

# Reserve Design to Maximize Species Persistence

Robert G. Haight · Laurel E. Travis

Received: 16 May 2006 / Accepted: 17 February 2007 / Published online: 15 March 2007  
© Springer Science + Business Media B.V. 2007

**Abstract** We develop a reserve design strategy to maximize the probability of species persistence predicted by a stochastic, individual-based, metapopulation model. Because the population model does not fit exact optimization procedures, our strategy involves deriving promising solutions from theory, obtaining promising solutions from a simulation optimization heuristic, and determining the best of the promising solutions using a multiple-comparison statistical test. We use the strategy to address a problem of allocating limited resources to new and existing reserves. The best reserve design depends on emigration, dispersal mortality, and probabilities of movement between reserves. When movement probabilities are symmetric, the best design is to expand a subset of reserves to equal size to exhaust the habitat budget. When movement probabilities are not symmetric, the best design does not expand reserves to equal size and is strongly affected by movement probabilities and emigration rates. We use commercial simulation software to obtain our results.

**Keywords** Optimization · Population viability · Reserve site selection · Resource allocation · Search heuristic · Simulation optimization · Stochastic population model

---

R. G. Haight (✉)  
USDA Forest Service, Northern Research Station,  
1992 Folwell Ave.,  
Saint Paul, MN 55108, USA  
e-mail: rhaight@fs.fed.us

L. E. Travis  
Department of Forestry,  
Virginia Polytechnic Institute and State University,  
Blacksburg, VA 24061, USA  
e-mail: ltravis@vt.edu

## 1 Introduction

Reserve site selection models typically maximize the number of species represented in selected sites subject to resource constraints and provide case-specific policy guidance for tradeoffs between conservation goals and reserve costs (see [3, 20, 21] for summaries of published studies). While reserve selection models maximize species richness, most do not assess the probabilities that species survive and thus provide no guarantee that species represented in protected sites will persist [4, 18, 25].

Optimization models for reserve design need to incorporate spatial population models, which are developed to understand and predict metapopulation dynamics, including probabilities of persistence. Spatial population models encompass many different modeling frames [2], including stochastic patch occupancy models, which model the presence/absence of species in habitat patches [17], and spatially explicit population models, which are individual-based, demographic models of local dynamics and dispersal behavior [19].

Determining reserve designs that maximize probability of persistence is difficult because spatial population models typically have nonlinear relationships and random variables that cannot readily be put into classical integer and mixed-integer programming formulations, which are the basis of most reserve selection models. Instead, tools are needed to join simulation and optimization to find good approximations of optimal reserve designs [25]. One approach is response surface analysis, in which a simple model of the probability of population persistence as a function of the quantity and quality of habitat is estimated from the results of a more complex spatial population model. The function for probability of persistence is combined with costs of habitat protection to determine cost-efficient protection

strategies [11, 12]. Another approach is simulation optimization in which the probability of persistence is estimated via stochastic simulation until a suitable approximation of the optimal reserve design is found [18]. The advantage of simulation optimization is that the spatial population model is used without modification in the search for the best set of sites. Unlike response surface analysis, no assumptions are made about the performance of a particular reserve design unless that reserve design is simulated. A disadvantage of simulation optimization is computational intensity: multiple replications or lengthy runs may be required to obtain a useful estimate of the probability of persistence for each set of sites that is evaluated.

The purpose of this paper is to present an optimization strategy for a reserve design problem in which probabilities of species persistence are predicted with a spatial population model. The problem is to decide how to spend limited resources to expand existing reserves and create new reserves to maximize the probability of species persistence. Probability of species persistence is predicted with a stochastic, individual-based, metapopulation model. The optimization strategy involves deriving promising solutions from the theory of reserve design [16], obtaining promising solutions from a simulation optimization heuristic [9], and determining the best of these promising solutions using multiple comparison procedures [1, 10]. We use only commercial software for our analysis, and we comment on its applicability to larger problems.

Many researchers have studied the effects of habitat loss and fragmentation on population persistence using spatial population models, and results strongly depend on assumptions about individual dispersal and establishment of home ranges, which link demography to landscape [6, 7, 24]. In wildlife studies, dispersal parameters such as emigration and dispersal mortality rates are expensive to estimate. We use our optimization strategy to solve a small reserve expansion problem and determine how emigration and dispersal mortality rates affect our approximations of optimal reserve design. If those approximations are insensitive to dispersal parameter values, then scarce resources need not be spent to obtain precise estimates.

## 2 Methods

### 2.1 Reserve Expansion Problem

Suppose we have a set of disjunct populations of an endangered species and a limited budget to increase habitat. By disjunct, we mean that populations live in separate habitat patches; however, individuals can move between patches. Habitat can be increased by expanding existing patches or making new patches. The site selection problem is to determine where to

locate new habitat to maximize the probability of metapopulation persistence over the management horizon. The model is formulated with the following notation:

|                      |   |
|----------------------|---|
| $i$                  | index of habitat patches  |
| $n$                  | number of patches (some may be size zero)                                       |
| $a_i$                | amount of habitat already present in patch $i$                                  |
| $b$                  | upper bound on budget   |
| $c_i$                | unit cost of increasing habitat in patch $i$                                    |
| $d_i$                | maximum amount of new habitat that may be added to patch $i$                    |
| $x_i$                | decision variable for the increase in habitat in patch $i$                      |
| $y_i$                | total amount of habitat in patch $i$ after increase; (i.e., $y_i = a_i + x_i$ ) |
| $P(y_1, \dots, y_n)$ | probability of metapopulation persistence                                       |

The optimization model is formulated as follows:

$$\text{Maximize : } P(y_1, \dots, y_n) \quad (1)$$

Subject to:

$$y_i = a_i + x_i \quad \text{for all } i \quad (2)$$

$$\sum_{i=1}^n c_i x_i \leq b \quad (3)$$

$$0 \leq x_i \leq d_i \quad \text{for all } i \quad (4)$$

The objective of the site selection problem (Eq. 1) is to maximize the probability of persistence of the metapopulation over the management horizon. The probability of persistence depends on the total amount of habitat in each patch, which is the sum of existing habitat and new habitat (Eq. 2). In our application, the units of habitat are integer values for the number of territories. Note that new patches may be added to the landscape when  $a_i = 0$  for some  $i$ . Equation 3 requires that spending for new habitat does not exceed the budget. In our application,  $d_i = \infty$  and  $c_i = 1$  for all  $i$  so the budget limits total number of territories that can be added independent of cost. Equation 4 bounds the amount of new habitat in each patch.

The solution of this site selection problem is complicated by the objective function, which is evaluated using a stochastic simulation model of the metapopulation. For given values of patch area,  $y_1, \dots, y_n$ , the probability of metapopulation persistence  $P(y_1, \dots, y_n)$  is estimated by the percent of independent replications of the stochastic simulation model in which metapopulation size is greater than zero at the end of a 100-year management horizon. Therefore, each candidate solution, characterized by increases in habitat area,  $x_1, \dots, x_n$ , must be evaluated using many replications of the stochastic simulation model.

## 2.2 Spatial Population Model

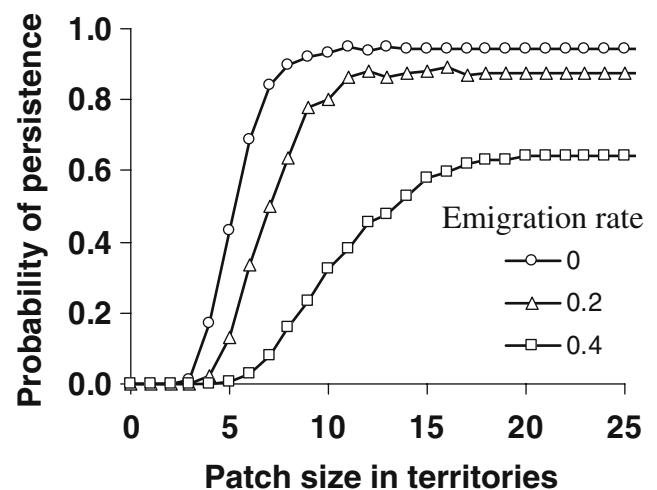
Metapopulation size and probability of persistence are simulated with a stochastic, individual-based model of a territorial species. The model is coded and implemented with Arena Professional simulation software (Rockwell Automation, Inc). Our example is for a hypothetical species living in up to six disjunct habitat patches. The size of each patch is measured by the number of potential territories. Three patches have existing habitat (nine, six, and three territories), and three patches have no existing habitat. The model is not spatially explicit at the patch level because territory shapes and locations are not included; however, the model is spatially explicit at the metapopulation level because patch locations affect dispersal between patches. The model is individually based because demographic events are computed one individual at a time. Variants of this type of model have been built for territorial species such as the Northern Spotted Owl (*Strix occidentalis caurina*) [15], gray wolf (*Canis lupus*) [13], and San Joaquin kit fox (*Vulpes macrotis mutica*) [12].

Our model is based on a demographic model of San Joaquin kit fox [12]. The model simulates birth, mortality, and dispersal of individuals in each patch on an annual cycle beginning winter. Each territory can support a pair of breeding adults and their offspring. Each individual is classified by sex and age, where juveniles are 0–12 months old and adults are >12 months old. Breeding pairs of adults produce offspring in winter, and reproduction is modeled as a two-step process. First, the success of each breeding pair is determined by comparing a random number obtained from a uniform 0–1 distribution to a success rate of 0.80. Then, the litter size of each successful pair is selected from a discrete probability distribution with mean 2.3 and range of 1–5. Gender of offspring is assigned randomly with equal probability. Mortality takes place in summer. Whether an individual dies is a Bernoulli random variable with probability depending on age. Mortality rates are 35% for juveniles and 25% for adults. All surviving juveniles disperse in autumn in search of mates and territories. Whether each juvenile is a short-distance disperser that searches for mates and territories only within the habitat patch or a long-distance disperser that leaves the patch is a Bernoulli random variable with mean equal to a given emigration rate. Short-distance dispersers settle with mates, if available. Otherwise, short-distance dispersers settle in vacant territories. In absence of either, the short-distance disperser dies. Long-distance dispersers are first evaluated for dispersal mortality using a Bernoulli trial. Survivors are assigned a new patch at random using probabilities of movement between source and destination patches. Once in new patches, long-distance dispersers occupy territories in the same process as short-distance dispersers. Adults do not disperse.

Model behavior is illustrated by plotting probability of population persistence in 100 years in a single patch as a function of patch size (Fig. 1) assuming that the initial population is at carrying capacity. With no emigration, probability of persistence reaches a maximum of 0.95 when the patch has 10 territories. With emigration, juveniles that leave the patch are assumed to die because there is only one patch. As a result, maximum probability of persistence drops and the minimum patch size for maximum persistence increases as the emigration rate increases.

Note that the plots of population persistence (Fig. 1) exhibit some stochastic variation. Our estimator of persistence is the proportion of 1,000 replications in which population size is greater than zero after 100 years. Standard errors of these estimates range up to 0.016. Increasing the number of replications to 2,500 or more would reduce the standard error to 0.010 or less and reduce or remove the wobble in the estimates of persistence. The effect of sample size on the precision of the estimator has implications for the performance of the optimization algorithm and will be discussed in the section on solution strategy.

The dispersal process in our model differs from dispersal assumptions in population models in which territories or habitat units are spatially explicit [6, 7, 24]. Those models divide the landscape into cells, each with demographic parameter values that depend on habitat quality. Emigration rate defines the probability of individuals leaving a cell, and movement of emigrants between cells is based on rules for dispersal and habitat selection. As a result, emigration from a group of cells in a patch of suitable habitat depends on patch size and shape. In contrast, we define emigration as long-distance dispersal where a fraction of juveniles leave the patch regardless of patch size and shape. Examples of



**Fig. 1** Probability of persistence in 100 years of a population living in a single patch as a function of patch size and emigration rate. Initial population size was at carrying capacity

mammal populations with long-distance dispersal include gray wolf [8] and San Joaquin kit fox [22].

### 2.3 Solution Strategy

The problem is to allocate a budget for the purchase of habitat equaling nine territories among the six habitat patches, which had 0, 9, 0, 3, 0, and 6 existing territories, respectively. The objective is to maximize the probability of metapopulation persistence in 100 years assuming that the initial population in each patch is at carrying capacity. With this setup, there are 2002 alternative reserve designs. Although we can easily simulate and compare all possible solutions, we cannot do so with enough replications to select the best solution with reasonable statistical confidence. Instead, our approach is to obtain a small set of promising solutions based on theory and simulation optimization and use additional simulations combined with a multiple comparison procedure to select the best of the set [1].

The theory of reserve design provides a solution to the problem of determining the number of reserves that maximizes the probability of population persistence when there are no existing reserves, a fixed amount of habitat can be reserved, and dispersal between patches is absent [16]. The optimal solution is to create a number of equal-sized reserves depending on the amount of habitat available. The model of McCarthy et al. [16] differs from ours in three respects: absence of dispersal, absence of existing reserves, and a specific functional form for the relationship between population persistence and reserve area. These differences prompted us to derive other promising solutions from theory.

We analyze a problem of allocating a fixed amount of habitat to patches of an existing reserve system to minimize metapopulation extinction risk (Appendix). No specific functional form is assumed for the relationship between extinction risk and patch area; the function is only assumed to be positive with strictly negative slope. Similar to the work of McCarthy et al. [16], there is no dispersal between patches. We derive the following rule for constructing a set of candidate solutions. For a problem with initial patch sizes  $(a_1, a_2, \dots, a_n)$ , select lower and upper patch sizes  $l$  and  $u$ . All patches with initial sizes between  $l$  and  $u$  are given values equal to  $u$  in the solution (i.e.,  $y_i = u$  for all  $i$  where  $l \leq a_i \leq u$ , and  $y_i = a_i$  otherwise). Here,  $l$  can be any value between 0 and the largest initial patch size, and  $u$  is selected so that the budget is exhausted. When there are multiple patches with the same initial size equal to  $l$ , there may be multiple solutions: one solution corresponding to augmenting one of the initially equal patches, one solution corresponding to augmenting two of them, and so forth. In the Appendix, we show that one of these solutions is likely to solve the necessary conditions of optimality. We also

discuss why the set of solutions may not contain the true optimum.

For our six-patch problem with  $(a_1, a_2, \dots, a_6) = (0, 9, 0, 3, 0, 6)$  and  $b = 9$ , we generated six solutions suggested by theory:

- Solution 1 = (0, 18, 0, 3, 0, 6); for  $6 < l \leq 9$  and  $u = 18$
- Solution 2 = (0, 12, 0, 3, 0, 12); for  $3 < l \leq 6$  and  $u = 12$
- Solution 3 = (0, 9, 0, 9, 0, 9); for  $0 < l \leq 3$  and  $u = 9$
- Solution 4 = (0, 9, 0, 6, 6, 6); for  $l = 0$  with 1 patch initially at 0 augmented and  $u = 6$
- Solution 5 = (0, 9, 4, 4, 4, 6); for  $l = 0$  with 2 patches initially at 0 augmented and  $u = 4$
- Solution 6 = (3, 9, 3, 3, 3, 6); for  $l = 0$  with 3 patches initially at 0 augmented and  $u = 3$

We use these six promising solutions as candidates for solving a more complex optimization problem in which individuals disperse between patches and the probability of metapopulation persistence is estimated with a stochastic simulation model, and we compare their performance to solutions obtained from a simulation optimization heuristic.

The simulation optimization heuristic uses a combination of scatter search and tabu search [9]. The heuristic is implemented in the OptQuest (OptTek Systems, Inc.) add-in to Arena Professional simulation software. The heuristic starts from a set of feasible solutions called reference points, collects data on the performance of those solutions by running multiple replications of the simulation model, and uses the information gathered to select new solutions. The process of selecting a set of solutions, testing them, and moving on to potentially better solutions, is repeated until a stopping criterion is met. The solution with the best performance to date is remembered by the algorithm at all times, and reported as the best solution when the algorithm terminates. Scatter searches use rounded linear combinations of the most promising reference points to generate new reference points. The scatter search is combined with tabu search, which keeps a list of solutions already visited by the search and restricts the search from revisiting these points. According to the software documentation, the combination of scatter and tabu searches allows generation of new reference points with inferior objective function values and prevents entrapment in local optima.

Computational intensity is an issue whenever a search heuristic is applied to stochastic problems. During the search, the objective function value of each candidate solution is computed by performing one or more replications of a stochastic simulation. The observations gathered are used to estimate the true objective function value. The variance of this estimator depends on both the level of variability inherent in the simulation model and the number of replications performed. The search algorithm may be misled if the estimated value of the objective function obtained from

stochastic simulation is far from the true value. Although increasing the number of replications to estimate the objective function value increases the likelihood that the estimate is close to the true value, the computational effort needed to perform the search increases dramatically.

To reduce computation time and maintain some of the accuracy of additional replications, we use an OptQuest feature to specify upper and lower bounds on the number of simulation replications to be performed for each solution encountered during the search. For each candidate solution, OptQuest performs the minimum number of replications and uses the resulting data to test whether the solution will likely outperform the best solution found to date. Additional replications take place only if the solution shows promise. Specifically, replications end when the 95% confidence interval about the mean performance of the solution being evaluated no longer contains the estimated best value to date. This test allows the search to focus computational effort on promising solutions. We use 30 and 3,000 as lower and upper bounds on the number of replications performed for each solution.

We use the solutions suggested by theory as starting solutions in OptQuest. The search heuristic evaluates those solutions first and uses the best as a standard for comparison. Combined with the feature allowing variable numbers of replications, computational effort is not spent on feasible solutions that are not performing as well as the best of the solutions suggested by theory. We run the search heuristic until all 2,002 of the potential reserve designs are used as reference points and evaluated with at least 30 simulation replications. Then, we use the best solution obtained with the heuristic in a statistical run-off with the solutions obtained from theory.

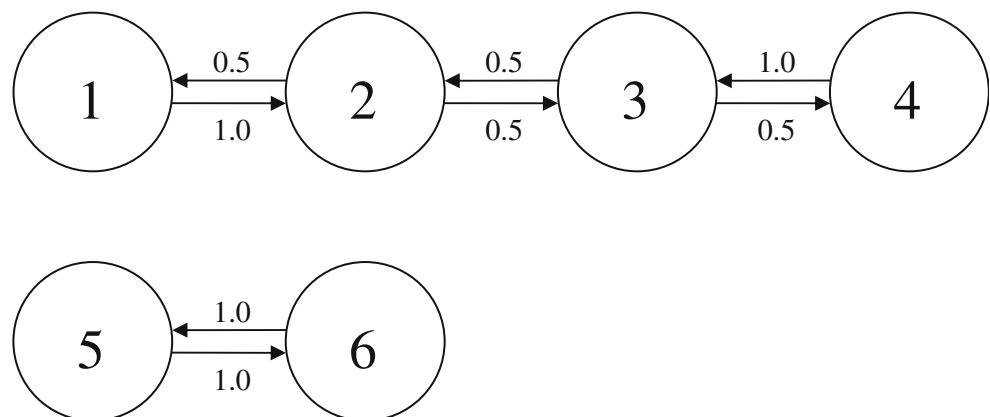
The statistical run-off involves additional simulations to select the best of the set of solutions obtained from theory and simulation optimization. We use a multiple comparison

procedure in OptQuest that accounts for the error that arises when making simultaneous inferences about differences in performance among multiple systems [10]. We stop the simulations when the probability of correct selection of the best of the set is 0.99. Occasionally, two or more solutions produce such close results that a winner cannot be picked at the 0.99 level even after 50,000 replications per solution. Then, we perform additional replications until one solution can be chosen at the 0.95 confidence level with an indifference zone parameter of 1%, which allows us to conclude at the 0.95 confidence level that the selected solution is either the best of the set or within 1% of the best.

## 2.4 Sensitivity Analysis

We use our optimization strategy to determine how the long-distance dispersal parameters affect the allocation of territories among the six habitat patches. In the first set of problems, we assume that movement probabilities are symmetric (i.e., long-distance dispersers from any given patch have equal probability of reaching any other patch). In the second set, we assume that movement probabilities are asymmetric: four patches are connected by positive dispersal probabilities in a linear array, and two patches are connected in a separate array (Fig. 2). Having two independent arrays of patches is an extreme case of asymmetry, which we chose to emphasize the differences between solution performance in the symmetric and asymmetric cases. In each set, we solve optimization problems using a range of long-distance dispersal (emigration) rates and dispersal mortality rates. For some values of the dispersal parameters, several promising solutions have probabilities of metapopulation persistence of 100% and the statistical run-off procedure cannot distinguish them. We only discuss results for cases where the best solution found has persistence less than 100%.

**Fig. 2** Patch layout with asymmetric probabilities of movement between patches



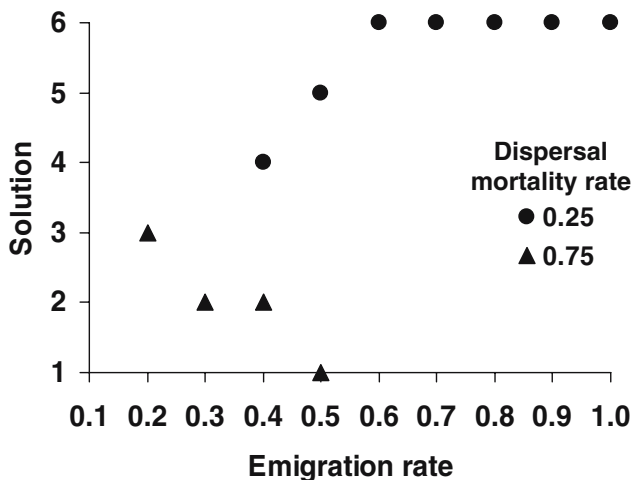
### 3 Results

#### 3.1 Symmetric Dispersal

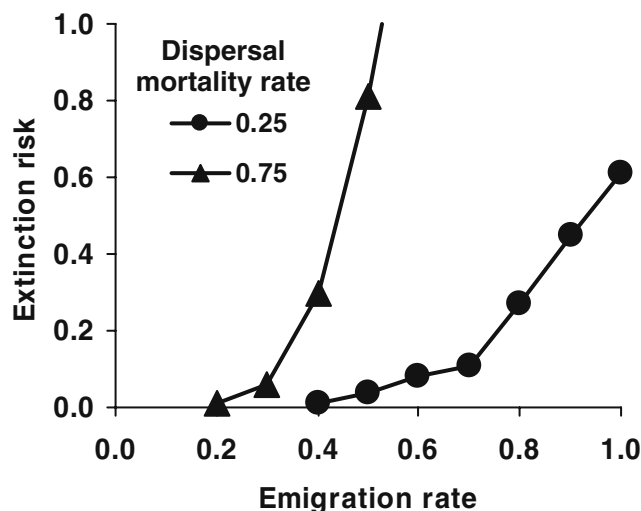
The best reserve design is always a solution suggested by theory: for each level of emigration and dispersal mortality, a subset of reserves is expanded to an equal size to exhaust the habitat budget (Fig. 3). In about half of the problems, the search heuristic converges to the best solution suggested by theory. In the other half, the solutions obtained with the search heuristic are slightly inferior. When emigration rate is low (<0.4 for dispersal mortality 0.25 and <0.2 for dispersal mortality 0.75), many reserve designs have no significant risk and we do not report results for those cases.

The best reserve design found is sensitive to changes in emigration and dispersal mortality rates (Fig. 3). When dispersal mortality is low (0.25), the best strategy for emigration rate 0.4 is a 4-patch solution in which two territories are added to three patches making them equal in size with six territories each. As emigration increases from 0.4 to 0.6, the best reserve design changes from a small number of large patches to a large number of small patches. With higher emigration, it is better to have a larger number of patches to accept the increasing numbers dispersers. The change from clumped to dispersed patches is incremental: for each increase in emigration rate, one additional patch is created.

With a high dispersal mortality rate (0.75), the best strategy found is to expand three existing reserves for emigration rate 0.2 (Fig. 3). As emigration increases from 0.2 to 0.5, the best strategy is to expand one or two existing



**Fig. 3** Best reserve design solution as a function of emigration rate and dispersal mortality rate for the case with symmetric movement probabilities. The solution numbers represent increasing numbers of patches and/or more equitable distribution of area among the patches. For our six-patch problem with initial patch sizes (0, 9, 0, 3, 0, and 6), solution 1=(0, 18, 0, 3, 0, and 6), solution 2=(0, 12, 0, 3, 0, and 12), solution 3=(0, 9, 0, 9, 0, and 9), solution 4=(0, 9, 0, 6, 6, and 6), solution 5=(0, 9, 4, 4, 4, and 6), and solution 6=(3, 9, 3, 3, 3, and 6)



**Fig. 4** Metapopulation extinction risk associated with the best reserve design as a function of emigration rate and dispersal mortality rate for the case with symmetric movement probabilities

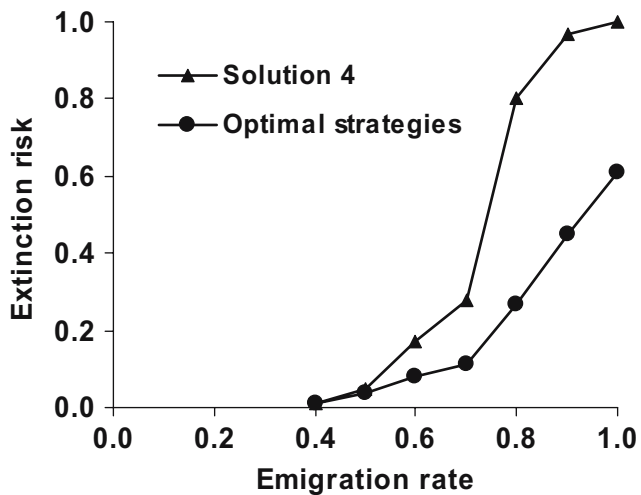
reserves. In these cases, the likelihood of dispersers reaching new patches is low and it is better to expand the carrying capacity of existing reserves to enhance the breeding success of juveniles that are not long-distance dispersers. Again, the change in strategy is incremental: for each increase in emigration rate between 0.2 and 0.5, resources are allocated to at most one less patch.

The level of extinction risk associated with the best strategies increases with emigration rate (Fig. 4) because emigrants are exposed to higher mortality rates than nonemigrants. As emigration rate increases, extinction risk increases slowly up to a threshold and then increases rapidly. For example, with a dispersal mortality rate of 0.25, extinction risk is less than 0.11 for emigration rates up to 0.7. With dispersal mortality of 0.75, extinction risk is less than 0.06 for dispersal rates up to 0.3.

The best reserve designs mitigate higher mortality associated with higher emigration rates. For example, with a long-distance dispersal mortality rate of 0.25, extinction risk of the 4-patch solution that is best for emigration rate 0.4 increases rapidly when that solution is used for higher emigration rates (Fig. 5). Compared with the performance of the 4-patch solution, risk is reduced up to 50 percentage points by creating a larger number of smaller patches.

#### 3.2 Asymmetric Dispersal

For the problem with asymmetric dispersal and a dispersal mortality rate of 0.25, extinction risk is zero for low emigration rates (<0.4), and many different reserve designs have no significant risk. Extinction risk increases for emigration rates >0.4 (Table 1), and these are the solutions that we examine. For each level of emigration, the solution obtained with the search heuristic departs from solutions



**Fig. 5** Reduction in metapopulation extinction risk associated with the best reserve design found compared with solution 4=(0, 9, 0, 6, 6, and 6) for the problem with symmetric movement probabilities and dispersal mortality of 0.25

suggested by theory and is superior in terms of minimizing extinction risk. For emigration rates of 0.4–0.7, the budget for habitat acquisition is allocated to patches 1 and 3 creating a linked series of four patches. For higher levels of emigration, the budget is allocated to expanding patches 5 and 6, creating a linked series of two patches separated from the others.

Extinction risks associated with the best reserve designs increase with emigration rate (Table 1) and are lower than minimum extinction risks obtained for problems with symmetric dispersal. The problem with asymmetric dispersal has fewer movement corridors, and focusing limited resources on expanding patches along those corridors provides greater refuge for migrating individuals and reduces metapopulation extinction risk.

### 3.3 Execution Time

The search heuristic and statistical run-off procedure were implemented on a Dell Pentium 4 laptop computer (CPU 2.4 GHz). Execution times were less than 24 h for each procedure. Execution time for the search heuristic depended on the total number of replications of the simulation model performed during the search. Solutions to problems with symmetric dispersal required 100,000–900,000 simulation replications. More replications were required to solve problems with relatively low probabilities of extinction because many solutions had close to the same level of performance. Similarly, execution time required to perform the statistical run-off of the best of the promising solutions to each problem depended on total number of simulation replications. Except in one case, the number of replications needed to obtain a statistically significant (99% confidence

level) estimate of the best of the promising solutions varied from 500 to 50,000 per solution. More replications were required with smaller differences in performance between solutions. Execution time required per replication increased as probability of extinction decreased. With high probabilities of extinction ( $>0.95$ ), 2.2 min were required per thousand replications. With low probabilities of extinction ( $<0.05$ ), execution time per thousand replications was 3.4 min.

## 4 Discussion

Reserve selection models typically focus on species representation and ignore population dynamics and likelihood of persistence [4, 18, 25]. To overcome this weakness, we develop an optimization strategy for problems that involve simulating population dynamics. We are aware of only one other attempt to address this type of problem. Moilanen and Cabeza [18] use a genetic search algorithm combined with a stochastic patch occupancy model to determine sites to protect to maximize the likelihood of metapopulation persistence. Our solution strategy addresses a reserve design problem in which probability of species persistence is estimated with a stochastic, individual-based model of the metapopulation. The strategy involves deriving promising solutions from the theory of reserve design, obtaining promising solutions from a search heuristic, and determining the best of these promising solutions using a multiple comparison statistical test.

The solution strategy attempts to overcome several difficult features of optimization problems that involve stochastic population dynamics. Most population models contain nonlinear relationships and random variables that are difficult to program into exact optimization procedures. As a result, our strategy involves simulation optimization in which the probabilities of persistence of candidate solutions are estimated via stochastic simulation until a suitable

**Table 1** Best reserve design found as a function of emigration rate for the case with asymmetric probabilities of movement between patches and a dispersal mortality rate of 0.25

| Emigration rate | Territories per patch |   |   |   |   |   | Extinction risk |
|-----------------|-----------------------|---|---|---|---|---|-----------------|
|                 | 1                     | 2 | 3 | 4 | 5 | 6 |                 |
| 0.4             | 1                     | 9 | 8 | 3 | 0 | 6 | 0.01            |
| 0.5             | 3                     | 9 | 6 | 3 | 0 | 6 | 0.01            |
| 0.6             | 4                     | 9 | 5 | 3 | 0 | 6 | 0.02            |
| 0.7             | 4                     | 9 | 5 | 3 | 0 | 6 | 0.04            |
| 0.8             | 0                     | 9 | 0 | 3 | 7 | 8 | 0.06            |
| 0.9             | 0                     | 9 | 0 | 3 | 7 | 8 | 0.08            |
| 1.0             | 0                     | 9 | 0 | 3 | 7 | 8 | 0.13            |

The initial numbers of territories per patch were 0, 9, 0, 3, 0, and 6.

approximation of the optimal reserve design is found. With simulation optimization, computational intensity is an issue because the estimate of the objective function value of each candidate solution must be computed with one or more replications of the stochastic population model. We address this issue by using results from the theory of reserve design as our first approximation of the optimal solution. Because solutions suggested by theory often perform well, many alternative solutions can be eliminated with few replications of the stochastic population model because they are clearly inferior. With simulation optimization, there is no guarantee that the best solution found is optimal because objective function values are estimated with error. We address this issue by composing a set of good candidate solutions and using a statistical run-off to find the best of the set within a given level of confidence. The run-off procedure can determine whether solutions found via simulation optimization are better than solutions based on theory or practical experience.

Our application demonstrates that reserve design problems can be addressed using commercial simulation optimization software; however, computation time is an issue. Our optimization problem with about 2,000 candidate solutions required up to 24 h of computation time. Computation time is related to the number of simulation replications performed during the search heuristic and statistical runoff. We used a very large number of simulation replications because we wanted to evaluate each candidate solution with the search heuristic and have a high degree of confidence in the best solution selected by the run-off procedure. Relaxing these design parameters reduces the required number of simulation replications and computation time dramatically.

The best reserve design depends on dispersal parameters including emigration rate, dispersal mortality rate, and the matrix of probabilities of movement between patches. When probabilities of movement between patches are equal, the best design is always a solution suggested by theory: a subset of reserves is expanded to equal size to exhaust the habitat budget. Which subset of reserves is expanded is strongly affected by the relative rates of emigration and dispersal mortality (Fig. 3). When both emigration and dispersal mortality rates are large, expanding a large patch is superior to creating many small patches because juveniles that do not emigrate have higher breeding success within a single large patch. This result is consistent with simulation studies in which reducing habitat fragmentation increases population persistence because fewer individuals disperse into the matrix (or low-quality habitat) where they are exposed to high mortality and low reproduction rates [6, 7, 24]. When emigration is high and dispersal mortality low, creating many small patches is superior to expanding a single large patch. In this case,

creating small patches increases the likelihood that emigrating juveniles find breeding habitat, which increases population persistence more than expanding a single large patch. This finding is consistent with results from patch occupancy models of metapopulations in which increasing dispersal success increases population persistence through recolonization of local extinctions [14]. When emigration is low (<40%), expanding a small number of existing patches to equal size is best. This strategy is best for a wide range of dispersal mortality rates, indicating that precise estimates of these parameters are not necessary in these ranges. Finally, when probabilities of movement between patches are asymmetric, the best reserve design is not a solution suggested by theory and strongly depends on the movement probabilities and emigration rates (Table 1).

Like most reserve design models, ours assumes that reserve expansion takes place all at once and ignores the reality that management decisions are sequential and depend on the states of the population observed over time. Researchers are beginning to address sequential reserve design problems to optimize conservation objectives subject to budget constraints and uncertainties about population dynamics and site degradation and loss [5, 23]. The idea is to develop an adaptive decision rule for conservation action depending on sites already protected, those currently available, the state of the population, and available funding. Optimal rules for small problems with up to seven sites can be obtained with stochastic dynamic programming and a stochastic patch occupancy model [23]. A challenge is to combine the tools of simulation optimization with stochastic, individual-based, metapopulation models to develop and evaluate adaptive decision rules.

Our application involves a relatively simple, individual-based simulation model for a hypothetical species. The model does not include important processes that may affect real populations in fragmented landscapes, such as movement of dispersers in the matrix, breeding success in habitat of various qualities, or density-dependent survival and reproduction rates. Individual-based models that incorporate these processes have been developed to determine the impacts of habitat amount and fragmentation on population persistence [6, 7, 24]. A challenge is to determine which processes governing species dynamics have significant impacts on optimal reserve design.

Determining a reserve design to maximize population persistence is a challenge when complex individual-based models that require lots of execution time for a single replication are combined with our optimization strategy. The execution time of our strategy depends in large part on the number of replications of the population model, and there are a number of ways to reduce replications. One way is to set up the reserve design problem with a small number of decision variables that can each have only a few possible values. Using



a small decision space reduces the number of potential solutions that need to be evaluated with the simulation optimization heuristic. A further reduction in number of replications can be obtained by reducing the degree of confidence needed to eliminate a potential solution from further consideration during simulation optimization.

Another way to reduce the number of replications of the population model is to reduce the confidence level and increase the indifference zone for selecting the best of the promising solutions during the run-off procedure following simulation optimization. The indifference zone is the maximum difference allowed between the estimated population persistence of the selected solution and the best of the set. The confidence level is the likelihood of selecting a solution whose difference from optimal performance is within the indifference zone. Increasing the indifference zone is the most direct and efficient way to reduce execution time, especially when detecting a small difference in performance is not very important given modeling error.

A final and extreme way to reduce execution time is to eliminate simulation optimization altogether. In this case, we could derive a set of promising solutions from theory and expert opinion and determine the best from this promising set using the multiple-comparison test. Although this approach avoids the execution time of the simulation optimization heuristic, the set of promising solutions does not include possibly superior solutions that could be found by the heuristic. If we only want to compare a small set of alternative solutions with a standard, the best strategy is to allocate computing resources to the multiple-comparison test and adjust its confidence level and indifference zone parameters accordingly.

## Appendix

This Appendix gives Kuhn–Tucker necessary conditions for optimality for a special case of the reserve design problem (Eqs. 1–4) in which there is no dispersal between patches. In this version of the problem, the objective is to minimize extinction risk, which depends on the sizes of habitat patches. We do not specify a function form for the extinction risk function and only require that the risk function be strictly positive with strictly negative slope. Given these assumptions, it is possible to derive a rule to obtain solutions that satisfy all but one of the Kuhn–Tucker conditions. Because we can obtain many solutions that satisfy this rule, we describe a heuristic to find solutions among them that are likely to satisfy the final Kuhn–Tucker condition. We call the solutions found in this way “solutions suggested by theory,” and we use them to

initialize the simulation optimization of more general problems with dispersal.

Suppose there exists an extinction risk function  $p$  such that  $1 - P(y_1, y_2, \dots, y_n) = \prod_{i=1}^n p(y_i)$ . Furthermore,  $p(\cdot) > 0$  and  $p'(\cdot)$  exists and is strictly negative. These assumptions will be satisfied when the patches are independent and identical and  $p(y_i)$  is the probability of extinction in patch  $i$  over the time horizon. Given independence of the fate of the individual patches, the product form of the objective function gives the probability of metapopulation extinction. Note that the absence of dispersal is required. Under these assumptions and when  $d_i = \infty$  and  $c_i = 1$ , the reserve design problem (Eqs. 1–4) is equivalent to:

$$\min \sum_{i=1}^n \ln(p(y_i)) \quad (\text{A1})$$

subject to:

$$\sum_{i=1}^n y_i - B \leq 0 \quad (\text{A2})$$

$$a_i - y_i \leq 0 \quad i = 1, \dots, n \quad (\text{A3})$$

Eq. A1 comes from applying a monotonic transformation  $\ln(\cdot)$  to the objective  $\min \prod_{i=1}^n p(y_i)$ . Eq. A2 comes from substituting Eq. 2 into Eq. 3 and defining  $B = b + \sum_{i=1}^n a_i$ . Eq. A3 comes from substituting Eq. 2 into Eq. 4 with  $d_i = \infty$ .

Defining  $\mathbf{y} = (y_1, \dots, y_n)$  and  $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_n)$ , the Lagrangian function for Eqs. A1–A3 is  $L(\mathbf{y}, \lambda, \boldsymbol{\gamma}) = \sum_{i=1}^n \ln(p(y_i)) + \lambda(\sum_{i=1}^n y_i - B) + \sum_{i=1}^n \gamma_i(a_i - y_i)$ , where  $\lambda$  and  $\boldsymbol{\gamma}$  are the dual variables corresponding to the budget constraints and the existing reserve size constraints, respectively. Because Eqs. A2 and A3 are linear, the Kuhn–Tucker constraint qualification holds, and the Kuhn–Tucker conditions for Eqs. A1–A3 are necessary conditions for optimality. These conditions are:

$$\frac{p'(y_i)}{p(y_i)} + \lambda - \gamma_i = 0 \quad i = 1, \dots, n \quad (\text{A4})$$

$$\left( \sum_{i=1}^n y_i - B \right) \lambda = 0 \quad (\text{A5})$$

$$(a_i - y_i) \gamma_i = 0 \quad i = 1, \dots, n \quad (\text{A6})$$

$$\sum_{i=1}^n y_i - B \leq 0 \quad (\text{A7})$$

$$a_i - y_i \leq 0 \quad i = 1, \dots, n \quad (\text{A8})$$

$$\lambda \geq 0 \tag{A9}$$

$$\gamma_i \geq 0 \quad i = 1, \dots, n \tag{A10}$$

Eq. A4 comes from setting the partial derivatives of the Lagrangian function to 0. Eqs. A5 and A6 comprise the complementary slackness condition; each constraint must either be binding at optimum or its corresponding dual variable must be zero. Eqs. A7 and A8 give primal feasibility, and Eqs. A9 and A10 state that the duals must be positive at optimum.

We cannot derive sets of  $(\mathbf{y}, \lambda, \gamma)$  that satisfy Eqs. A4–A10 without specifying the functional form of  $p(\cdot)$ . We do not specify the form of the extinction risk function because we use a stochastic simulation model to determine extinction risk. Nevertheless, we can derive  $(\mathbf{y}, \lambda, \gamma)$  that satisfy Eqs. A4–A9 and examine them to see if they are likely to satisfy Eq. A10 and be Kuhn–Tucker points.

To find a solution  $(\mathbf{y}, \lambda, \gamma)$  to Eqs. A4–A9, select an arbitrary set of indices  $A \in \{1, 2, \dots, n\}$ . These are indices corresponding to the patches whose size will be augmented. Let  $I$  be the set of indices corresponding to patches that are left at their initial value. Now consider the point  $\mathbf{y}$  where:

$$\begin{aligned} y_i &= a_i & i \in I \\ y_j &= \frac{B - \sum_{i \in I} a_i}{|A|} & j \in A \end{aligned} \tag{A11}$$

where  $|A|$  is the number of elements in  $A$ . Notice that the above solution corresponds to leaving the patches with subscripts in  $I$  equal to their initial values, and dividing the remaining resources among the patches with subscripts in  $A$ . The numerator in Eq. A11 represents the resources remaining after allocation to the patches in set  $I$ . The set  $A$  must be selected so that:

$$\frac{B - \sum_{i \in I} a_i}{|A|} > a_j \quad \forall j \in A \tag{A12}$$

If this is not the case, the set of augmented patches  $A$  is simply too large relative to the budget  $B$ , and dividing the resources among the set does not suffice to bring the common resulting size up to the largest initial size of the patches corresponding to the set  $A$ . In this case, removing one or more indices from  $A$  will suffice to generate a set  $A$  such that Eq. A12 is satisfied. Note that there will always be some sets  $A$  that satisfy Eq. A12 because any set of size 1 will work.

**Theorem A1** *Given the problem described by Eq. A1–A3, a set of indices  $A$  satisfying Eq. A12, and a solution  $\mathbf{y}$  defined by Eq. A11, there exist  $\lambda$  and  $\gamma$  such that Eq. A4–A9 are satisfied.*

*Proof*

$$\sum_{k \in I \cup A} y_k = \sum_{i \in I} a_i + |A| \left( \frac{B - \sum_{i \in I} a_i}{|A|} \right) = B$$

so Eqs. A5 and A7 are satisfied. Equation A8 follows from Eq. A12. Note that  $p'(y_k)/p'(y_l) < 0$  for all  $k \in I \cup A$  because we assumed that  $p'(\bullet) > 0$ , and  $p'(\bullet) < 0$ . Furthermore,  $p'(y_j)/p(y_j)$  is equal for all  $j \in A$  because  $y_j$  is equal for all  $j \in A$  by Eq. A11. Define

$$\begin{aligned} \lambda &= \frac{-p'(y_j)}{p(y_j)} & j \in A \\ \gamma_j &= 0 & j \in A \\ \gamma_i &= \frac{p'(y_i)}{p(y_i)} + \lambda & i \in I \end{aligned}$$

With these definitions, Eqs. A4, A6, and A9 hold. □

Using Eq. A11, we can find solutions  $(\mathbf{y}, \lambda, \gamma)$  that satisfy necessary conditions (Eqs. A4–A9) by selecting a subset of reserves and expanding them to equal size to exhaust the habitat area budget. There are many ways to select the subset of reserves to be expanded and hence a large number of solutions satisfying Eqs. A4–A9. To reduce the number of solutions evaluated, we order the reserves according to initial patch size and augment a subset of consecutive reserves from this ranking. Specifically, we set upper ( $u$ ) and lower ( $l$ ) bounds on initial patch sizes, and augment those patches between  $l$  and  $u$  until they are size  $u$  (see body of paper). By varying  $l$  and setting  $u$  to exhaust budget, we construct a set of “solutions suggested by theory,” all of which satisfy Eq. A11 and hence Eqs. A4–A9. Setting  $u$  so as not to overextend the budget guarantees that Eq. A12 will be satisfied.

The focus on a subset of consecutive reserves is heuristic and motivated by the likelihood (and experimental observation) that  $\ln(p(y))$  will have only one steep area, where marginal returns to reserve expansion are greatest. The location of this steep area will be a function of problem parameter values, so we use all possible lower bounds  $l$  to generate our set of promising points. One or more of these points is likely to have a good match between the augmented patch sizes and the steep area of  $\ln(p(y))$  and hence satisfy:

$$\frac{p'(y_i)}{p(y_i)} > \frac{p'(y_j)}{p(y_j)} \quad \forall i \in I, j \in A \tag{A13}$$

Intuitively, when  $p'(y_j)/p(y_j)$  is large in absolute value (highly negative), the marginal impact on the objective function of augmenting patch  $j$  is large. If the point is selected so that the marginal impact of augmenting the chosen patches is larger than those left at their initial

values, then the final Kuhn–Tucker condition, Eq. A10, will be satisfied.

It is important to note that, without additional assumptions on the function form of  $p(\cdot)$ , we cannot prove that the solutions suggested by theory contain all possible solutions to the Kuhn–Tucker Eqs. A4–A10. Furthermore, our application includes dispersal of individuals between habitat patches, which violates one of the assumptions on which Theorem A1 is based. Therefore, although the set of solutions suggested by theory is a convenient and promising place to start the simulation optimization, there is no guarantee that the true optimal solution lies within this set.

## References

- Boesel, J., Nelson, B. L., & Kim, S.-H. (2003). Using ranking and selection to “clean up” after simulation optimization. *Operations Research*, *51*, 814–825.
- Burgman, M. A., Lindenmayer, D. B., & Elith, J. (2005). Managing landscapes for conservation under uncertainty. *Ecology*, *86*, 2007–2017.
- Cabeza, M., & Moilanen, A. (2001). Design of reserve networks and the persistence of biodiversity. *Trends in Ecology and Evolution*, *16*, 242–248.
- Cabeza, M., & Moilanen, A. (2003). Site-selection algorithms and habitat loss. *Conservation Biology*, *17*, 1402–1413.
- Costello, C., & Polasky, S. (2004). Dynamic reserve site selection. *Resource and Energy Economics*, *26*, 157–174.
- Fahrig, L. (2001). How much habitat is enough? *Biological Conservation*, *100*, 65–74.
- Flather, C. H., & Bevers, M. (2002). Patchy reaction–diffusion and population abundance: the relative importance of habitat amount and arrangement. *The American Naturalist*, *159*, 40–56.
- Gese, E. M., & Mech, L. D. (1991). Dispersal of wolves (*Canis lupus*) in northeastern Minnesota, 1969–1989. *Canadian Journal of Zoology*, *69*, 2946–2955.
- Glover, F., & Laguna, M. (1997). Tabu Search (408 p). Kluwer Academic Publishers.
- Goldman, D., & Nelson, B. L. (1998). Statistical screening, selection, and multiple comparison procedures in computer simulation. In: D. J. Medeiros, E. F. Watson, J. S. Carson, & M. S. Manivannan (eds.), *Proceedings of the 1998 Winter Simulation Conference* (pp. 159–166). Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.
- Haight, R. G., Cypher, B., Kelly, P. A., Phillips, S., Possingham, H. P., Ralls, K., et al. (2002). Optimizing habitat protection using demographic models of population viability. *Conservation Biology*, *16*, 1386–1397.
- Haight, R. G., Cypher, B., Kelly, P. A., Phillips, S., Ralls, K., & Possingham, H. P. (2004). Optimizing reserve expansion for disjunct populations of San Joaquin kit fox. *Biological Conservation*, *117*, 61–72.
- Haight, R. G., Mladenoff, D. J., & Wydeven, A. P. (1998). Modeling disjunct gray wolf populations in semi-wild landscapes. *Conservation Biology*, *12*, 879–888.
- Hanski, I. (1994). A practical model of metapopulation dynamics. *Journal of Animal Ecology*, *63*, 151–162.
- Lamberson, R. H., Noon, B. R., Voss, C., & McKelvey, K. S. (1994). Reserve design for territorial species: the effects of patch size and spacing on the viability of the Northern Spotted Owl. *Conservation Biology*, *8*, 185–195.
- McCarthy, M. A., Thompson, C. J., & Possingham, H. P. (2005). Theory for designing nature reserves for single species. *The American Naturalist*, *165*, 250–257.
- Moilanen, A. (2004). SPOMSIM: software for stochastic patch occupancy models of metapopulation dynamics. *Ecological Modelling*, *179*, 533–550.
- Moilanen, A., & Cabeza, M. (2002). Single-species dynamic site selection. *Ecological Applications*, *12*, 913–926.
- Possingham, H. P., & Davies, I. (1995). ALEX: a model for the viability analysis of spatially structured populations. *Biological Conservation*, *73*, 143–150.
- ReVelle, C. S., Williams, J. C., & Boland, J. J. (2002). Counterpart models in facility location science and reserve selection science. *Environmental Modeling and Assessment*, *7*, 71–80.
- Rodrigues, A. S. L., & Gaston, K. J. (2002). Optimisation in reserve selection procedures—why not? *Biological Conservation*, *107*, 123–129.
- Schwartz, M. K., Ralls, K., Williams, D. F., Cypher, B. L., Pilgrim, K. L., & Fleischer, R. C. (2005). Gene flow among San Joaquin kit fox populations in a severely changed ecosystem. *Conservation Genetics*, *6*, 25–37.
- Westphal, M. I., Pickett, M., Getz, W. M., & Possingham, H. P. (2003). The use of stochastic dynamic programming in optimal landscape reconstruction for metapopulations. *Ecological Applications*, *13*, 543–555.
- Wiegand, T., Revilla, E., & Moloney, K. A. (2005). Effects of habitat loss and fragmentation on population dynamics. *Conservation Biology*, *19*, 108–121.
- Williams, J. C., ReVelle, C. S., & Levin, S. A. (2005). Spatial attributes and reserve design models: a review. *Environmental Modeling and Assessment*, *10*, 163–181.