

Planning for Biodiversity Conservation Using Stochastic Programming

Trevon Fuller,¹ David P. Morton,² and Sahotra Sarkar¹

¹ Section of Integrative Biology, University of Texas at Austin, 1 University Station, C0390, Austin, Texas 78712–1180, USA; tfuller@mail.utexas.edu, sarkar@mail.utexas.edu

² Graduate Program in Operations Research, University of Texas at Austin, 1 University Station, C2200, Austin, Texas 78712–0292, USA; morton@mail.utexas.edu

Summary. Rapid species extinctions and the loss of other biodiversity features worldwide have prompted the development of a systematic planning framework for the conservation of biodiversity. Limited resources (~ 40 million USD annually) are available for conservation, particularly in the developing countries that contain many of the world’s hotspots of species diversity. Thus, conservation planning problems are often represented as mathematical programs in which the objective is to select sites to serve as conservation areas so that the cost of the plan is as small as possible and adequate habitat is protected for each species. Here, we generalize this approach to allow for uncertainty in the planning process. In particular, we assume that the species to be protected disperse after the conservation areas are established and that planners cannot anticipate with certainty the species’ future locations when selecting the conservation areas. This uncertainty is modeled by including random variables in the mathematical program. We illustrate the approach by designing a network of conservation areas for birds in southern Quebec.

Key words: Conservation areas, reserve selection, stochastic programming, conservation biology, biodiversity.

9.1 Introduction

Conservation areas are broadly defined as sites administered for the protection of threatened species and other features of biodiversity. However, many conservation areas throughout the world were created not because of their biodiversity content but because they had little economic value when established [9]. As a consequence, there is growing evidence at the global scale that existing conservation areas do not represent threatened species adequately [11]. Thus, tools are needed for selecting and refining conservation areas worldwide. Globally, approximately 40 million USD is spent annually to protect biodiversity hotspots [10]. To make the best use of this limited funding, conservation areas must be selected in such a way that the cost of acquiring and managing the land and the foregone opportunity cost to local human communities are as small as possible.

An effective way to formulate this planning problem is to represent it as a linear integer mathematical program. In general, a mathematical program is an optimization model that takes as input data parameters. Solving the mathematical program amounts to selecting values for decision variables that optimize an objective function while obeying one or more constraints. In our setting, the objective is to minimize the cost of the selected sites. Alternatively, the objective may be to maximize the number of species (or other biodiversity surrogates) protected subject to a budgetary constraint. Typically, the program includes constraints to ensure that the selected sites contain sufficient habitat for each species. Though this approach has proven useful in many planning contexts [15], it makes two unrealistic assumptions: that there is no uncertainty in the data parameters of the mathematical program and that the conservation areas are selected all at once.

Here, we relax these assumptions by modeling conservation planning problems using stochastic programming, which is a branch of operations research concerned with optimization under uncertainty [1]. The uncertainty is represented by having some of the data parameters in the mathematical program be random variables whose values are determined by a random experiment. Formally, each random variable is a mapping from the sample space Ω to \mathbb{R} and the outcomes of the random experiment constitute a σ -algebra F on Ω in the probability space (Ω, F, P) , where P is a probability measure on F [7]. In a two-stage stochastic program, the first-stage decision must be taken before the specific values of the random variables are known. Then a random experiment is conducted and the values of the random variables are disclosed. In the second stage, a recourse decision is taken to respond to, or compensate for, the results of the random experiment. The objective of the stochastic program is to minimize the cost of the first-stage decision plus the expected value of some function of the first-stage decision and the random variables. This two-stage framework has been used profitably in many areas of environmental planning, including acid rain control [14], water supply reliability modeling [8], and biodiversity conservation [6, 13].

In our two-stage stochastic integer program, sites are selected, based on their species composition, to be included in a network of conservation areas in stage one (Appendix, Fig. 9.1). There is no uncertainty about the locations of the species in stage one. However, the stochastic program includes random variables that represent the locations of the species after the first stage. After the stage-one decision, a random experiment is conducted and the species' new locations are disclosed. This models the relocation of species due, e.g., to anthropogenic habitat disturbance. The sample space Ω of the stochastic program consists of all possible scenarios of species dispersal. F is a collection of subsets of Ω and the function $P : F \rightarrow [0, 1]$ assigns probabilities to the dispersal scenarios. In the second stage, we determine whether the sites selected in the first stage still cover each species adequately. If the species are no longer covered, the stage-one decision incurs a penalty based on the "site shortage" for each species. The site shortage is the targeted number of sites for the species less the number of selected sites that contain the species. The objective is to minimize the expected value of this penalty. The stochastic program requires a particular decision-making sequence—stage-one decision, random experiment, recourse decision—but places no restrictions on the duration of the stage-one activities or the elapsed time between the first and

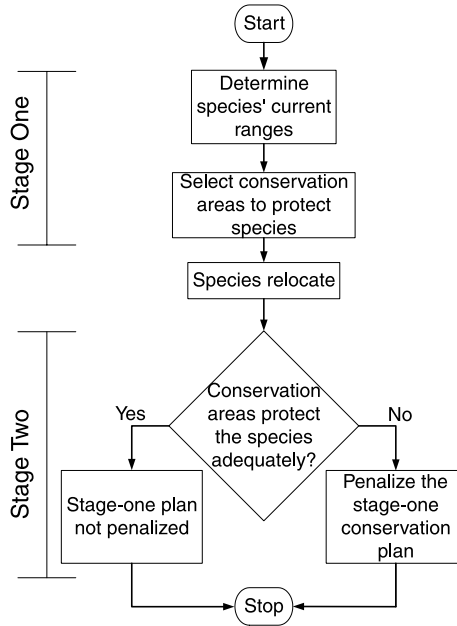


Fig. 9.1. Flowchart of conservation planning with species relocation.

second stages. This is particularly useful in conservation planning because it may take several years to establish the conservation areas (the stage-one decision) and to assess their performance (the stage-two decision).

This chapter makes two contributions. First, our stochastic program allows general targets of representation. The target of representation for a species is the number of populations of the species that should be included in the conservation areas. Previous models for conservation planning under uncertainty assume that a target of one representation is adequate [3, 6, 13]. However, if only one population of a species is protected and some conservation areas are destroyed or subject to poaching, the species may go extinct. For this reason, targets of 20 to 50 representations may be suitable for at-risk species [12]. Our stochastic program permits planners to select a suitable target for each species from one population up to all populations of the species. Second, we find the optimal solution for a substantially larger conservation planning problem that involves uncertainty than has been reported in the literature. In the context of deterministic conservation planning, decision problems with up to 1,906 sites have been solved optimally [5]. In the context of conservation planning under uncertainty, optimal solutions have been found only for much smaller problems (up to 146 sites, 116 species, and 100 scenarios of species relocation [13]). We report results for a stochastic programs that is substantially larger with respect to the number of sites, species, and species relocation scenarios.

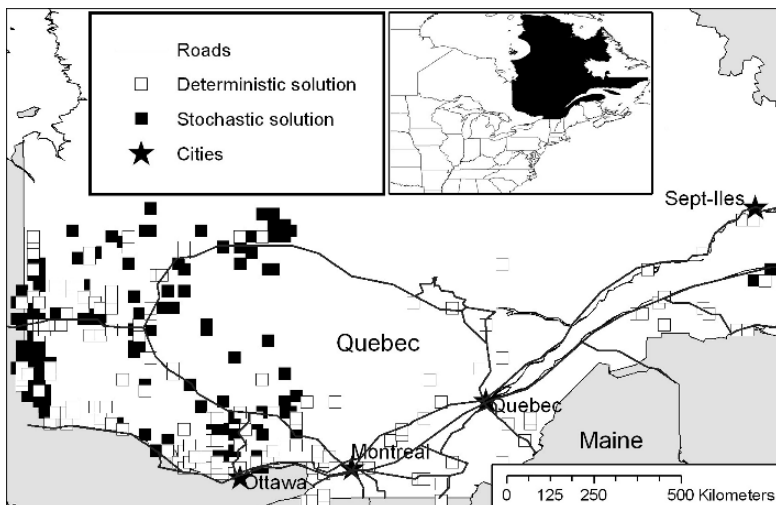


Fig. 9.2. Conservation areas for birds in southern Quebec selected using linear integer programming (white squares) and stochastic programming (black squares).

9.2 Case Study: Bird Conservation in Quebec

As part of an effort to expand the network of conservation areas of the Canadian province of Quebec, breeding bird nesting data was collected in southern Quebec between 1984 and 1994 [12]. At the $0.02^\circ \times 0.02^\circ$ scale of longitude and latitude, the data set contains 2,049 sites and 242 bird species. First, we solved the integer linear program with deterministic input parameters to find the minimum set of sites required to represent 10% of the habitat of each bird species. The optimal set of conservation areas contained 156 sites (Fig. 9.2, white squares). Next, we simulated 500 scenarios of bird dispersal after stage one and solved the resulting stochastic program with a stage-one budget of 156 sites (Appendix). The solution to the stochastic program contains more sites in southwestern Quebec (Fig. 9.2, black squares) than the solution to the deterministic planning problem. However, both solutions contain many sites adjacent to roads in western Quebec. This may be due to a bias toward roads in the bird data set because sites adjacent to roads are easier to survey. We solved the stochastic program both (i) by converting it into a mixed-integer program and solving it using a branch-and-bound algorithm and (ii) using the L-shaped method [1]. The L-shaped method decomposes the two-stage program into a master program and several sub-programs, each of which corresponds to a different scenario of relocation for the bird species. At each iteration of the algorithm, the solution of the subproblems is used to construct a piecewise linear approximation of the stage-one objective function. The mixed-integer program had 121,242 constraints and 123,049 decision variables but the L-shaped method required solving a master program with 2,049 binary decision variables and 242 initial constraints. One new constraint was added to the master program at each iteration and the algorithm converged after 7 iterations. Both methods obtained

the same optimal solution but the L-shaped method was 1.77 times faster (running time: 4,325 seconds on a 1.7GHz Dell Xeon computer with 1 GB of RAM). The mathematical program was coded in the GAMS modeling language and solved with the CPLEX 9.0 solver.

9.3 Conclusion

Unlike our two-stage program, previous models for conservation planning under uncertainty [3, 6, 13] contain integer decision variables in the second (and subsequent) stages and therefore cannot be solved with the L-shaped method. In some previous optimization models [3, 13], the objective was to maximize the number of species covered, whereas our objective is to minimize the shortage of each species from its target. Whether the min or max formulation is more suitable depends on the planning context.

Here, we represent conservation area selection problems that include uncertainty as stochastic programs. Alternatively, such conservation planning problems can be formulated as stochastic dynamic programs [3]. However, the decision-making structure encountered in conservation planning may be more amenable to solution by stochastic programming than by stochastic dynamic programming. In conservation planning, decisions are made infrequently because the establishment of a new conservation area requires years of effort, and stochastic dynamic programming is often better suited for problems with many time stages in which decisions are made in each stage. We refer the reader to [2, 4] for further discussion comparing stochastic programming and approaches in stochastic dynamic programming and optimal control.

Appendix

Sets

- $i \in I$ species
- $j \in J$ sites
- $\omega \in \Omega$ species dispersal scenarios

Data Parameters

- c_j cost of site j in stage 1. $c \in \mathbb{R}_+^{|J|}$
- t_i target for species i . $t_i \in [0, 1, 2, \dots, |J|]$
- b_1 stage 1 budget.

Random Data

- p^ω probability of scenario ω . $p^\omega \in [0, 1]$, $\sum_{\omega \in \Omega} p^\omega = 1$
- b_{ij}^ω 1 if species i is in site j in scenario ω . 0 otherwise
- $b^\omega \in \{0, 1\}^{|I| \times |J|}$

Decision Variables

- x_j 1 if site j is selected in stage one. 0 otherwise. $x \in \{0, 1\}^{|J|}$
- y_i^ω tally the site shortage for species i in scenario ω . $y_i^\omega \in [0, t_i]$

Formulation

$$\min_x \sum_{\omega \in \Omega} p^\omega Q(x, b^\omega) \quad (9.1)$$

$$\text{s.t. } \sum_{j \in J} c_j x_j \leq b_1 \quad (9.2)$$

$$x_j \in \{0, 1\}, \quad j \in J, \quad (9.3)$$

where

$$Q(x, b^\omega) = \min_{y^\omega} \sum_{i \in I} y_i^\omega \quad (9.4)$$

$$\text{s.t. } y_i^\omega \geq t_i - \sum_{j \in J} b_{ij}^\omega x_j, \quad i \in I \quad (9.5)$$

$$y_i^\omega \geq 0, \quad i \in I. \quad (9.6)$$

The stochastic program (9.1)–(9.3) is expressed in terms of the stage-one decision variables, which select sites to be protected. Constraint (9.2) ensures that the cost of the sites selected in stage one does not exceed the budget. Constraint (9.3) states that each site must be selected or not selected in stage one. The objective function in (9.4) sums the tallies of site shortages for each species, i.e., the (positive) amount by which the species' target exceeds the number of sites selected in stage one that contain the species under scenario ω . The overall objective function in (9.1) then takes the expected value of the number of species-site shortages over all b^ω scenarios. Together, constraints (9.5) and (9.6) capture $\max\{t_i - \sum_{j \in J} b_{ij}^\omega x_j, 0\}$. This is the total number of sites short species i is relative to its target. For the Quebec data set, we generated scenarios of species dispersal by having each bird species randomly select one of the sites adjacent to its location in stage one and relocate to the selected site in stage two.

References

1. Birge, J. R., Louveaux, F.: *Introduction to Stochastic Programming*. Springer, Berlin (1997).
2. Cheng, L., Subrahmanian, E., Westerberg, A.: A comparison of optimal control and stochastic programming from a formulation and computation perspective. *Computers and Chemical Engineering*, **29**, 149–164 (2004).
3. Costello, C., Polasky, S.: Dynamic reserve site selection. *Resource and Energy Economics*, **26**, 157–174 (2004).
4. Dupačová, J., Sladký, K.: Comparison of multistage stochastic programs with recourse and stochastic dynamic programs with discrete time. *Zeitschrift für angewandte Mathematik und Mechanik*, **82**, 753–765 (2002).
5. Fischer, D. T., Church, R. L.: The SITES reserve selection system: a critical review. *Environmental Modeling and Assessment*, **10**, 215–228 (2005).
6. Haight, R. G., Snyder, S. A., Revelle, C. S.: Metropolitan open-space protection with uncertain site availability. *Conservation Biology*, **19**, 327–337 (2005).
7. Karr, A. F.: *Probability*. Springer, Berlin (1993).
8. Lund, J. R.: Derived estimation of willingness to pay to avoid probabilistic shortage. *Water Resources Research*, **31**, 1367–1372 (1995).

9. Margules, C., Pressey, R.: Systematic conservation planning. *Nature*, **405**, 243–253 (2000).
10. Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., Ken, J.: Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858 (2000).
11. Rodrigues, A. S., and twenty others.: Effectiveness of the global protected area network in representing species diversity. *Nature*, **428**, 640–643 (2004).
12. Sarakinos, H., Nicholls, A. O., Tubert, A., Aggarwal, A., Margules, C. R., Sarkar, S.: Area prioritization for biodiversity conservation in Québec on the basis of species distributions. *Biodiversity and Conservation*, **10**, 1419–1472 (2001).
13. Snyder, S. A., Haight, R. G., ReVelle, C. S.: Scenario optimization model for dynamic reserve site selection. *Environmental Modeling and Assessment*, **9**, 179–187 (2004).
14. Watanabe, T., Ellis, H.: Stochastic programming models for air quality management. *Computers and Operations Research*, **20**, 651–663 (1993).
15. Williams, J. C., ReVelle, C. S., Levin, S. A.: Using mathematical optimization models to design nature reserves. *Frontiers in Ecology and the Environment*, **2**, 98–105 (2004).