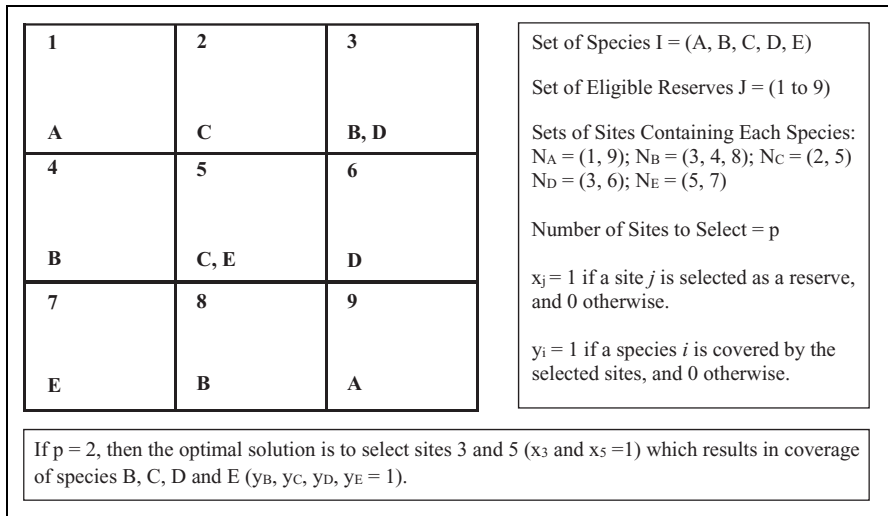


Figure 1. Illustrative example of the maximal covering species problem (equations [1–4]).

$$\sum_{j \in J} x_j = p. \quad (3)$$

$$x_j y_i \in \{0, 1\}. \quad (4)$$

Where,

i, I is the index and set of species to be protected; j, J is the index and set of eligible sites that can be selected as protected reserves; p is the number of sites that can be selected for the reserve system; N_i is the set of sites j that contain species i

$$x_j = \begin{cases} 1 & \text{If a site } j \text{ is selected for the reserve system} \\ 0 & \text{Otherwise} \end{cases}.$$

$$y_i = \begin{cases} 1 & \text{If a species } i \text{ is covered by the selected sites} \\ 0 & \text{Otherwise} \end{cases}.$$

The objective of equation (1) of the MCSP is to maximize the number of unique species represented or “covered” by the selection of a fixed number of eligible reserve sites. Note, one structural difference between the MCLP and the MCSP is that the a_i coefficients of the MCLP (e.g., a_i is the population served at demand node i) are not part of the MCSP formulation. Thus, the MCSP is a special case of the MCLP formula with all $a_i = 1$. Equation (2) specifies that a species is considered covered if at least one of the eligible sites in which it is located is selected for the

reserve system. The concept of the distance or time standard that was used in determining coverage in the MCLP is replaced in the MCSP by the notion that species' coverage is determined by their presence or absence in selected sites. Equation (3) enforces a limit on the number of sites that can be selected for inclusion in the reserve system. This constraint structure can be modified to limit the total area or cost of selected sites to address the situation in which sites vary in these attributes and/or budgets for site selection are limited (Ando et al. 1998; Camm et al. 1996). Equation (4) specifies integer restrictions for the two sets of variables. However, as Church and ReVelle (1974) recognized with the MCLP, the y_i coverage variables need not explicitly be defined as binary variables when solving as a linear program with the branch and bound algorithm for them to solve as such in both the MCLP and MCSP. The structure of the model and the integer restrictions on the x_j variables force the coverage variables, y_i , to be integer as long as they are defined as nonnegative variables with an upper bound of 1. Thus, the integrality restrictions on the y_i coverage variables may be relaxed as follows:

$$0 \leq y_i \leq 1. \quad (5)$$

Relaxing the integrality restriction on the coverage variables reduces the number of integer variables in the problem which may lead to an easier optimization model to solve.

The MCSP, and variants of it, has been utilized by researchers and practitioners from many disciplines who have applied the models to a variety of countries, species, land types, and spatial scales of interest. Many linkages have been found between classic location science models and reserve site selection problems (ReVelle, Williams, and Boland 2002). We highlight the major application areas and extensions of the MCSP, focusing explicitly on reserve site selection models that are based upon a maximal covering construct or typology. Other review articles have been published, which focus on additional aspects of reserve site selection and reserve design models (e.g., Rodrigues and Gaston 2002b; Williams, ReVelle, and Levin 2005; Billionnet 2013).

Minimal Uncovering Species Problem

A mathematically equivalent, but inverse, specification of the MCSP has been formulated (Church, Stoms, and Davis 1996), which is based upon the minimum uncovering location formulation (Church and ReVelle 1974). The objective of this formulation is to minimize the number of species that are left uncovered by the selected set of sites. This formulation relies upon the following variable substitution:

$$u_i = 1 - y_i \quad \forall i \in I. \quad (6)$$

Where

$$u_i = \begin{cases} 1 & \text{If a species } i \text{ is not covered by the selected sites} \\ 0 & \text{Otherwise} \end{cases}.$$

The uncovering version of the formulation is notable in that it offers a potential computational advantage over the conventional maximal covering approach. As was realized by Church, Stoms, and Davis (1996), the “uncovering” variable (u_i) does not have to be explicitly constrained to be a binary variable or even upper-bounded at a value of 1 in order for it to solve with binary values. Instead, the u_i variables can be defined as nonnegative, thereby reducing the number of binary variables in this formulation. This lack of a required upper bound on the u_i variables has been suggested as a way to make the minimal uncovering version of the MCSP more computationally efficient than the MCSP, as was found with the uncovering version of the MCLP (Church and ReVelle 1974).

Church, Stoms, and Davis (1996) and Arthur et al. (1997) both formulated and solved specifications of the MCSP in the uncovering format. While they each reported quick solution times, neither conducted a comparative analysis in terms of solution time and effort to the equivalent MCSP formulation. Snyder, ReVelle, and Haight (2004) formulated and solved versions of the MCSP and minimal uncovering species problem with constraints on the *area* of selected sites rather than the number of sites and failed to find any computational superiority associated with an uncovering formulation. Thus, the computational impact of employing the uncovering version of MCSP has been mixed, although not well tested.

Modified Coverage Constraints

The early MCSP applications focused primarily on maximizing representation of species in selected reserves. However, researchers began to realize that focusing on representation of species might not necessarily lead to long-term survival or persistence of species or the protection of ecological and biological processes that maintain biodiversity (Faith and Walker 1996; Salomon, Ruesink, and DeWreede 2006). In response to this emerging concern, MCSP models have been developed with different objective functions designed to promote persistence of species, including the amount, quality, and spatial arrangement of habitat features that species need in order to persist, as well as representation of biodiversity surrogates (Larsen, Bladt, and Rahbek 2009), phylogenetic diversity (Rodrigues and Gaston 2002a), environmental diversity (Faith and Walker 1996), and assemblage diversity (Araújo, Denham, and Williams 2004).

One approach to promoting species' persistence is to augment the MCSP with backup coverage constraints (ReVelle, Williams, and Boland 2002), an idea derived from the work of Hogan and ReVelle (1986) to develop an MCLP with secondary coverage requirements when siting emergency services. Malcolm and ReVelle (2005) formulate and solve an application of the MCSP with backup constraints. In this model, a species has backup representation in the system of reserves if it is covered by, or represented in, two or more protected sites. Although providing backup coverage does not address species habitat requirements directly, backup coverage guarantees that a species is still covered in the event that a natural or human-caused

catastrophe makes a given site uninhabitable. In addition to the primary coverage decision variable y_i , a variable, z_i , is created as the backup coverage indicator, where z_i takes on a value of 1 if species i is covered by at least two reserve sites, and 0 if covered by one or no reserve sites. The model is as follows:

$$\text{Maximize } \sum_{i \in I} z_i. \quad (7)$$

Subject to:

$$\sum_{j \in N_i} x_j \geq y_i + z_i \quad \forall i \in I. \quad (8)$$

$$y_i \geq z_i \quad \forall i \in I. \quad (9)$$

The objective of equation (7) is to maximize the number of species that are covered more than once. Together, equations (8) and (9) ensure that backup coverage for a species occurs only when at least two sites selected contain species i . Equation (9) ensures that the backup coverage variable is equal to 1 only if the primary coverage variable is also equal to 1 for a species i . A constraint limiting the number of selected sites to p was also enforced, as were the integrality conditions. With this formulation, it is possible to trade-off primary with secondary coverage by specifying a two-objective model to maximize backup coverage $\sum_{i \in I} z_i$ and maximize primary coverage $\sum_{i \in I} y_i$. That is, we may have a choice between representing all species at

least once and representing some species at least twice but leaving other species unrepresented. In their application, Malcolm and ReVelle (2005) found that backup coverage could be achieved for many species with little impact or reduction in the number of species with primary coverage for a specified number of reserve sites. Computational experience with this model was not discussed, however, so it is not clear whether the alteration of the conventional covering constraint influenced solution properties.

Incorporating habitat quality is another means of addressing species' persistence in reserve site selection decisions. Church et al. (2000) were the first to include habitat quality in an MCSP. Augmenting the data for the MCSP, they rated the quality of habitat of each species population present in a site, based on factors such as the size of the population; the amount of resources available for food, cover, and reproduction; and the presence of competitors. This model follows from the weighted benefit coverage model developed by Church and Roberts (1983) as an extension of the MCLP. The model selects sites to protect to maximize species representation as well as the quality of habitat for those species:

$$\sum_{i \in I} \sum_{k \in K} w_i^k y_i^k. \quad (10)$$

Subject to:

$$\sum_{j \in N_i^k} x_j \geq y_i^k \quad \forall i \in I, \forall k \in K. \quad (11)$$

$$\sum_{k \in K} y_i^k \leq 1 \quad \forall i \in I. \quad (12)$$

Where habitat quality is represented by an index k (e.g., $k = 1, 2$, or 3 , representing low, medium, and high quality) and N_i^k is a set of sites that contain species i with level of habitat quality k . Let y_i^k be a 0-1 coverage variable for whether or not a species i is covered by a protected site of habitat quality k . The objective of equation (10) maximizes the weighted sum of species represented in the protected sites, where the weights ($w_i^1 > w_i^2 > w_i^3$) correspond to the levels (low, medium, and high) of habitat quality that are available in the protected sites. Equation (11) enforces the logic of covering: a species is covered with habitat quality k ($y_i^k = 1$) if at least one site with that species and habitat quality is selected for protection. Equation (12) requires the coverage of a species to be counted in at most one quality class. The model also enforces an upper bound of p on the number of sites selected. By applying the model with a range of objective weights, one can generate a series of optimal trade-off alternatives between the total number of species protected and the quality of the coverage obtained for those species for a given number of protected reserves. For modest problem sizes, Church et al. (2000) report solution times ranging from half a second to 10 minutes. Moreover, in a comparative analysis to a basic MCSP formulation, alternate optima to the base MCSP solution were identified in which for a given number of selected reserves, total representation was the same, but higher quality site representation was obtained. Thus, this formulation extends the base MCSP, without significant computational burden, to allow for a more informed selection of reserve sites which account not only for species' representation, but also the quality of the habitat for each species as a means of promoting species persistence.

More sophisticated treatment of habitat quality has been developed in further extensions of the MCSP. For example, Burns, Tóth, and Haight (2013) extend the use of habitat information in the MCSP by developing a reserve site selection model to maximize the number of populations covered for a given species, where a population is considered covered if minimum areas of desired habitat features are present within desired spatial conditions in the protected sites. In spite of using a coverage constraint that was significantly more complex than the basic MCSP coverage constraint, the authors report trivial solution time for a realistic-sized site selection problem. Thus, in general, modifications to the MCSP to address issues of species' persistence have tended to result in coverage constraints that are more complex and with nonbinary coefficients which could increase computational burden. However, for at least some of these models that employ modified coverage constraints, solution times have not been an issue.

Incorporating Spatial Design Features with Species' Coverage

One shortcoming of the MCSP formulation is that it does not consider the spatial distribution of selected sites. As a consequence, solutions from the MCSP may consist of a set of scattered reserves with little spatial coherence, which may increase the difficulty and expense of reserve management, as well as do little to support persistence of species (Cabeza and Moilanen 2001). Further, scattered reserves are particularly troublesome when they are surrounded by a matrix of land uses and cover types that adversely impact species' persistence, viability, and movement. For example, when a reserve system consists of disjoint areas of protected habitat, the distance between reserves may influence species' mobility and viability.

Reserve design attributes such as reserve proximity, connectivity, and shape can be formulated as objectives and/or constraints in integer programming models for site selection (Williams, ReVelle, and Levin 2005; Billionnet 2013). When combined with an objective for species coverage, the trade-offs between the spatial and coverage objectives can be explored and synergies or conflict between achievements of these two types of objectives can be identified. However, for the most part, reserve design and reserve site selection model development have been largely separate lines of inquiry with the reserve design models tending to employ model structures and constructs without a coverage element or connection to the MCLP. Exceptions to this can be found in Rothley (1999) who developed a three-objective model to maximize coverage of rare plant species, while also maximizing the total area of the selected reserve system and maximizing a measure of connectedness of the selected reserve sites. This model was further refined in Rothley (2006) to maximize species' coverage while maximizing a design objective that combined patch size and connectivity subject to MCSP constraints. We suggest that research is still needed on ways to combine coverage objectives with reserve design features within site selection models.

Multi-objective Maximum Covering Species Problem

Multi-objective extensions of the MCSP have been specified and solved to address a variety of reserve site selection and design goals in addition to species' representation such as spatial attributes (Rothley 1999), resource limitations and opportunity costs (Church, Stoms, and Davis 1996; Polasky, Camm, and Garber-Yonts 2001; Murdoch et al. 2007), habitat quality (Church et al. 2000), species-specific habitat requirements (Malcolm and ReVelle 2005), species' richness and rarity metrics (Memtsas 2003), and community proximity to the reserve system (Ruliffson et al. 2003; Önal and Yanprechaset 2007).

Church, Stoms, and Davis (1996) were the first to report solving a multi-objective statement of the MCSP. That is, they solved the base MCSP formulation iteratively for increasing values of p (number of reserve sites), trading-off species'

representation with the number of reserve sites selected utilizing the constraint method of solution (Haimes, Lasdon, and Wismer 1971). The limit on the number of selected sites was used as a proxy budget constraint.

An alternate way to specify a site, area, or budget-constrained MCSP is to cast this resource limitation as a second objective and solve via the multi-objective weighting method (Zadeh 1963). Snyder, ReVelle, and Haight (2004) solved an example of this type of formulation, where the objective was:

$$\text{Maximize} \left(w_1 \times \left(\sum_{i \in I} y_i \right) \right) - \left(w_2 \times \left(\sum_{j \in J} a_j x_j \right) \right). \quad (13)$$

The objective of equation (13) is to maximize the number of land types covered by the selected set of sites (y_i) while minimizing the total area of reserve system ($a_j x_j$). With this formulation and the multi-objective weighting solution method, the objective function weights (w_1 , w_2) are systematically varied, and the problem resolved to produce a trade-off curve between the number of land types covered and the total area of the selected sites. While the constraint-based approach (Snyder, Tyrrell, and Haight 1999) and the multi-objective weighting approach (Snyder, ReVelle, and Haight 2004) to a bi-criteria specification of the MCSP are both potential means of generating an estimate of the noninferior set of solutions, there are some potential computational issues to consider.

The weighting method may offer computational efficiencies (ReVelle 1993; Snyder, ReVelle, and Haight 2004) in that it allows constraints that are not likely to be amenable to producing integer solutions (e.g., budget or area constraints) to be placed in the objective function, thereby promoting the “integer-friendly” structure (ReVelle 1993) of the constraint set. However, an issue known as gap points can occur when solving a bi-criteria integer optimization model via the weighting method (Cohon 1978). Gap points are noninferior solutions that cannot be identified via the weighting method because the surface of the trade-off curve may not be convex or concave due to the integrality of the integer decision variables. Thus, using the weighting method to generate an estimate of the noninferior set and then the constraint method to hone in on areas on the trade-off curve of greatest interest may be a useful strategy when solving multi-objective instances of the MCSP. Tóth, McDill, and Rebain (2006) describe additional means of solving bi-criteria integer problems that may prove useful for modelers trying to solve increasingly complex specifications of multi-objective MCSP problems.

Probabilistic Extensions—Maximal Expected Species Covering Model

The MCSP assumes that species’ presence in each site is known with certainty; however, in many cases, this information is not known with certainty and is expressed as a probability of occurrence. In these situations, decision makers may be concerned

about the likelihood that species are represented in the selected reserves sites and want to maximize the number of species covered with a minimum reliability. That is, a species is considered covered only if its probability of presence in the selected sites reaches or exceeds a specified threshold, for example, 95 percent (Haight, ReVelle, and Snyder 2000). This threshold approach utilizes the “safe minimum standard” approach to conservation (Bishop 1978) in which a threshold can be identified as the minimum acceptable probability of success with respect to the conservation objective.

The threshold approach can be viewed as a maximal covering problem because continuous probabilities of species presence in the selected sites are converted to dichotomous 0, 1 variables and then summed. Let p_{ij} be the probability that species i is present in site j where $0 \leq p_{ij} < 1$ and p_{ij} is independent of the probability of occurrence in neighboring sites. Then, the probability that species i is not present in the sites selected for protection is a product of the absence probabilities over all sites, $\prod_{j \in J} (1 - p_{ij})^{x_j}$, where x_j is the 0-1 decision variable for whether or not site j is selected for protection. Defining y_i as a zero-one variable for whether or not species i is present in the selected sites with probability α_i , we formulate the model as:

$$\text{Maximize } \sum_{i \in I} y_i. \quad (14)$$

Subject to:

$$\prod_{j \in J} (1 - p_{ij})^{x_j} \leq (1 - \alpha_i)^{y_i} \quad \forall i \in I. \quad (15)$$

The objective of the equation (14) maximizes the number of species represented with α -reliability. Equation (15) says that species i is counted as being represented in the reserve ($y_i = 1$) only if the probability of system-wide absence is less than $(1 - \alpha_i)$. Note that if $y_i = 0$, then the right-hand side of equation (15) is one, and the expression no longer constrains the model. This constraint can be linearized by taking logarithms and producing a covering constraint:

$$y_i \leq \sum_{j \in J} \frac{x_j \ln(1 - p_{ij})}{\ln(1 - \alpha_i)}. \quad (16)$$

This model also enforces a budget constraint and integrality requirements. This probabilistic species-covering problem is a generalization of the maximum availability location problem (ReVelle and Hogan 1989) in which a constant busy fraction replaces $(1 - p_{ij})$ in equation (15) and the problem is to maximize the number of demand nodes covered with a specified level of reliability. An alternative formulation is to maximize the expected number of species covered (Polasky et al. 2000), where species coverage is a continuous variable representing the probability that the species is present in at least one of the selected sites. Then, the sum of the coverage

probabilities equals the expected number of species covered. While this objective is nonlinear and cannot be converted to an equivalent linear integer program, Camm et al. (2002) and Billionnet (2011) develop linear approximations that can be solved using standard mixed-integer programming software.

Dynamic Reserve Selection with Uncertain Site Availability

The MCSP assumes that site selections are made all at once and protection takes place rapidly before site degradation or loss. In practice, however, decisions take place sequentially as funds and political support become available. Further, land availability is dynamic: sites currently available may be developed if protection is delayed, or sites not immediately available may be open for protection later.

To address the issue of site availability, Snyder, Haight, and ReVelle (2004) develop a two-period maximal covering model for sequential site selection in which uncertainty about future site availability is represented with a set of probabilistic scenarios. Although scenario optimization is commonly used to model uncertainty in the parameters of facility location models (Owen and Daskin 1998), we believe Snyder, Haight, and ReVelle's (2004) application is the first use of scenario optimization with the maximal covering problem. The two-period problem maximizes the expected number of species covered at the end of the second period subject to upper bounds p_1 and p_2 , on the number of reserve sites selected in each period. 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 site development scenarios, S , indexed by s . Each scenario s is one possible development outcome represented by a vector of 0-1 parameters d_{js} for all sites $j \in J$ identifying which sites are undeveloped and available for protection ($d_{js} = 1$) and which sites are not ($d_{js} = 0$) in the second period. Associated with each scenario s is a probability of occurrence, p_s . The model has two sets of 0-1 site-selection variables. The first set includes the protection choices x_{1j} for all sites $j \in J$ 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 x_{2js} all sites $j \in J$ in the second period under each development scenario s . The 0-1 variable y_{is} counts whether species i is represented in protected sites in scenario s , where a species is represented if it is present in at least one site selected for protection. The model is formulated as follows:

$$\text{Maximize } \sum_{s \in S} \left(p_s \sum_{i \in I} y_{is} \right). \quad (17)$$

Subject to:

$$x_{1j} + x_{2js} \leq 1 \quad \forall j \in J, \forall s \in S. \quad (18)$$

$$x_{2js} \leq d_{js} \quad \forall j \in J, \forall s \in S. \quad (19)$$

$$\sum_{j \in J} x_{1j} + \leq p_1. \quad (20)$$

$$\sum_{j \in J} x_{2js} \leq p_2 \quad \forall s \in S. \quad (21)$$

$$\sum_{j \in N_i} (x_{1j} + x_{2js}) \geq y_{is} \quad \forall i \in I, \forall s \in S. \quad (22)$$

$$x_{1j}, x_{2js}, y_{is} \in \{0, 1\}. \quad (23)$$

The objective of equation (17) maximizes the expected number of species represented by the set of selected sites in period one and by the selected sites in each scenario in period two. Equation (18) specifies that site j can at most be selected for protection in either period one or period two, but not both. Equation (19) specifies that site j can only be selected for protection in period 2 in scenario s if site j is undeveloped and available for protection in that scenario. Equations (20) and (21) limit the total number of sites selected for protection in periods 1 and 2 under each scenario to specified upper bounds. Equation (22) defines the conditions under which species i is represented. This constraint stipulates that in order for a species to be represented in scenario s , at least one site that contains that species must be selected for protection either in the first period or in scenario s in the second period. The result is a set of sites for protection in period 1 and a set of sites for protection in period 2 under each development scenario. This two-period problem is readily solvable using integer programming methods (Snyder, Haight, and ReVelle 2004) and provides information about how uncertain site availability affects current site selection decisions (Haight, Snyder, and ReVelle 2005).

Another way to treat uncertainty about site availability and loss is to reformulate the MCSP to minimize expected biodiversity loss (O'Hanley, Church, and Gilles 2007). Species are assumed to survive not only in protected habitat but also unprotected habitat provided the habitat is not destroyed by development. Because development of unprotected sites is uncertain, species survival outside of the protected sites is also uncertain and the objective is to select sites for protection to minimize the expected number of species lost to development in the unprotected sites. Let q_j be the probability of destruction of site j over some future period after site selection. Then, the probability β_i that an unprotected species i will be extirpated is equal to the joint probability that every site $j \in N_i$ is destroyed, or $\beta_i = \prod_{j \in N_i} q_j$. Let the

variable $y_i = 0$ if species i is present in one or more of the protected sites and $y_i = 1$ one if species i is absent from the protected sites. Then, the problem is to minimize the expected number of species in unprotected sites that are subsequently lost to development (the expected coverage loss [ECL] problem):

$$\text{Minimize } \sum_{i \in I} \beta_i y_i. \quad (24)$$

Subject to:

$$\sum_{j \in N_i} x_j \geq 1 - y_i \quad \forall i \in I. \quad (25)$$

Equation (25) defines the conditions under which species i is present in the protected sites. This constraint stipulates that for $y_i = 0$, at least one site that contains that species must be selected for protection (i.e., the right-hand side of equation [25] is less than or equal to 0). Note that, if $y_i = 0$, then the value of expected loss for species i in the objective function drops to 0, meaning species' protection removes any risk of extirpation. The model also enforces a budget constraint.

The ECL problem represents an alternative formulation of the MCSP (i.e., in its minimization form as shown in Church, Stoms, and Davis 1996) but with an additional set of weights β_i corresponding to the probability that species i will be extirpated if left unprotected. As noted by O'Hanley, Church, and Gilles (2007), the MCSP could be viewed as a generalization of the ECL formula in which all of the β_i are assumed equal to 1. Given this, they suggest that solution of the ECL is not likely to be any harder to achieve than the conventional MCSP. However, the addition of the budget constraint over a constraint on the number of sites to protect could negatively influence solution properties.

Heuristic Approaches to Solving the MCSP

In recent years, increasing emphasis has been placed on developing and solving reserve site selection models for large problem applications that have broad spatial extent, large numbers of species and potential reserves, and the ability to interface with geographic information system for map displays of results (Bladt et al. 2009; Larsen, Londoño-Murcia, and Turner 2011). Such problem instances can result in very large combinatorial problems when formulated as integer programming problems. Since the MCLP and MCSP are in a class of models known as Nondeterministic Polynomial-time hard (Megiddo, Zeman, and Hakimi 1983), solution by exact optimization methods can be problematic. Specifically, a proven bound on the computational effort required to solve every problem instance to optimality cannot be expressed as a polynomial of the problem characteristics. Thus, it is possible that there could be instances of the MCSP, as with the MCLP, which cannot be solved in reasonable time frames using exact optimization methods. Given this, another trend that has emerged in the field is the development of heuristic solution

techniques to solve landscape-level specifications of the MCSP (e.g., Csuti et al. 1997). A variety of heuristic techniques have been developed for solution of MCSP formulations, including: Tabu search (Kincaid, Easterling, and Jeske 2008; Ciarleglio, Barnes, and Sarkar 2009), interchange heuristics (Rosing, ReVelle, and Williams 2002), greedy adding heuristics (Clemens, ReVelle, and Williams 1999; Önal 2003; Vanderkam, Wiersma, and King 2007), and simulated annealing (Ball, Possingham, and Watts 2009).

Moreover, Church and ReVelle (1976) recognized that the MCLP could be stated as a special case of the related *p*-median problem (Hakimi 1964), in which the distance between each demand node and its closest facility is replaced by 0 if it is less than the maximal service distance, and 1 otherwise. Given this relationship, specialized *p*-median algorithms and heuristics can be used to solve the MCLP as well as the MCSP (Church and ReVelle 1976; Eilon and Galvão 1978; Rosing, ReVelle, and Schilling 1999; Teitz and Bart 1968). However, research on how well *p*-median heuristics perform on MCSP formulations and data sets is limited (e.g., Gerrard et al. 1997; Woodhouse et al. 2000; Araújo, Densham, and Williams 2004; Hortal, Araújo, and Lobo 2009). More exploration of this topic is needed, particularly given the trend toward solution of large, complex instances of reserve site selection problems via heuristics.

Observations and Next Steps

The MCSP has been applied and modified in numerous ways since its inception: eligible reserve sites have been differentiated by size, cost, and quality; budget and area restrictions have been added; and objectives beyond species' representation such as cost, public accessibility, and spatial reserve characteristics have been optimized. However, research on reserve site selection models is still needed in a number of areas, and we suggest that models with coverage constructs are still useful ways to approach this important conservation biology problem. For one, new theories are emerging from the field of conservation biology about the biodiversity elements that are important to consider when making land conservation decisions, such as the protection of phylogenetic diversity (Kukkala and Moilanen 2013). Continuing to evolve and refine MCSP-type formulations to address these new conservation paradigms will be an important direction for decision modelers. Moreover, we suggest that greater focus is needed on models and modeling approaches that combine species coverage objectives with spatial design objectives. Neither of these general objectives is likely sufficient to adequately address long-term habitat needs of species of interest, yet they have typically been addressed as separate fields of inquiry in the decision modeling literature.

Another new direction will be to explore whether MCSP-type models can be used to inform habitat conservation or restoration decisions in the face of disturbance regimes such as climate change, wildfire, and pest and disease outbreaks. A looming concern for reserve managers and conservation planners is whether existing reserves

will continue to provide the protection for species they were designed to in the future under different climate projections. Moreover, decisions about where to locate future protected habitat reserves will need to recognize that the matrix of habitat location and species' location is not likely to remain stable over time.

Beyond new application areas of the MCSP, we also suggest that more research is needed to explore whether formulation "efficiencies" that have previously been suggested in the literature for both the MCLP and the MCSP actually lead to better solution times for different problem instances of the MCSP. For example, many MCSP articles we reviewed are still explicitly requiring the coverage variables to be binary variables rather than nonnegative variables with an upper bound of 1 as recommended by Church, Stoms, and Davis (1996). More comparative research is needed with different instances and data sets of the MCSP to determine whether the relaxation of the integrality conditions on the coverage variables lead to the solution efficiencies found when applied to MCLP problems. We are aware of only one study that included a comparative analysis of solution time and effort when coverage variables in an MCSP formulation were declared binary and when the integrality restrictions were relaxed (Snyder, ReVelle, and Haight 2004). In this research, solution times were found to be quicker when the coverage variables were declared binary. More comparative work of this nature is needed to determine whether relaxed integrality restrictions are useful in the context of MCSP, particularly with models that include constraints on area or budget rather than number of sites. Further, in spite of suggestions that formulation of the MCLP in its uncovering version offers computational advantages (Church and ReVelle 1974), little use of this modeling structure in the reserve site selection literature has been made. More comparative research between the covering and uncovering formulations when applied to MCSP applications is needed to establish whether computational efficiencies encountered with the uncovering version of MCLP problem also apply when utilized to solve various specifications of the MCSP. Camm et al. (1996) outline suggestions for strategic preprocessing of species' presence-absence data sets in order to reduce problem size, similar to the row and column "reduction" techniques suggested by Toregas and ReVelle (1973) for the location set-covering problem. Few of the manuscripts we reviewed make any mention of the utilization of such data reduction techniques.

In addition to these structural formulation efficiencies, optimization software solver settings can be modified to capitalize on solution features of the MCLP and MCSP. Önal (2003) recognized that the optimality gap in the branch and bound solver for an MCSP application can be set to a value just less than 1 and still guarantee optimal solutions. This is possible because the objective function of the MCSP can only take on integer values (e.g., number of species covered). Thus, no improvement in the objective function value is possible when the absolute optimality gap falls below 1. This allows for an earlier stopping criteria and potentially quicker solution times when solving instances of the MCSP in which the number of species is being optimized. Again, we find little documented evidence in articles we reviewed that

this computational feature of the MCSP is being capitalized on when solved via branch and bound methods.

Additional research is needed to evaluate how the structure of MCSP data sets may influence solution time and effort when solved via different solution methods, and whether specialized algorithms or modeling constructs can be developed to take advantage of such data structure. MCLP problem sets tend to have a data structure with a spatial logic; for example, a data set with transitivity associated with location of potential servers and demand points on a geographical network (Rosing, ReVelle, and Williams 2002). However, this data structure may be missing from the spatial arrangement of species' presence/absence data and potential reserve site location utilized in MCSP problem specifications (Rosing, ReVelle, and Williams 2002). This may make instances of the MCSP more difficult to solve than similarly sized instances of the MCLP and could mean that specialized MCLP algorithms (Church and ReVelle 1974; Galvão and ReVelle 1996; ReVelle, Scholssberg, and Williams 2008) may not be as effective when applied to the data of the MCSP.

Finally, we would urge those interested in continuing to apply and modify the MCSP to explore the rich literature on the MCLP which has been developing over the past forty years for tips on formulations, model structure, and algorithmic enhancements that may be brought to bear anew in reserve site selection and design applications. Quantitative decisions models like the MCLP and MCSP can provide managers valuable insights into complex resource allocation problems. Continuing to evolve the MCSP to allow land managers to address the complexities of real-world habitat conservation decision making is an important effort that will support informed, effective, and efficient land protection decisions.

Acknowledgment

We are grateful to Richard Church, Alan Murray, and the reviewers for providing valuable feedback and suggestions on this article.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Funding was provided by the USDA Forest Service, Northern Research Station.

References

- Ando, A., J. D. Camm, S. Polasky, and A. Solow. 1998. "Species Distributions, Land Values and Efficient Conservation." *Science* 279:2126–28.
- Araújo, M. B., P. J. Densham, and P. H. Williams. 2004. "Representing Species in Reserves from Patterns of Assemblage Diversity." *Journal of Biogeography* 31:1037–50.

- Arthur, J. L., M. Hachey, K. Sahr, M. Huso, and A. R. Kiester. 1997. "Finding All Optimal Solutions to the Reserve Site Selection Problem: Formulation and Computational Analysis." *Environmental and Ecological Statistics* 4:153–65.
- Ball, I. R., H. P. Possingham, and M. Watts. 2009. "Marxan and Relatives: Software for Spatial Conservation Prioritisation." In *Spatial Conservation Prioritisation: Quantitative Methods and Computational Tools*, Chapter 14, edited by A. Moilanen, K. A. Wilson, and H. P. Possingham, 185–95. Oxford, UK: Oxford University Press.
- Billionnet, A. 2011. "Solving the Probabilistic Reserve Selection Problem." *Ecological Modelling* 222:546–54.
- Billionnet, A. 2013. "Mathematical Optimization Ideas for Biodiversity Conservation." *European Journal of Operational Research* 231:514–34.
- Bishop, R. C. 1978. "Endangered Species and Uncertainty: The Economics of the Safe Minimum Standard." *American Journal of Agricultural Economics* 60:10–18.
- Bladt, J., N. Strange, J. Abildtrup, J.-C. Svenning, and F. Skov. 2009. "Conservation Efficiency of Geopolitical Coordination in the EU." *Journal for Nature Conservation* 17: 72–86.
- Burns, E. S., S. F. Tóth, and R. G. Haight. 2013. "A Modeling Framework for Life History-based Conservation Planning." *Biological Conservation* 158:14–25.
- Cabeza, M., and A. Moilanen. 2001. "Design of Reserve Networks and the Persistence of Biodiversity." *Trends in Ecology and Evolution* 16:242–48.
- Camm, J. D., S. K. Norman, S. Polasky, and A. R. Solow. 2002. "Nature Reserve Selection to Maximize Expected Species Covered." *Operations Research* 50:946–55.
- Camm, J. D., S. Polasky, A. Solow, and B. Csuti. 1996. "A Note on Optimal Algorithms for Reserve Site Selection." *Biological Conservation* 78:353–55.
- Chung, C.-H. 1986. "Applications of the Maximal Covering Location Planning (M.C.L.P.) Model." *The Journal of the Operational Research Society* 37:735–46.
- Church, R. L., R. Gerrard, A. Hollander, and D. Stoms. 2000. "Understanding the Trade-offs between Site Quality and Species Presence in Reserve Site Selection." *Forest Science* 46: 157–67.
- Church, R., and C. ReVelle. 1974. "The Maximal Covering Location Problem." *Papers of the Regional Science Association* 32:101–18.
- Church, R. L., and C. S. ReVelle. 1976. "Theoretical and Computational Links Between P-median, Location Set-covering, and Maximal Covering Location Problem." *Geographical Analysis* 8:406–15.
- Church, R., and K. Roberts. 1983. "Generalized Coverage Models and Public Facility Location." *Papers of the Regional Science Association* 53:117–35.
- Church, R. L., D. M. Stoms, and F. W. Davis. 1996. "Reserve Selection as a Maximal Covering Location Problem." *Biological Conservation* 76:105–12.
- Ciarleglio, M., J. W. Barnes, and S. Sarkar. 2009. "ConsNet: New Software for the Selection of Conservation Area Networks with Spatial and Multi-criteria Analyses." *Ecography* 32: 205–9.
- Clemens, M. A., C. S. ReVelle, and J. C. Williams. 1999. "Reserve Design for Species Preservation." *European Journal of Operational Research* 112:273–83.

- Cohon, J. 1978. *Multiobjective Programming and Planning*. New York: Academic Press.
- Csuti, B., S. Polasky, P. H. Williams, R. L. Pressey, J. D. Camm, M. Kershaw, A. R. Kiester, B. Downs, R. Hamilton, M. Huso, and K. Sahr. 1997. "A Comparison of Reserve Selection Algorithms Using Data on Terrestrial Vertebrates in Oregon." *Biological Conservation* 80:83–97.
- Eilon, S., and R. D. Galvão. 1978. "Single and Double Vertex Substitution in Heuristic Procedures for the P-median Problem." *Management Science* 24:1763–66.
- Faith, D. P., and P. A. Walker. 1996. "Environmental Diversity: On the Best-possible Use of Surrogate Data for Assessing the Relative Biodiversity of Sets of Areas." *Biodiversity and Conservation* 5:399–415.
- Galvão, R. D., and C. S. ReVelle. 1996. "A Lagrangian Heuristic for the Maximal Covering Location Problem." *European Journal of Operational Research* 88:114–23.
- Gerrard, R., R. Church, D. M. Stoms, and F. W. Davis. 1997. "Selecting Conservation Reserves Using Species-covering Models: Adapting the ARC/INFO GIS." *Transactions in GIS* 2:45–60.
- Haight, R. G., C. S. ReVelle, and S. A. Snyder. 2000. "An Integer Optimization Approach to a Probabilistic Reserve Site Selection Problem." *Operations Research* 48:697–708.
- Haight, R. G., S. A. Snyder, and C. S. ReVelle. 2005. "Metropolitan Open-space Protection with Uncertain Site Availability." *Conservation Biology* 19:327–37.
- Haimes, Y. Y., L. S. Lasdon, and D. A. Wismer. 1971. "On a Bicriterion Formulation of the Problems of Integrated System Identification and System Optimization." *IEEE Transactions on Systems, Man and Cybernetics* 1:296–97.
- Hakimi, S. 1964. "Optimum Location of Switching Centers and the Absolute Centers and Medians of a Graph." *Operations Research* 12:450–59.
- Hogan, K., and C. S. ReVelle. 1986. "Concepts and Applications of Backup Coverage." *Management Science* 32:1434–44.
- Hortal, J., M. B. Araújo, and J. M. Lobo. 2009. "Testing the Effectiveness of Discrete and Continuous Environmental Diversity as a Surrogate for Species Diversity." *Ecological Indicators* 9:138–49.
- Kincaid, R. K., C. Easterling, and M. Jeske. 2008. "Computational Experiments with Heuristics for Two Nature Reserve Site Selection Problems." *Computers and Operations Research* 35:499–512.
- Kukkala, A. S., and A. Moilanen. 2013. "Core Concepts of Spatial Prioritisation in Systematic Conservation Planning." *Biological Reviews* 88:443–64.
- Larsen, F. W., J. Bladt, and C. Rahbek. 2009. "Indicator Taxa Revisited: Useful for Conservation Planning." *Diversity and Distributions* 15:70–79.
- Larsen, F. W., M. C. Londoño-Murcia, and W. R. Turner. 2011. "Global Priorities for Conservation of Threatened Species, Carbon Storage, and Freshwater Services: Scope for Synergy?" *Conservation Letters* 4:355–63.
- Lee, G., and A. T. Murray. 2010. "Maximal Covering with Network Survivability Requirements in Wireless Mesh Networks." *Computers, Environment and Urban Systems* 34:49–57.

- Li, X., Z. Zhao, X. Zhu, and T. Wyatt. 2011. "Covering Models and Optimization Techniques for Emergency Response Facility Location and Planning: A Review." *Mathematical Methods of Operations Research* 74:281–310.
- Malcolm, S. A., and C. S. ReVelle. 2005. "Models for Preserving Species Diversity with Backup Coverage." *Environmental Modeling and Assessment* 10:99–105.
- Megiddo, M., E. Zeman, and S. L. Hakimi. 1983. "The Maximum Coverage Location Problem." *SIAM Journal on Algebraic and Discrete Methods* 4:253–61.
- Memtsas, D. P. 2003. "Multiobjective Programming Methods in the Reserve Selection Problem." *European Journal of Operational Research* 150:640–52.
- Moilanen, A., K. A. Wilson, and H. P. Possingham, eds. 2009. *Spatial Conservation Prioritization: Quantitative Methods and Computational Tools*. New York: Oxford University Press, p. 304.
- Murdoch, W., S. Polasky, K. A. Wilson, H. P. Possingham, P. Kareiva, and R. Shaw. 2007. "Maximizing Return on Investment in Conservation." *Biological Conservation* 139: 375–88.
- Murray, A. T., K. Kamyoung, J. W. Davis, R. Machiraju, and R. Parent. 2007. "Coverage Optimization to Support Security Monitoring." *Computers and Urban Systems* 31:133–47.
- O'Hanley, J. R., R. L. Church, and J. K. Gilles. 2007. "Locating and Protecting Critical Reserve Sites to Minimize Expected and Worst-case Losses." *Biological Conservation* 134:130–41.
- Önal, H. 2003. "First-best, Second-best and Heuristic Solutions in Conservation Reserve Site Selection." *Biological Conservation* 115:55–62.
- Önal, H., and P. Yanprechaset. 2007. "Site Accessibility and Prioritization of Nature Reserves." *Ecological Economics* 60:763–73.
- Owen, S. H., and M. S. Daskin. 1998. "Strategic Facility Location: A Review." *European Journal of Operational Research* 111:423–47.
- Pimm, S. L., and J. H. Lawton. 1998. "Planning for Biodiversity." *Science* 279:2068–69.
- Plastria, F., and L. Vanhaverbeke. 2007. "Aggregation without Loss of Optimality in Competitive Location Models." *Networks and Spatial Economics* 7:3–18.
- Polasky, S., J. D. Camm, and B. Garber-Yonts. 2001. "Selecting Biological Reserves Cost Effectively: An Application to Terrestrial Vertebrate Conservation in Oregon." *Land Economics* 77:68–78.
- Polasky, S., J. D. Camm, A. R. Solow, B. Csuti, D. White, and R. Ding. 2000. "Choosing Reserve Networks with Incomplete Species Information." *Biological Conservation* 94:1–10.
- ReVelle, C. S. 1993. "Facility Siting and Integer-friendly Programming." *European Journal of Operational Research* 65:147–58.
- ReVelle, C. S., and K. Hogan. 1989. "The Maximum Availability Location Problem." *Transportation Science* 23:192–200.
- ReVelle, C. S., M. Scholssberg, and J. C. Williams. 2008. "Solving the Maximal Covering Location Problem with Heuristic Concentration." *Computers and Operations Research* 35:427–35.
- ReVelle, C. S., J. C. Williams, and J. J. Boland. 2002. "Counter-part Models in Facility Location Science and Reserve Selection Science." *Environmental Modeling and Assessment* 7:71–80.

- Rodrigues, A. S. L., and K. J. Gaston. 2002a. "Maximising Phylogenetic Diversity in the Selection of Networks of Conservation Areas." *Biological Conservation* 105:103–11.
- Rodrigues, A. S. L., and K. J. Gaston. 2002b. "Optimisation in Reserve Selection Procedures – Why Not?" *Biological Conservation* 107:123–29.
- Rosing, K. E., C. S. ReVelle, and D. A. Schilling. 1999. "A Gamma Heuristic for the P-median Problem." *European Journal of Operational Research* 117:522–32.
- Rosing, K. E., C. S. ReVelle, and J. C. Williams. 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* 9:741–50.
- Rothley, K. D. 2006. "Finding the Tradeoffs between the Reserve Design and Representation." *Environmental Management* 38:327–37.
- Ruliffson, J. A., R. G. Haight, P. H. Gobster, and F. R. Homans. 2003. "Metropolitan Natural Area Protection to Maximize Public Access and Species Representation." *Environmental Science and Policy* 6:291–99.
- Salomon, A. K., J. L. Ruesink, and R. E. DeWreede. 2006. "Population Viability, Ecological Processes and Biodiversity: Valuing Sites for Reserve Selection." *Biological Conservation* 128:79–92.
- Schilling, D. A., C. ReVelle, J. Cohon, and D. J. Elzinga. 1980. "Some Models for Fire Protection Locational Decisions." *European Journal of Operational Research* 5:1–7.
- Snyder, S. A., R. G. Haight, and C. S. ReVelle. 2004. "A Scenario Optimization Model for Dynamic Reserve Site Selection." *Environmental Modeling and Assessment* 9:179–87.
- Snyder, S. A., C. S. ReVelle, and R. G. Haight. 2004. "One and Two-objective Approaches to an Area-constrained Habitat Reserve Site Selection Problem." *Biological Conservation* 119:565–74.
- Snyder, S. A., L. E. Tyrrell, and R. G. Haight. 1999. "An Optimization Approach to Selecting Research Natural Areas in National Forests." *Forest Science* 45:458–69.
- Teitz, M. B., and P. Bart. 1968. "Heuristic Methods for Estimating Generalized Vertex Median of a Weighted Graph." *Operations Research* 16:955–61.
- Toregas, C., and C. ReVelle. 1973. "Binary Logic Solutions to a Class of Location Problem." *Geographical Analysis* 5:145–55.
- Tóth, S. F., M. E. McDill, and S. Rebain. 2006. "Finding the Efficient Frontier of a Bi-criteria, Spatially Explicit, Harvest Scheduling Model." *Forest Science* 52:93–107.
- Vanderkam, R. P. D., Y. F. Wiersma, and D. J. King. 2007. "Heuristic Algorithms Vs. Linear Programs for Designing Efficient Conservation Reserve Networks: Evolution of Solution Optimality and Processing Time." *Biological Conservation* 137:349–58.
- Williams, J. C., C. S. ReVelle, and S. A. Levin. 2005. "Spatial Attributes and Reserve Design Models: A Review." *Environmental Modeling and Assessment* 10:163–81.
- Woodhouse, S., A. Lovett, P. Dolman, and R. Fuller. 2000. "Using a GIS to Select Priority Areas for Conservation." *Computers, Environment and Urban Systems* 24:79–93.
- Zadeh, L. 1963. "Optimality and Non-scalar Valued Performance Criteria." *IEEE Transactions Automatic Control* AC-8:59.