Wildlife Conservation Planning Using Stochastic Optimization and Importance Sampling

Robert G. Haight and Laurel E. Travis

ABSTRACT. Formulations for determining conservation plans for sensitive wildlife species must account for economic costs of habitat protection and uncertainties about how wildlife populations will respond. This paper describes such a formulation and addresses the computational challenge of solving it. The problem is to determine the cost-efficient level of habitat protection that satisfies a viability constraint for a sensitive wildlife population. The viability constraint requires a high probability of attaining a population size target. Because of the complexity of wildlife prediction models, population survival probabilities under alternative protection plans must be estimated using Monte Carlo simulation. The computational challenge arises from the conflicting effects of sample size: fewer replications used to estimate survival probability increases the speed of the search algorithm but reduces the precision of the estimator of the optimal protection plan. Importance sampling is demonstrated as a simulation technique for reducing estimator variance for a given sample size, particularly when the tail of the population distribution is of critical importance. The method is demonstrated on a hypothetical problem involving gray wolf management in the Great Lakes region of the United States. In comparison to random sampling, importance sampling produces a 21-fold reduction in the variance of the estimator of the minimum-cost protection plan. Results from the optimization model demonstrate the extreme sensitivity of the minimum-cost protection plan to the structure of the growth model and the magnitude of environmental variation. This sensitivity is not widely recognized in the literature on wildlife habitat planning and is a strong reason for using optimization methods that can handle stochastic population models with a wide range of structures. For. Sci. 43(1):129–139.

Additional Key Words. Importance sampling, metapopulation, Monte Carlo simulation, population modeling, retrospective optimization, wildlife management.

Protecting habitat for sensitive wildlife species is a primary concern of national forest planners in the United States (Salwasser et al. 1984, Wilcove 1993). Determining appropriate conservation plans is difficult because decision-makers and the public are concerned not only with foregone economic benefits of additional habitat protection but also with uncertainty about how wildlife populations respond to habitat protection. Consequently, methods for developing conservation plans must account for both economic costs and biological uncertainties.

This paper presents such a method. The problem is to determine the minimum-cost level of habitat protection that satisfies a viability requirement for a sensitive wildlife population. The formulation has several strengths. First, it is consistent with an approach increasingly used to evaluate habitat protection plans for endangered wildlife species (Shaffer 1981, Boyce 1992). That approach uses population viability analysis to predict the likelihood that an endangered population will survive over some time horizon (e.g., Lamberson et al. 1992, 1994). Second, the formulation recognizes that the likelihood of population survival is an incremental quantity that depends on the amount of habitat protected thereby allowing the estimation of the incremental tradeoffs between the cost of habitat protection and the likelihood of population survival (Conrad and Salas 1993, Montgomery et al. 1994, Montgomery 1995, Haight 1995).

Robert Haight is Research Forester, USDA Forest Service, North Central Forest Experiment Station, St. Paul, MN 55108 (haigh003@maroon.tc.umn.edu); and Laurel Travis is Assistant Professor, Information and Computer Sciences Department, Metropolitan State University, Minneapolis, MN.

Acknowledgments: This research was supported by the North Central Forest Experiment Station. We thank Mike Bevers, Alan Murray, Barry Nelson, Bruce Schmeiser, Mike Taaffe, and the anonymous referees for many suggestions that improved the manuscript.

Manuscript received September 29, 1995. Accepted April 5, 1996.
This article was written by a U.S. Government employee and is therefore in the public domain.

Finally, the formulation meets the intent of USDA Forest Service regulations that require protection of wildlife populations at minimum cost.

We address the computational challenge of solving this habitat protection problem. The viability constraint requires a high probability of attaining a population size target. Models for predicting survival probability are usually spatially explicit and assume nonlinear density-dependent growth (see Simberloff 1988, Boyce 1992, and Dunning et al. 1995 for reviews). Because of this complexity, survival probabilities under alternative protection plans must be estimated using Monte Carlo simulation. The computational challenge arises from the conflicting effects of sample size: fewer replications used to estimate survival probability increases the speed of the search algorithm used to find optimal protection plans but reduces the precision of the estimator of the optimal plan. Importance sampling is demonstrated as a simulation technique for reducing estimator variance for a given sample size thereby improving the efficiency of estimating the minimum-cost protection plan with a desired level of precision.

Investigating the sensitivity of minimum-cost habitat protection to assumptions contained in the wildlife prediction model is an important component of conservation planning (e.g., Montgomery 1995). Sensitivity analysis is required because observations of population dynamics under alternative habitat configurations are usually not available to determine the reliablity of model forecasts (Conroy et al. 1995). Two components of wildlife models that greatly affect population predictions are the structure of the relationship between population growth and crowding and the magnitude of environmental variation affecting population growth (Boyce 1992). In the context of a hypothetical problem involving the protection of a gray wolf *(Canis lupus)* population in the upper Great Lakes region of the United States, we demonstrate the importance of these model assumptions on wolf population predictions and the amount of habitat required to meet a viability requirement for the wolf population.

The paper proceeds as follows. The first section presents the habitat protection problem. The second section describes an approach to optimization and highlights stochastic simulation methods that improve search efficiency. The third section describes an application of the problem formulation and solution method to a hypothetical problem involving the management of gray wolves. We conclude with a discussion of the results and avenues for further research.

The Habitat Protection Problem

To demonstrate our methodology, we formulated a relatively simple habitat protection problem. The formulation is for a landscape composed of areas of prime habitat for a sensitive wildlife population. The areas are fragmented and surrounded by unsuitable habitat. For example, habitat areas could be areas of forest surrounded by agricultural land. For simplicity, attributes of the habitat other than size are constant over time. The formulation could include predictions of habitat attributes such as vegetation structure and composition that affect habitat quality over time.

The formulation assumes that periodic predictions of wildlife population size can be made with a system of difference equations. Each habitat area supports a portion of the regional wildlife population. During each time period, the population in each area grows and then disperses. Population growth is a function of the current population size, the size of the habitat area, and random growth conditions. The number of dispersers depends on crowding, and the destination of the dispersers depends on the distance to and size of neighboring areas.

The formulation assumes that economic development activities such as forestry and farming are land uses that compete with preservation by destroying habitat. The problem is to determine how much of the existing habitat should be preserved for the sensitive wildlife population and how much should be used for economic development. Although the formulation assumes that economic development destroys habitat, this need not be the case. The formulation could include activities that have varying degrees of either positive or negative effects on habitat quality and abundance.

The habitat protection problem is formulated to be consistent with USDA Forest Service planning regulations. These regulations require that enough habitat be protected to maintain a viable wildlife population. At the same time, the regulations require the development of cost-efficient protection plans. Assuming that the most significant cost of habitat protection is foregone economic development, we determine the minimum-cost protection plan subject to a viability constraint for the sensitive population. The problem formulation is:

$$\begin{aligned} \min_{0 \le x \le u} \quad & cx \\ s.t. \quad & \Pr[N(x;E) \ge n] > p \end{aligned} \tag{1}$$

where

u ≡ the vector of initial sizes of habitat areas

x ≡ the vector of decision variables for habitat areas

c ≡ the vector of unit costs of preservation

E ≡ the matrix of random environmental effects

$N(x; E)$ ≡ the population size at the end of the horizon

n ≡ the target population size

p ≡ the margin of safety.

The decision variables x are the sizes of the habitat areas reserved from development. Assuming there are k areas, $u, x,$ and c are vectors of k positive real numbers.

Time enters the problem through the prediction of wildlife population size. Assuming there are T time periods, E is a $k \times T$ matrix of random variables (one for each habitat area and time period) representing environmental

effects on population growth (e.g., weather, human-caused mortality, and prey abundance). An example of distributions for the random variables comprising E will be given in the application, but no particular functional form is required by the solution technique. The random variable $N(x; E)$ is the sum of the period T ending populations across all habitat areas and depends on the sizes of the habitat areas and random environmental effects. Because there is no closed form expression for $N(x; E)$, the probability distribution of $N(x; E)$ is obtained using the deterministic simulation of the difference equations for population growth and dispersal within a Monte Carlo simulation of random environmental effects. The simulation methods described in the next section can be used with a wide range of population models.

The viability constraint is the foundation of the protection problem. A population is viable if its predicted size at the end of the horizon exceeds a population size target n. The target reflects the decision-maker's judgment about a lower bound on the viable population size at the end of the planning horizon. Because of random environmental effects, the predicted population size is a random variable, and whether or not the size target is exceeded must be viewed probabilistically. The resulting viability constraint is called a *safety rule* (Lichtenberg and Zilberman 1988) because it limits to some small value, $1-p$, the probability of violating the size target n. The lower bound p is a margin of safety that represents the decision-maker's aversion to uncertainty about reaching the target. Because the consequences of not having a viable population are dire, the margin of safety is typically set high (e.g., 0.950, 0.990, or even 0.999). If the constraint is not satisfied when all of the available habitat is preserved, the problem is infeasible and the decision-maker's viability goals must be revised or remain unmet.

Solution Method

The approach to solving the habitat protection problem involves a search algorithm that systematically evaluates the costs and population survival probabilities of alternative protection plans. This section shows how the survival probability $\Pr[N(x; E) \geq n]$ is estimated using Monte Carlo simulation, discusses the effects of sample size on estimation and optimization, and introduces two methodologies—importance sampling, and retrospective optimization—for improving the efficiency of optimization.

Estimating Pr $[N(x; E) \geq n]$

Estimating $\Pr[N(x; E) \geq n]$ for a given vector of reserve sizes x involves the deterministic simulation of the difference equations for population growth and dispersal within a Monte Carlo simulation of random environmental effects. Each deterministic simulation is the computation of $N(x; \varepsilon)$ where ε is one realization of the matrix of random variables comprising E. We call ε a *scenario* of environmental effects. For fixed x and ε, $N(x; \varepsilon)$ is obtained by computing population growth and dispersal in each time period and summing the populations across reserves to obtain the ending population size.

An estimator of $\Pr[N(x; E) \geq n]$ is computed using Monte Carlo simulation. We randomly select a set of d scenarios, denoted $\{\varepsilon\}$, and compute $N(x; \varepsilon)$ under each scenario. Each $N(x; \varepsilon)$ is considered a *success* or *failure* depending on whether or not it exceeds the target n. Because the outcome is a Bernoulli random variable, the mean of d Bernoulli trials [i.e., the percent of d scenarios for which $N(x; \varepsilon) \geq n$] is an unbiased estimator of the actual value of $\Pr[N(x; E) \geq n]$. Letting Y_i be the outcome of the ith scenario ($Y_i = 1$ if $N(x; \varepsilon) \geq n$ and $Y_i = 0$ if $N(x; \varepsilon) < n$), the mean of the d Bernoulli trials is

$$\hat{p} = \frac{\sum_{i=1}^{d} Y_i}{d} \tag{2}$$

The estimator $\hat{p}$ is compared with the margin of safety p to estimate whether the vector of reserve sizes x satisfies the constraint $\Pr[N(x; E) \geq n] > p$.

Sample Size

The number of scenarios in $\{\varepsilon\}$ affects the precision of the estimator $\hat{p}$. Precision is measured by the confidence interval half-width for $\hat{p}$, which is proportional to

$$\hat{\sigma}(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{d-1}}.$$

Because the viability constraint requires a small probability, $1-p$, of violating the population size target, feasible protection plans fail to meet the target only under rare scenarios of relatively severe environmental effects. The estimator $\hat{p}$ depends on the number of these rare scenarios in the set $\{\varepsilon\}$, and the precision of the estimator increases as d increases. Because $\hat{p}$ is sensitive to the number of rare scenarios in $\{\varepsilon\}$, a large sample size d is required for a desirable level of precision. For example, if $1-p = 0.01$, a sample size of about 10,000 scenarios is required for $\hat{\sigma}(\hat{p})$ to be less than 10% of $1-p$.

The precision of the estimators of $\Pr[N(x; E) \geq n]$ under alternative protection plans affects the precision of the estimator of the minimum-cost feasible protection plan. The estimator of the optimal plan is the mean of solutions obtained using different sets $\{\varepsilon\}$. If the number of scenarios in each $\{\varepsilon\}$ is small, estimates of $\Pr[N(x; E) \geq n]$ under any given x will vary widely across sample sets $\{\varepsilon\}$. This variability, in turn, will cause the set of feasible plans and the minimum-cost plan to vary widely across sample sets $\{\varepsilon\}$. Consequently, increasing the number of scenarios in $\{\varepsilon\}$ increases the precision of estimators of $\Pr[N(x; E) \geq n]$ and the optimal protection plan.

The number of scenarios in $\{\varepsilon\}$ also has important consequences for the speed of the optimization algorithm. A variety of optimization techniques could be used to determine the cost-minimizing protection plan. For instance, the constraint $\Pr[N(x; E) \geq n] > p$ could be brought into the

objective function using a Lagrangian multiplier, and a gradient search could be performed. Any technique will require the evaluation of $\Pr[N(x; E) \geq n] > p$ for many values of x. Thus, any reduction in the number of scenarios in $\{\varepsilon\}$ will greatly increase the speed of the search algorithm. The next subsection describes a sampling technique that reduces the standard error of the estimator $\hat{p}$ for a given number of scenarios in $\{\varepsilon\}$ thereby reducing the need for a large sample.

Importance Sampling

A class of simulation tools used to address problems in which solutions are sensitive to rare, stochastic events is called *importance sampling* (see Nelson and Schmeiser 1983 for an intuitive discussion of these techniques). Importance sampling methods force a larger number of the rare events from the underlying distribution into a sample that is used for simulation. Inference from the simulations is done in a way to correct for sampling bias.

We use an importance sampling method called *Russian roulette* (see Hammersley and Handscomb 1964) to obtain the set of scenarios $\{\varepsilon\}$ used for simulation and optimization. The method involves repeatedly drawing scenarios and classifying each according to the size of the population it is likely to produce. Suppose we define two classes. Class 1 includes worst-case scenarios that are judged likely to produce one of the smallest population sizes regardless of the degree of habitat protection. Class 2 includes not-bad scenarios. Details on how these judgments are made will be given in the application section. Suppose we keep 100% of the worst-case scenarios and 1% of the not-bad scenarios. The sampling stops when a set of d scenarios is obtained, which requires, say, D trials.

Because the set $\{\varepsilon\}$contains more worst-case scenarios and fewer not-bad scenarios than would be obtained in an ordinary random sample, the scenarios need to be weighted according to the number of scenarios they represent. Let m_j be the number of scenarios remaining in class j after discards where

$$\sum_{j=1}^{2} m_j = d$$

and let M_j be the total number of trials in class j where

$$\sum_{j=1}^{2} M_j = D.$$

Let w_j be the reciprocal of the proportion of scenarios in class j that are kept (i.e., $w_j = M_j/m_j$). Because we kept every worst-case scenario, $w_1 = 1$. Because we discarded 99% of the not-bad scenarios, each kept scenario represents 100 not-bad scenarios and $w_2 = 100$.

The sample of scenarios $\{\varepsilon\}$ and their weights are used to estimate the success probability $\Pr[N(x; E) \geq n]$ of a given protection plan x. Recall that, with random sampling, $\hat{p}$ is the percent of scenarios in $\{\varepsilon\}$ for which $N(x; \varepsilon) \geq n$ (Equation 2). With importance sampling, the estimator depends on the scenario weights. Letting Y_{ij} be the outcome of the ith scenario in class j ($Y_{ij} = 1$ if $N(x; \varepsilon) \geq n$ and $Y_{ij} = 0$ if $N(x; \varepsilon) < n$)

$$\hat{p} = \frac{\sum_{j=1}^{2} w_j \left(\sum_{i=1}^{m_j} Y_{ij} \right)}{D} \tag{3}$$

which recognizes that each scenario was obtained from a sample of size D. Even though the set $\{\varepsilon\}$ contains a larger proportion of worst-case scenarios than would be obtained in an ordinary random sample, the weights assure that $\hat{p}$ is not biased (i.e., $E\{\hat{p}\} = E\{\Pr[N(x; E) \geq n]\}$). This design for obtaining a sample set $\{\varepsilon\}$ gives a better estimator for $\Pr[N(x; E) \geq n]$ (i.e., an estimator with lower standard error) than a sample set of the same size obtained with random sampling. We will show in the application that importance sampling also produces a more precise estimator for the optimal habitat protection plan.

Although the importance sampling technique provides a better estimator for $\Pr[N(x; E) \geq n]$, the method requires more computational effort than random sampling because a larger number of scenarios must be drawn and evaluated. In the next subsection we describe an optimization technique that allows us to use our sample, once drawn, to maximum advantage.

Retrospective Optimization

Retrospective optimization is a technique that reduces the overhead involved in repeatedly computing sets of scenarios $\{\varepsilon\}$. In the course of an optimization run, many values of x are considered. Rather than drawing a new sample $\{\varepsilon\}$ to estimate $\Pr[N(x; E) \geq n]$ for each x, one sample set $\{\varepsilon\}$ is drawn and used for all values of x examined. This repeated use of the same sample set reduces overhead because the sample set is only drawn once. This technique was developed for the optimization of Monte Carlo simulations by Healy and Schruben (1991) and was further improved by Rubinstein and Shapiro (1993), Gürkan et al. (1994), and Chen and Schmeiser (1995).

Confidence Intervals

In stochastic optimization, a single optimization run gives only a single point estimate of the optimal solution. With no confidence interval, the decision-maker has no way of knowing whether other optimization runs would produce similar or dramatically different results. Such point estimates are not sufficient for decision-making. We compute a confidence interval for the optimal solution by repeating the entire optimization for several sample sets $\{\varepsilon\}$ (see Figure 1 for flowchart). The half-width of the confidence interval decreases as the number of optimizations increases and the variance of the estimator of the optimal x decreases. Because importance sampling should reduce the variance of the estimator of optimal x, fewer replications of the optimization should be needed to produce a narrow confidence interval. In the following application, we demonstrate this variance reduction by comparing solutions obtained using importance sampling and random sampling methods.

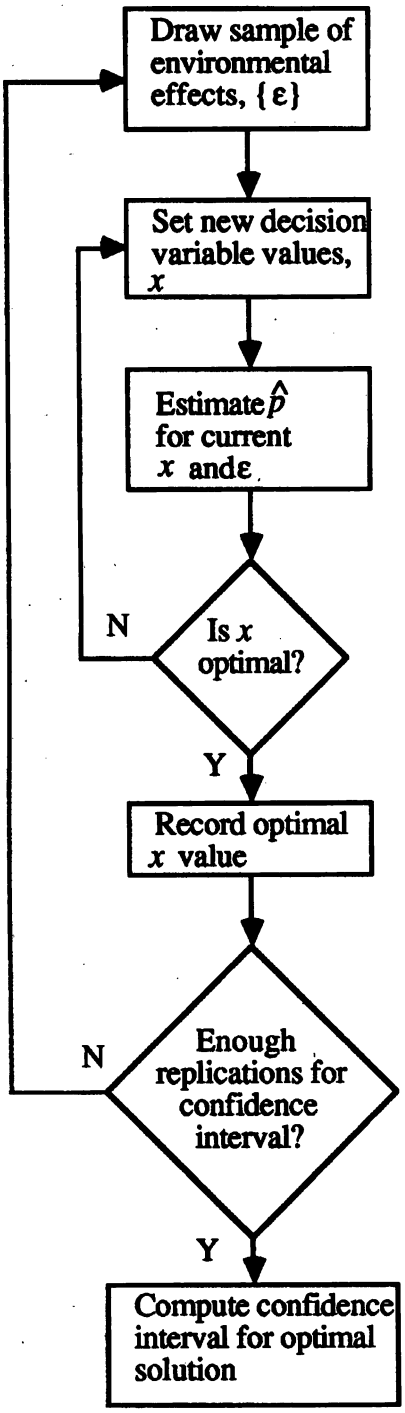

Figure 1. Flowchart of retrospective optimization technique. The inside loop represents iterations of a single optimization run performed with a single sample set {ε}. The outside loop represents several optimization runs performed with different sample sets to obtain a confidence interval for the optimal solution. If retrospective optimization were not used, the drawing of the sample set {ε} would be inside both loops.

Application: Gray Wolf Management

The numerical solution method is demonstrated on a habitat protection problem for gray wolves, an endangered species currently repopulating the upper Great Lakes region of the United States. This section describes a model for wolf dynamics and a habitat protection problem involving three separate wolf populations linked by dispersal. The problem is used to demonstrate three things:

- the variance reduction using a solution method based on importance sampling,
- the tradeoffs between the risk of population extinction and cost of habitat protection,
- the effects of population growth assumptions on the cost of habitat protection.

Background

Except for small breeding populations in northern Minnesota and Isle Royale, gray wolves were extirpated from the eastern United States by 1960 (Mech 1995). Since receiving legal protection in 1974 under the Endangered Species Act of 1973, their numbers and distribution in the upper Great Lakes region have greatly increased. Wolves in northern and central Minnesota currently number around 2,000 (Mech 1995). Individuals from Minnesota have recolonized areas of northern Wisconsin and upper Michigan where they currently number over 100 (Mech 1995). Assuming that the population in Wisconsin and Michigan does not drop below 100, the criterion for recovery of the eastern timber wolf population (U.S. Fish and Wildlife Service 1992) should be satisfied before the year 2000.

An important problem facing managers of wolves in the Lake States involves the protection and enhancement of small isolated populations (Fuller 1995). The landscape of northern Wisconsin and upper Michigan is a mixture of forest, agricultural, and developed land under a variety of public and private ownerships. Within this landscape, wolves have settled in forested areas with few roads or human settlements (Mladenoff et al. 1995). Because favorable habitat occurs in small fragmented areas, the recovering population includes local populations that are isolated in the sense that immigration is relatively infrequent (Wydeven 1993, Hammill 1993). Protecting isolated populations is important not only for recovery, but also for postrecovery management. After the wolf is declassified, management agencies may choose a zoning approach that includes identifying relatively small areas of favorable habitat where wolves may live free from human exploitation (Mech 1995).

From detailed studies of wolf populations throughout North America (see Fuller 1995 for review), managers have gained an understanding of the relationships between factors that affect wolf population growth and persistence. These factors include mortality, dispersal, and habitat carrying capacity (Fuller 1995). In fragmented populations like those that inhabit northern Wisconsin and Michigan, dispersal is extremely important because immigration can greatly improve the persistence of small isolated populations subject to high mortality rates (e.g., Fuller 1989).

Hypothetical Problem

This background along with knowledge of the current and potential wolf distribution in the upper Great Lakes region (Mladenoff et al. 1995) provides the context for a

hypothetical problem involving the protection and enhancement of isolated wolf populations. Suppose there are two areas of favorable wolf habitat that are reserved from development and surrounded by land unsuitable for wolves. Although one area is saturated with wolves (e.g., northeastern Minnesota), the other area is just beginning to be recolonized (e.g., northern Wisconsin). Between the two broad areas is a small island of favorable habitat that has been colonized by migrants and is the primary source of immigration into the newly colonized reserve. The island, which is not a reserve, is under threat of development from farming, forestry, and real-estate interests. The problem is how much of the island should be protected from development in order to create a viable wolf population in the newly colonized reserve.

The problem has features that typify environmental disputes. Protecting more of the island habitat enhances the viability of the colonizing population in the second reserve by increasing the rate of immigration, which in turn provides a buffer against chance mortality events. However, increased protection comes at a cost in terms of foregone economic development. Estimating the tradeoffs between the degree of protection afforded the island habitat and the cost of protection should enhance the decision process.

A Model for Wolf Dynamics

To address this problem, we built a spatially explicit population model for gray wolf dynamics. The model has three assumptions that are common to wildlife models used in conservation planning: the regional population is separated into a number of subpopulations that interact through dispersal, the growth of each subpopulation is a nonlinear function of crowding, and growth is subject to random environmental effects.

The model assumes that there are three areas of prime wolf habitat. Let $x = (x_1, x_2, x_3)$ be the sizes of the habitat areas where x_1 and x_3 are large protected areas and x_2 is a small island of favorable habitat between x_1 and x_3. Areas x_1 and x_3 are fixed and each equal to 5,000 km^2. Assuming that a pack of five wolves requires a territory of 125 km^2 (Fuller 1995), the carrying capacity of each area is 200 wolves. The initial area of x_2, the decision variable, is 1,250 km^2 with a carrying capacity of 50 wolves. The initial population sizes for x_1, x_2, and x_3, are 150, 25, and 10 wolves, respectively.

The model uses density-dependent growth and dispersal equations for the wolf population in each area. Equation parameters and subsequent growth and dispersal predictions reflect habitat quality as described below.

The growth of each population is modeled with a modified logistic equation. Let $N_{i,t}(x; E)$ be the population size in area i at the beginning of period t. We compute $N_{i,t+1}(x; E)$ by first computing population growth and then dispersal. Letting $M_{i,t}(x; E)$ be the postgrowth but predispersal population size, the growth dynamics are:

$$M_{i,t}(x;E) = N_{i,t}(x;E)\exp\{R_i(N_{i,t},x_i) + E_{i,t}\} \qquad (4)$$

where $R_i(N_{i,t}, x_i)$ is the instantaneous per-capita population growth rate and $E_{i,t} \sim N(0, \sigma_i^2)$ is the environmental variation in the growth rate in patch i period t.

The per-capita growth rate R_i depends on crowding. The expression for crowding, $N_{i,t}(x; E)/k_i(x_i)$, is the ratio of current population size $N_{i,t}$ and carrying capacity $k_i(x_i)$ where $k_i(x_i)$ is an increasing function of reserve size x_i. The central assumption in a logistic model of population dynamics is that per-capita growth rate decreases linearly with crowding and becomes negative when population size exceeds carrying capacity:

$$R_i(N_{i,t},x_i) = r_i - r_i\left(\frac{N_{i,t}(x;E)}{k_i(x_i)}\right) \qquad (5)$$

where r_i is the intrinsic rate of population increase at low population sizes. However, in some populations there is a crowding threshold below which per-capita growth rate decreases with decreasing crowding. This so-called *Allee* effect can have a variety of causes including the uncertainty of mates finding each other in a sparse population (Dennis 1989). The Allee effect is modeled by subtracting from the per capita growth rate obtained with the logistic model a term for growth reduction because of mating shortage so that

$$R_i(N_{i,t},x_i) = r_i - r_i\left(\frac{N_{i,t}(x;E)}{k_i(x_i)}\right) - \left(\frac{a}{a + N_{i,t}(x;E)}\right) \qquad (6)$$

where a represents the level of crowding at which the mating success rate is 0.50 (Dennis 1989). With the Allee term, per-capita growth rate increases with crowding, peaks, and then drops below zero as the population exceeds carrying capacity (Figure 2).

In the baseline problem, we used Equation (6) to model population growth rate because the likelihood of population survival diminishes in small wolf populations in

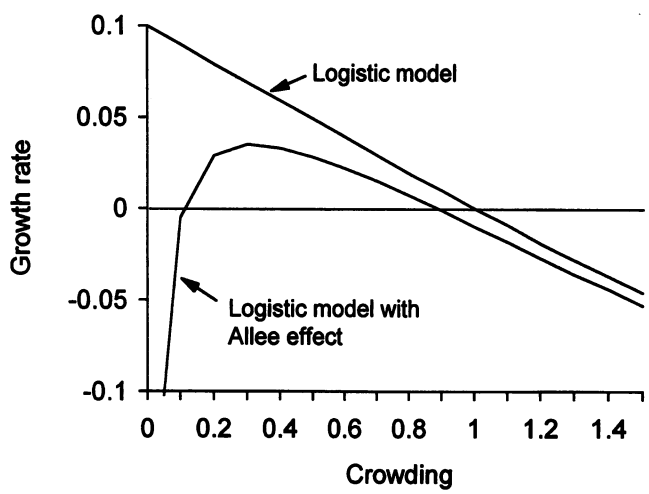

Figure 2. Relationship between instantaneous per capita growth rate and crowding using a logistic model with an Allee effect. The intrinsic rate of population increase [r_i in Equation (6)] is 0.10 and the Allee parameter [a in Equation (6)] is 0.01.

which lone wolves have difficulty finding mates (Fuller 1995). For each population, the intrinsic rate of population growth r_i is 0.10 and the Allee parameter a is 0.01. With these parameter values, population growth rates peak at about 4% when crowding is about 40% and drop as population size approaches carrying capacity (Figure 2). This range of growth rates is consistent with observations of wolf populations that are subject to moderate levels of human-caused mortality (Fuller 1989).

The population growth rate is subject to random variation caused by unpredictable changes in environmental conditions [$E_{i,t}$ in Equation (4)]. With adequate food and protection from human persecution, wolf populations have great reproductive potential with maximum observed growth rates as high as 50% per year in colonizing populations (Fuller 1995). At the other extreme, populations subject to intense human-caused mortality can decrease as much as 50% per year (Fuller 1995). We assume that each population is subject to a moderate degree of environmental variation ($\sigma_i = 0.05, i = 1, \ldots, 3$) so that the population growth rate ranges between ±15% most of the time.

Following population growth, individuals disperse between areas. The goals of dispersing wolves are to find mates and enough resources to propagate, and the distance traveled usually depends on mate and resource availability (Mech 1987). If dispersers find mates and resources locally, they settle close to their natal territory. Long-distance dispersal (i.e., between reserve areas in our model) is more common when an area is saturated. Consequently, we assume that the per-capita rate of emigration, $D_{i,t}$, increases linearly with crowding:

$$D_{i,t}(x;E) = \beta_i \frac{M_{i,t}(x;E)}{k_i(x)} \tag{7}$$

We assume only small proportions of the populations in the large protected areas (x_1 and x_3) become long-distance dispersers ($\beta_1 = 0.02$), while a larger percent of the population in the smaller area x_2 move between reserves ($\beta_2 = 0.10$).

The destination of dispersers is modeled with a matrix of destination parameters. The fraction of dispersers from area j that are in area i after dispersal is a constant γ_{ji}, where

$$\sum_i \gamma_{ji} = 1.$$

We assume that all of the dispersers from areas 1 and 3 move to area 2. Dispersers from area 2 are equally likely to reach areas 1 or 3. The equations for dispersal link the populations and complete the periodic growth cycle:

$$\begin{aligned} N_{i,t+1}(x;E) &= M_{i,t}(x;E) - M_{i,t}(x;E)D_{i,t}(x;E) \\ &\quad + \sum_{j=1}^{3} \gamma_{j,i} M_{j,t}(x;E) D_{j,t}(x;E) \end{aligned} \tag{8}$$

Problem Formulation

The baseline problem is to determine the minimum area of x_2 required to create a viable wolf population in x_3. The viability parameters are a population size target $n = 30$ wolves and a margin of safety $P = 0.99$. The management horizon is 50 yr. The problem is:

$$\begin{aligned} \min_{0 \le x_2 \le 1250} \quad & x_2 \\ \text{s.t.} \quad & \Pr[N_{3,50}(x;E) \ge 30] > 0.99 \end{aligned} \tag{9}$$

In sensitivity analyses, we examine the cost effects of changing the population size target, the margin of safety, and parameter values of the population growth equation for wolves in area 3.

Solution Method

With only one decision variable, the problem is relatively easy to solve. Because both the objective function and the left-hand-side of the constraint are increasing functions of the decision variable, the optimal solution is the level of x_2 that just satisfies the constraint. We used binary search to find an estimate of the optimal value of x_2. The solution is an estimate because it is obtained using a given sample of scenarios $\{\varepsilon\}$. The binary search is repeated for each of 10 different sets of scenarios producing 10 estimates of the optimal solution. Our estimator for the optimal solution is the mean of these estimates.

Recall that our importance sampling procedure for determining the set of scenarios $\{\varepsilon\}$ is a version of Russian roulette that creates a sample in which a large proportion of the scenarios in $\{\varepsilon\}$ are relatively rare and severe (in terms of negative effects on population growth) sequences of random environmental effects. Our procedure for classifying severity is based on percentiles of the estimated probability distribution of ending population size in area x_3 assuming full protection of area x_2 (our results show no sensitivity to this assumption). This distribution is obtained by randomly drawing 500 scenarios ε and computing $N_{3,50}(x; \varepsilon)$ for each ε. We use percentiles 0.02, 0.05, and 1.00 for the classification. For example, each scenario that produces an ending population below the 0.02 percentile is classified as a worst-case scenario.

To compute the sample set $\{\varepsilon\}$, we repeatedly draw a scenario ε, compute $N_{3,50}(x; \varepsilon)$, and classify using the percentiles of the probability distribution. We keep all scenarios that produce population sizes below the 0.02 percentile, 10% of the scenarios that produce populations sizes between the 0.02 and 0.05 percentiles, and 0.1% of the scenarios that produce population sizes above the 0.05 percentile. Sampling is stopped when 500 scenarios are obtained. With this design, roughly 80% of the scenarios in $\{\varepsilon\}$ will be worst-case scenarios.

Results

Variance Reduction from Importance Sampling

To estimate the variance reduction from importance sampling, we computed 10 estimates of the optimal solution to

the baseline problem for each of two methods for obtaining $\{\varepsilon\}$: importance sampling and random sampling. A sample size of 500 was used for each method.

Importance sampling resulted in a 21-fold reduction in the variance of the estimate of the optimal solution (expressed as the percent of x_2 preserved as wolf habitat) relative to random sampling (Table 1). The 95% confidence interval for the estimate of the optimal solution obtained with importance sampling was almost 5 times smaller than the confidence interval obtained with random sampling and did not contain the mean solution obtained with random sampling. This suggests that the point estimate obtained from random sampling is not a particularly good estimate of the true value of the optimal solution (i.e., we are 97.5% sure that the true value is strictly greater than the point estimate from random sampling).

While importance sampling resulted in a 21-fold reduction in variance, it used much more execution time for the 10 optimization runs than did random sampling (Table 2). For an average optimization run, importance sampling used 9.5 minutes to generate a sample and 0.5 minutes to search for the optimal solution. In contrast, random sampling required practically no setup time and the same search time. Consequently, 10 optimization runs using random sampling were completed in 5 minutes whereas importance sampling required 100 minutes.

Although importance sampling required more execution time for a fixed number of optimization runs than did random sampling, importance sampling would require fewer optimization runs to obtain a fixed confidence interval width for the optimal solution. The confidence interval half-width for the estimator of the optimal solution is proportional to $\sigma/\sqrt{n}$ where n is the number of optimization runs and σ is the standard deviation of the solutions. Thus, a 21-fold increase in variance of solutions obtained with random sampling would result in a 21-fold increase in the number of optimization runs required to obtain a confidence interval of the same width as importance sampling. We estimated the execution time required for random sampling to produce a confidence interval width equal to that obtained with importance sampling using 10 replications (Table 2). In this case, both sampling methods would require about the same overall execution time.

Importance sampling should display greater efficiency as problem size increases beyond a simple one-dimensional problem. As the amount of time required to search for an optimal solution increases in larger problems, the overhead for importance sampling will be a less significant proportion of the total execution time. For example, if search time for one optimization run is 60 min. for both sampling methods, importance sampling would require only 16% more execution time for 10 optimization runs compared with random sampling (Table 3). To get the same confidence interval width, random sampling would require 18 times more execution time than importance sampling. Consequently, variance reduction via importance sampling can make larger problems tractable.

Table 1. Comparison of estimates of the minimum-cost protection plan (expressed as the percent of habitat area reserved from development) obtained using importance sampling and random sampling.

Sampling method	Minimum-cost plan		
	Mean	Variance	95% CI
Importance	0.689	1.18×10^{-3}	0.664, 0.714
Random	0.639	2.57×10^{-2}	0.524, 0.754

Table 2. Computational results (minutes of execution time) for a simple model solved using random sampling and importance sampling.

Sampling method	Execution time for 10 optimization runs[a]	Execution time for equal confidence intervals[b]
Random	5	105
Importance	100	100

[a] For each optimization run, importance sampling required 9.5 min. of setup time whereas random sampling required no setup time. The search algorithm required 0.5 min. regardless of the sampling method.

[b] Execution time for random sampling is computed so that the confidence interval would be the same as that produced by importance sampling using 10 optimization runs. Random sampling is assumed to yield a 21-fold increase in the variance of the optimal solutions for a given number of optimization runs compared with importance sampling.

Risk-Cost Tradeoffs

We used multiple runs of the optimization model to estimate the costs of meeting alternative population targets and margins of safety for the wolf population in area x_3 (Figure 3). Cost is expressed as the percent of x_2 preserved as wolf habitat. The population size target and margin of safety represent risk parameters: the risk of population extinction increases as the population size target and margin of safety decrease. The cost curves give policy makers a feel for how sharply cost changes as a function of these risk parameters. Great cost reduction can be obtained by reducing the population target from 30 to 20 wolves or reducing the margin of safety from 99% to 95%. This sensitivity of economic cost to changes in the risk parameters lets policy makers know that the parameter values must be set carefully. The willingness of policy makers and society to accept increased extinction risk in trade for economic gain will play a large part in choosing the values of the risk parameters.

Table 3. Estimates of execution time (minutes) required to solve a larger conservation problem using random sampling and importance sampling.

Sampling method	Execution time for 10 optimization runs[a]	Execution time for equal confidence intervals[b]
Random	600	12,600
Importance	695	695

[a] For each optimization run, importance sampling is assumed to requir 9.5 min. of setup time whereas random sampling required no setu time. The search algorithm is assumed to require 60 min. regardless (the sampling method.

[b] Execution time for random sampling is computed so that the confidenc interval would be the same as that produced by importance samplin using 10 optimization runs. Random sampling is assumed to yield a 2 fold increase in the variance of the optimal solutions for a given numb of optimization runs compared with importance sampling.

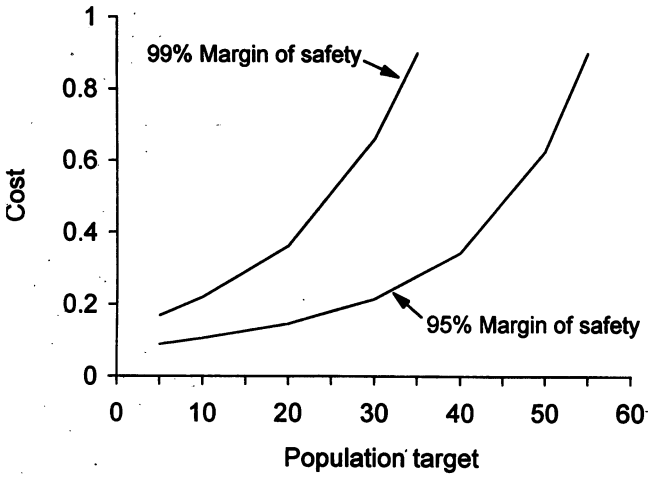

Figure 3. Cost of habitat protection as a function of the parameters of the viability constraint.

Cost Effects of Growth Assumptions

It is well known that assumptions about population growth and dispersal can greatly impact population survival probability (e.g., Stacey and Taper 1992). In our model, the cost of habitat protection is highly sensitive to changes in the parameters of the growth and dispersal equations. To demonstrate this sensitivity, we modified the baseline problem by systematically changing the values of two growth parameters for the wolf population in area 3. Intrinsic rate of growth (r_3) varied between 0.0 and 0.2, and the degree of environmental variation (σ_3) varied between 0.0 and 0.1. The minimum-cost level of habitat protection required to meet the population viability constraint in the baseline problem was calculated for each combination of parameter values (Figure 4). The cost of habitat protection increases with decreasing values of r_3 and increasing values of σ_3. For $r_3 \leq 1.08$, population growth is too low to meet the population size target regardless of the degree of environmental variation. For $r_3 \geq 1.16$, population growth is high enough so that area 2 needs no protection. For $1.16 < r_3 < 1.08$, cost depends on the degree of environmental variation. With low levels of variation, population growth is sufficient to reach the target without immigration. However, as environmental variation increases, high levels of immigration are needed to balance the effects of wide population growth fluctuations.

These results demonstrate that management actions affecting the intrinsic rate of growth and environmental variation will influence the cost of habitat protection. Examples of management actions that affect population growth and variability include increased enforcement of wolf poaching laws and restrictions on deer hunting in sparse deer years. If cost functions for these activities were incorporated in the optimization model, the optimal allocation of cost between management actions and land protection could be determined.

The Allee effect has a big impact on the cost of habitat protection. To see this, we repeated the sensitivity analysis on intrinsic rate of growth and environmental variation using the pure logistic growth model [Equation (5)] for the populations in the three habitat areas. Without the Allee effect, the costs of habitat protection are dramatically less across the range of parameter values (Figure 5). Because the logistic model predicts much higher growth rates for small populations than does the Allee model (Figure 2), the initial population in area 3 reaches the population size target with no immigration from area 2 for most combinations of parameter values. These results emphasize that assumptions about the nature of density dependence in growth rates can have a major effect on population size predictions and the levels of habitat protection required to attain a population size target.

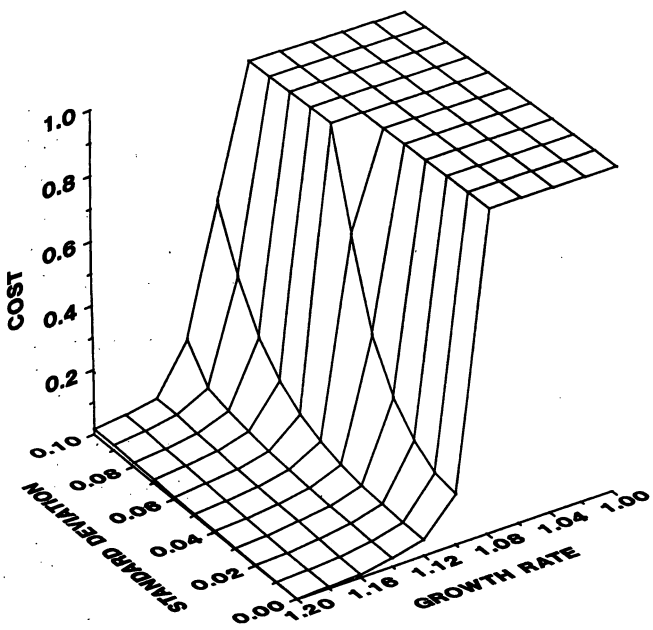

Figure 4. Cost of habitat protection as a function of the intrinsic per capita growth rate and degree of environmental variation (expressed as the standard deviation of the variability in population growth rate) using a logistic growth model with an Allee effect.

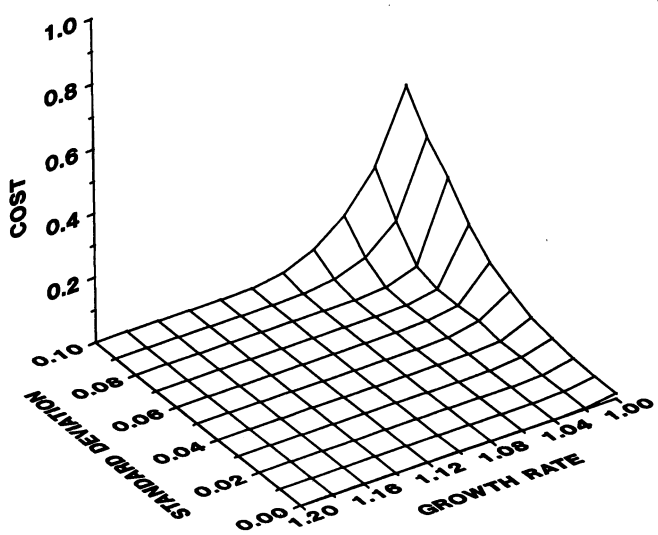

Figure 5. Cost of habitat protection as a function of the intrinsic per capita growth rate and degree of environmental variation (expressed as the standard deviation of the variability in population growth rate) using a logistic growth model with no Allee effect.

Conclusions

Problem Formulation

The formulation we present for determining cost-effective habitat protection plans for sensitive wildlife populations has two important features that distinguish it from the ways in which habitat protection is handled in most timber harvest scheduling models. First, the formulation includes uncertainty in wildlife response to habitat protection. It does so by incorporating a stochastic model of wildlife dynamics directly in the optimization framework. Accounting for uncertainty allows the estimation of incremental tradeoffs between habitat protection costs and the likelihood of population survival (Conrad and Salas 1993, Montgomery et al. 1994, Montgomery 1995, Haight 1995). Estimating the likelihood of population survival is consistent with the population viability analysis approach that is increasingly used to evaluate habitat protection plans for endangered wildlife populations (Shaffer 1981, Boyce 1992). Second, the optimization formulation accommodates population models with a wide variety of difference equation structures. This flexibility allows investigation of the sensitivity of minimum-cost protection plans to assumptions contained in the wildlife prediction model. As the analysis of wolf management indicated, predictions of population size and minimum required habitat protection can be very sensitive to assumptions about population growth rate and environmental variability. Sensitivity analysis is extremely important because there is usually a great deal of uncertainty about the appropriate model structure and parameters for a given species (Conroy et al. 1995).

It is difficult to use harvest scheduling models based on mixed-integer programming alone to study how management actions affect the dynamics of wildlife populations. These formulations usually address habitat protection by regulating aggregate areas of forest cover-types (e.g., Hof and Raphael 1993) or spatial layout of logging blocks (e.g., Hof and Joyce 1993, Weintraub et al. 1994, Yoshimoto et al. 1994) without explicit consideration of wildlife dynamics. When wildlife dynamics are incorporated (e.g., Hof et al. 1994), the dynamics are typically assumed to be linear and deterministic. Instead, combining mixed-integer programming and simulation might be appropriate. The analysis would begin with a mixed-integer programming model that allocates a fixed amount of acreage to wildlife habitat reserves while meeting protection and output requirements for other resources. Then, simulation analysis would evaluate and adjust the reserve areas using stochastic models of wildlife dynamics. Dual values from the linear programming analysis would be used as coefficients for cost minimization. The reserve areas obtained from the simulation analysis would then be used as constraints in the mixed-integer programming model to adjust activities for the other forest resources. This entire analysis could be done for multiple risk levels if knowledge of tradeoffs between risks and costs were required.

Solution Method

Incorporating wildlife dynamics in a habitat protection problem makes the problem difficult to solve because simulation must be used to estimate probabilities of population survival under alternative protection plans. Two important conclusions can be drawn from our investigation of simulation methods. First, because the population viability constraint requires a very small probability of extinction, a large number of repetitions of the wildlife model is required to obtain high levels of precision of the estimators of survival probabilities and minimum-cost protection plans. Second, for a given sampling intensity, importance sampling provides more precise estimators of survival probabilities and minimum-cost protection plans than does random sampling. Although importance sampling requires greater execution time to generate a sample, this overhead will become less important for problems with more decision variables that require greater execution time for the search algorithm. The variance reduction resulting from importance sampling may allow larger problems to be solved with reasonable precision.

Future Work

Clearly, using simulation to evaluate the risks and costs of wildlife conservation plans limits the size of optimization problems that can be solved. Nevertheless, prudent formulation of the decision problem and search methodology should make larger problems tractable.

Although our formulation is relatively simple, the solution method is simulation based, giving a wide range of possible extensions. For example, the formulation could include predictions of vegetation attributes that affect habitat quality over time and activities such as tree planting or selective logging that have varying degrees of either positive or negative effects on habitat quality. Because the number of decision variables must be kept small for optimization tractability, tradeoffs between spatial and temporal detail must be explored.

Methodological work is needed to investigate the performance of search algorithms for problems with more than one decision variable. A straightforward approach is to incorporate the viability constraint in the objective function using a Lagrangian multiplier or penalty function. However, because the estimator for $\Pr[N(x; E) \geq n]$ is the mean of a number of Bernoulli trials, the estimator is not a continuous function of x. Thus, search methods based on gradient approximation can give erratic results. One way to avoid this problem is to replace the viability constraint $\Pr[N(x; E) \geq n] > p$ with an approximation that is a continuous function of x and employ a gradient method for optimization. Another way is to use Bernoulli trials to estimate $\Pr[N(x; E) \geq n]$ but to employ search methods that do not require continuity to estimate the optimal habitat protection plan.

An alternative approach to conservation planning is to compare a small number of predefined conservation plans with different values for a range of decision variables. When the measures of performance are functions of random variables and sensitive to rare events, huge numbers of simulations are required to increase the precision of performance estimators. Importance sampling can be extremely useful in this setting by reducing the simulations needed to obtain precise performance estimators thereby increasing the statistical reliability of plan comparisons.

Literature Cited

BOYCE, M.S. 1992. Population viability analysis. Annu. Rev. Ecol. Syst. 23:481–506.

CHEN, H., AND B.W. SCHMEISER. 1995. Retrospective approximation algorithms for stochastic root finding. Purdue University, Dep. of Stat., Sch. of Indus. Eng. Res. Rep. SMS95-1.

CONRAD, J.M., AND G. SALAS. 1993. Economic strategies for coevolution: Timber and butterflies in Mexico. Land Econ. 69:404–415.

CONROY, M.J., Y. COHEN, F.C. JAMES, Y.G. MATSINOS, AND B.A. MAURER. 1995. Parameter estimation, reliability, and model improvement for spatially explicit models of animal populations. Ecol. Applic. 5:17–19.

DENNIS, B. 1989. Allee effects: Population growth, critical density, and the chance of extinction. Natur. Resour. Model. 3:481–538.

DUNNING, J.B., JR., ET AL. 1995. Spatially explicit population models: Current forms and future uses. Ecol. Applic. 5:3–11.

FULLER, T.K. 1989. Population dynamics of wolves in north-central Minnesota. Wildl. Monogr. 105. 41 p.

FULLER, T.K. 1995. Guidelines for gray wolf management in the northern Great Lakes region. Tech. Publ. #271. Internat. Wolf Cent., Ely, MN. 19 p.

GÜRKAN, G., A.Y. ÖZGE, AND S.M. ROBINSON. 1994. Sample-path optimization in simulation. P. 247–254 *in* Proc. 1994 Winter Simul. Conf. Tew, J.D., et al. (eds.). Inst. of Elec. and Electron. Eng., Piscataway, NJ.

HAIGHT, R.G. 1995. Comparing extinction risk and economic cost in wildlife conservation planning. Ecol. Appl. 5:767–775.

HAMMERSLEY, J.M., AND D.C. HANDSCOMB. 1964. Monte Carlo methods. Methuen, London.

HAMMILL, J. 1993. Wolves in Michigan: A historical perspective. Internat. Wolf 3(1):22–23.

HEALY, K., AND L.W. SCHRUBEN. 1991. Retrospective simulation response optimization. P. 954–957 *in* Proc. 1991 Winter Simul. Conf., Nelson, B.L., W. D. Kelton, and G. M. Clark (eds.). Inst. of Elec. and Electron. Eng., Piscataway, NJ.

HOF, J.G., AND L.A. JOYCE. 1993. A mixed integer linear programming approach for spatially optimizing wildlife and timber in managed forest ecosystems. FOR. SCI. 39:816–834.

HOF, J.G., AND M.G. RAPHAEL. 1993. Some mathematical programming approaches for optimizing timber age-class distributions to meet multispecies wildlife population objectives. Can. J. For. Res. 23:828–834.

HOF, J., M. BEVERS, L. JOYCE, AND B. KENT. 1994. An integer programming approach for spatially and temporally optimizing wildlife populations. For. Sci. 40:177–191.

LAMBERSON, R.H., R. MCKELVEY, B.R. NOON, AND C. VOSS. 1992. A dynamic analysis of northern spotted owl viability in a fragmented forest landscape. Conserv. Biol. 6:505–512.

LAMBERSON, R.H., B.R. NOON, C. VOSS, AND K.S. MCKELVEY. 1994. Reserve design for territorial species: The effects of patch size and spacing on the viability of the northern spotted owl. Conserv. Biol. 8:185–195.

LICHTENBERG, E., AND D. ZILBERMAN. 1988. Efficient regulation of environmental health risks. Quart. J. Econ. 103:167–178.

MECH, L.D. 1987. Age, season, distance, direction and social aspects of wolf dispersal from a Minnesota pack. P. 55-74 *in* Patterns of dispersal among mammals and their effects on the genetic structure of populations. Chepko-Sade, B.D., and Z. Halpin, eds. University of Chicago Press, Chicago, IL.

MECH, L.D. 1995. The challenge and opportunity of recovering wolf populations. Conserv. Biol. 9:270–278.

MLADENOFF, D.J., T.A. SICKLEY, R.G. HAIGHT, AND A.P. WYDEVEN. 1995. A regional landscape analysis and prediction of favorable gray wolf habitat in the northern Great Lakes region. Conserv. Biol. 9:279–294.

MONTGOMERY, C. 1995. Economic analysis of the spatial dimensions of species preservation: the distribution of northern spotted owl habitat. FOR. SCI. 41:67–83.

MONTGOMERY, C.A., G.M. BROWN, JR., AND D.M. ADAMS. 1994. The marginal cost of species preservation: the northern spotted owl. J. Environ. Econ. Manage. 26:111–128.

NELSON, B.L., AND B.W. SCHMEISER. 1983. Variance reduction: Basic transformation. P. 255–258 *in* Proc. of 1983 Winter Simul. Conf. Inst. of Elec. and Electron. Eng., Piscataway, NJ.

RUBINSTEIN, R.Y., AND A. SHAPIRO. 1993. Discrete event systems: Sensitivity analysis and stochastic optimization by the score function method. Wiley, New York.

SALWASSER, H., S.P. MEALEY, AND K. JOHNSON. 1984. Wildlife population viability: A question of risk. Trans. of North American Wildl. and Natur. Resour. Conf. 49:421–439.

SHAFFER, M.L. 1981. Minimum population sizes for species conservation. BioScience 31:131–134.

SIMBERLOFF, D. 1988. The contribution of population and community biology to conservation science. Annu. Rev. Ecol. Syst. 19:473–511.

STACEY, P.B., AND M. TAPER. 1992. Environmental variation and the persistence of small populations. Ecol. Appl. 2:18–29.

U.S. FISH AND WILDLIFE SERVICE. 1992. Recovery plan for the eastern timber wolf. United States Fish and Wildl. Serv., St. Paul, MN.

WEINTRAUB, A., F. BARAHONA, AND R. EPSTEIN. 1994. A column generation algorithm for solving general forest planning problems with adjacency constraints. For. Sci. 40:142–161.

WILCOVE, D. 1993. Getting ahead of the extinction curve. Ecol. Appl. 3:218–220.

WYDEVEN, A.P. 1993. Wolves in Wisconsin: Recolonization under way. Internat. Wolf 3(1):18–19.

YOSHIMOTO, A., J.D. BRODIE, AND J. SESSIONS. 1994. A new heuristic to solve spatially constrained long-term harvest scheduling problems. For. Sci. 40:365–396.