SINGLE-SPECIES DYNAMIC SITE SELECTION

ATTE MOILANEN 1 AND MAR CABEZA

Department of Ecology and Systematics, Division of Population Biology,
P.O. Box 17 (Arkadiankatu 7), FIN-00014 University of Helsinki, Finland

Abstract. Methods for designing regional reserve networks mostly concentrate on providing maximal representation of species occurring in the region. Representation-based methods, however, typically consider a static snapshot of species incidences, and the spatial dynamics of the species are ignored. It has been empirically demonstrated that reserves designed using representation do not guarantee another important goal of reserve design: long-term persistence. The question studied here is the following: Which subset of sites do you select to maximize the long-term persistence of a species living in a metapopulation, given that each site has a cost and the amount of resource (e.g., money) available is limited? We present an optimization method, which uses a combination of evolutionary optimization (a genetic algorithm) and local search to find the optimal selection of sites. The quality of each candidate solution is evaluated using a spatially realistic metapopulation model, the incidence function model. The proposed method is applied to a metapopulation of the false heath fritillary butterfly, an endangered species in Finland. With this data set, the proposed estimation method produces intuitively acceptable and consistent results within minutes of computation time. Sites favored by the algorithm are located in three patch clusters, and they tend to be inexpensive and initially occupied. Expensive and/or very isolated patches are rarely selected into the optimal site selection.

Key words: economic constraint; genetic algorithm; incidence function model; local search; long-term persistence; Melitaea diamina; metapopulation; objective function; reserve network design; site selection algorithm.

INTRODUCTION

The design of regional reserve networks has two strategic goals: to efficiently represent the full spectrum of regional diversity within a system of protected areas, and to ensure the long-term persistence of biological diversity (Frankel and Soulé 1981). In general, conservation planning deals with representation and persistence of the species in separate contexts. Most of the current reserve design methods are based on analysis of species richness or endemism. They consider a single static snapshot of incidence patterns with the aim of finding a set of sites that complement each other well and when combined produce a complete representation of species. These methods aim at solutions that are efficient (sensu Pressey and Nicholls 1989) in the sense that a defined minimum level of representation is achieved with minimum cost or area (see Williams [2000] for review). Representation-based approaches typically ignore environmental change and ecological dynamics (Cabeza and Moilanen 2001). There are three studies of the long-term persistence of species in reserve networks designed with representation-based methods (see Appendices A and B). Margules et al. (1994) empirically evaluated a method of identifying a minimum set of sites that would capture rare or uncommon species. They compared the minimum set with the current occupancy after 11 years and found that the original selection of sites was inadequate for preserving all species. They continued by identifying a minimum set of sites that included more than one original population of each species, and again not all of the species would have persisted for 11 years. Similar conclusions were found also by Virolainen et al. (1999) and Rodrigues et al. (2000).

The other approach to reserve design, population viability analysis (PVA; Soulé 1987, Possingham et al. 2000), focuses mainly on single endangered species. Population viability analysis analyses the dynamics of populations, giving recommendations about minimum population levels and conservation areas required to ensure persistence. Population viability analysis has been mainly carried out on the so-called “charismatic megavertebrates,” such as the grizzly bear (Schaffer 1981) or the Spotted Owl (Marcot and Holthausen 1987), but lately, metapopulation dynamics based PVA has also been applied, for example, to birds (Akçakaya and Atwood 1997), marsupials (Possingham et al. 1994, Lindenmayer and Possingham 1995), and butterflies (Murphy et al. 1990). However, none of these studies are directly linked to site selection.

From comparisons between representation- and persistence-oriented approaches, it has become apparent that these two alternatives, when considered independently, can lead to divergent conservation strategies (Pimm and Lawton 1998). Although some representation methods have tried to implicitly include some
persistence criteria (inclusion of PVA in Conservation Options and Decision Analysis [CODA], Bedward et al. 1992; Schonewald-Cox criteria for viability of mammalian predators, Kiester et al. 1996; adjacency rule, sites close together to maximize persistence, Nicholls and Margules 1993), these approaches typically ignore the spatial dynamics of populations. There is no guarantee that a method based on representation will produce a site selection that guarantees the long-term persistence of the species. A species may be represented by one or a few small sites (populations), but persistence may require sets of larger and/or better interconnected sites.

Lately, the need to incorporate spatiotemporal dynamics into site selection procedures has been emphasized (Pressey et al. 1993, Pressey 1994, Andelman et al. 2000, Cabeza and Moilanen 2001). Metapopulation models have been suggested as appropriate tools for reserve design if the goal is the long-term persistence of endangered species in fragmented habitats (Burgman et al. 1993, Hanski 1998). In this study we address the problem of site selection for a metapopulation species. We present an optimization method that searches for the optimal subset of sites that maximizes the persistence of the focal species living in a metapopulation with explicit spatial structure. We adopt an economic constraint approach, that is, we assume that site selection is restricted by some resource (e.g., money) and that each patch has a given cost. An alternative would have been a biodiversity constraint approach, where the aim of optimization is to achieve with minimum cost an acceptably low level of extinction probability (akin to PVA methods; Lindenmayer and Possingham 1995). This approach was discarded because of substantial difficulties in the parameterization of regional stochasticity (Hanski 1991), which leads to substantial uncertainty in the estimation of the metapopulation extinction probability. Technically, the optimal subset of sites is sought with an optimization method that combines a genetic algorithm (Holland 1975) with local search. The proposed method is applied to a metapopulation of the false heath fritillary butterfly (Melitaea diamina), an endangered species in Finland.

METHODS

A classical metapopulation consists of a number of local populations living in a network of spatially distinct habitat patches (Levins 1969, see Hanski [1998] for a review). All local populations in a classical metapopulation are in risk of local extinction, and thus the metapopulation persists regionally in a balance of local extinctions and recolonization of empty habitat. The aim of this study is to find a subset of patches that gives maximal persistence for a metapopulation, given that not all patches can be selected due to a resource constraint. This is not a trivial optimization problem. The number of different selections that can be made from a patch network grows very quickly as a function of patch number, and evaluating all possible selections is quite impossible. Furthermore, the value of a site selection has to be evaluated using a stochastic metapopulation model, which further complicates the optimization problem.

Here, we first define the quantity to be optimized, the objective function. Then the proposed optimization approach is described in detail. The problem of site selection is exceptionally well suited for optimization with a binary genetic algorithm (GA; Holland 1975, Goldberg 1989), because the selected subset of patches can be naturally represented by a string of ones and zeros, which is the coding used by binary GA’s. (A one indicates a selected site and a zero indicates a nonselected site.) We combine the GA with a local search procedure that greatly accelerates optimization speed. Next we briefly describe the metapopulation model we use, which is the incidence function model of Hanski (1994).

The objective function

The goal of optimization here is to find an optimal subset ($S_{\text{opt}}$) of patches that gives the metapopulation highest persistence (lowest extinction risk) for a given number of years ($T$). The solution $S_{\text{opt}}$ has to satisfy the condition $\text{cost}(S_{\text{opt}}) \leq R$, where $R$ is the maximum amount of available resources, be it money or something else. The extinction risk, $F(S)$, of a selection of patches ($S$) is evaluated from a number ($N$) of stochastic simulation runs that start from a given initial state ($O_0$), which would most often be the observed pattern of patch occupancy. The simplest quantity to be minimized would be the proportion of replicate simulations that go extinct before time $T$. There are two problems with this measure. First, if replicates only rarely go extinct, a very large number of simulation runs is needed to evaluate the extinction probability of the metapopulation reliably, which will slow down the optimization considerably. Second, the simple measure is unable to distinguish between solutions that are always persistent. Here a more efficient objective function is used. We define that the value of site selection $S$ for time period $T$, given observed initial state $O_0$ and metapopulation model $M$ with parameter set $\theta$ as

$$F(S) = F(S, O_0, M, \theta, T) = \frac{1}{NT} \sum_{n=1}^{N} \sum_{t=1}^{T} P[\phi_{n,t+1} | X_n]$$

in which $X_n$ is the simulated patch occupancy pattern at time $t$ in replicate simulation $n$, and $\phi_{n,t+1}$ denotes the state of extinction at time $t + 1$. $F(S)$ denotes the value of the objective function; it is the average one-step global extinction probability of the metapopulation calculated over the simulation runs. Note that this formulation assumes that nonselected habitat is lost for the species. For our sample butterfly metapopulation (see Materials), this assumption is justified, as unman-
aged sites become unsuitable for the species due to overgrowth in a relatively short time.

The one-step global extinction probability of the metapopulation, \( \mu(X) \), is calculated as the probability of simultaneous extinction of all local populations:

\[
\mu(X) = P[d_{n+1} | X] = \prod_{i=1}^{m} \left[ 1 - C_i(t) \right] \text{ if } p_i(t) = 0 \\
\left[ 1 - C_i(t)E_i \right] \text{ if } p_i(t) = 1
\]

in which \( p_i(t) \) is the occupancy state of patch \( i \) at time \( t \), \( C_i(t) \) is the probability of patch \( i \) being colonized, \( E_i \) is the probability of patch \( i \) going extinct independently of colonization from other patches (the intrinsic extinction probability), and \( 1 - C_i(t)E_i \) is the extinction probability of patch \( i \) when considering the rescue effect (Brown and Kodric-Brown 1977).

Eq. 1 is related to, but not always identical to, the extinction risk of the metapopulation. It has the following properties: \( F(S) \rightarrow 1 \) when the metapopulation goes extinct almost immediately; \( F(S) \rightarrow 0 \) when the metapopulation is highly persistent. When the metapopulation goes extinct in a simulation at time \( t \), \( F(S) \) is increased by \( (T - t)/(NT) \), and thus the value of \( F(S) \) is always greater than the proportion of time the metapopulation is extinct in the simulations. Importantly, Eq. 1 is able to differentiate between two solutions that persist until the end of all simulation runs; \( F(S) \) is smaller for the solution that is more persistent. We now describe the optimization method used to find the optimal site selection. A nontechnically oriented reader may skip these details without breaking continuity.

**The genetic algorithm (GA)**

Evolutionary optimization, including genetic algorithms (GA), uses principles of evolution to solve optimization problems. These algorithms operate on a population of solutions (here site selections), which are often called individuals or chromosomes. Each individual is assigned a fitness value based on the value of the objective function for that individual. The fitness of an individual determines its probability of taking part in reproduction; individuals with high fitness reproduce on average more often than individuals with low fitness [here low \( F(S) \) translates to high fitness]. This fitness-proportionate selection guides the search toward promising sections of the search space. An evolutionary optimization algorithm typically proceeds generation by generation, with individuals of generation \( n \) producing the individuals of generation \( n + 1 \). In addition to selection, so-called variation operators are needed. The two most common variation operators are recombination (often called crossover) and mutation. The recombination operator combines information in two (or more) parent individuals into two child individuals that exhibit properties of both parents. Mutation makes one or more random changes in the individual. Variants of evolutionary optimization algorithms have been applied to a wide range of optimization problems. The review by Bäck et al. (1997) provides a good starting point for research into evolutionary algorithms.

In this study, the optimization was done using a combination of a GA and local search (see Moilanen [2001] for an introduction and review). The GA operates as a global search engine, which enables the algorithm to escape local optima while giving its own direct contribution to the search process. Local search provides quick improvement to the value of the objective function when possible. The following algorithm implements a GA–local search combination algorithm. Parts of the algorithm will be discussed in detail.

**The genetic algorithm–local search hybrid algorithm**

Define \( P^{(n)} \) as the population in generation \( n \), \( G \) as the number of generations in the evolutionary optimization algorithm, \( K_{\text{max}} \) as the population size, and \( K^0 \) as the population size in generation \( n \).

1. Create the initial population \( P^{(1)} \).
2. FOR \( n = 1, 2, \ldots, G \) DO
   1. Evaluate the fitness of individuals in \( P^{(n)} \) (includes local search phase).
   2. Copy the individual with the highest fitness in \( P^{(n)} \) into \( P^{(n+1)} \). Set \( K^{(n+1)} = 1 \). (This is the elitist strategy.)
   3. WHILE \( K^{(n+1)} < K_{\text{max}} \) DO
      1. Select two parents, \( p_1 \) and \( p_2 \), from \( P^{(n)} \) using 20% truncation selection.
      2. Produce two children, \( c_1 \) and \( c_2 \), by uniform crossover between \( p_1 \) and \( p_2 \).
      3. Mutate \( c_1 \) and \( c_2 \) with probability \( p_m \).
      4. Add \( c_1 \) and \( c_2 \) to \( P^{(n+1)} \). Set \( K^{(n+1)} = K^{(n+1)} + 2 \).
   4. Terminate if current best solution is good enough, or after \( G \) generations.

The standard run parameters used in this study were \( K_{\text{max}} = 100 \), \( G = 25 \), and \( p_m = 1/L \), where \( L \) is the number of patches in the patch network. Fifty replicate simulations were used in fitness evaluation (\( N = 50 \)), and the time period of interest (\( T \)) was 25 yr.

The GA variant used in this study applies standard GA techniques. The algorithm starts with initialization, in which solutions are generated by randomly including and excluding sites with probability 0.5. Next the algorithm proceeds generation by generation. In each generation \( n \), the population for the next generation \( n + 1 \) is generated based on the fitnesses of individuals in the current population. In recombination we used truncation selection (Mühlenbein and Schlierkamp-Voosen 1993), in which the best \( T\% \) of individuals are eligible for reproduction, and the parent individuals are drawn randomly with equal probabilities from this \( T\% \) fraction. Truncation selection is one of the rank-based selection schemes, which have the advantage of being
 independent of the distribution of the fitness values of individuals. We also used so-called elitist selection (Grefenstette 1986), in which the current best individual is copied unchanged to the population of the next generation. Elitist selection guarantees asymptotic convergence of the optimization and it has usually been found to improve the convergence speed of the algorithm.

In binary GA’s, mutation means a change from zero to one, or vice versa. A simple guideline is to have \( p_m = 1/L \), where \( L \) is the length of the chromosome (Mühlenbein 1992). The other variation operator, recombination (see e.g., Goldberg 1989), is the search operator that most clearly distinguishes evolutionary optimization algorithms from other stochastic optimization methods (Eshelman and Shaffer 1993b).

Here, uniform crossover was used. Two children are produced by randomly dividing the genes (here sites) present in parents (bold and lightface characters indicate which elements are swapped): Parents: 11000000 Children (e.g.): 11011000 11111100 11000100

**Fitness evaluation and resource limitation**

The value of a subset of patches is evaluated as follows:

**ALGORITHM** fitness evaluation

1. Perform resource limitation on \( S \) giving solution \( S^R \).
2. Evaluate \( F(S^R) \) using Eq. 1.
3. Based on information gathered from step (2), perform local search on \( S^R \) giving \( S^L \).
4. Evaluate the \( F(S^L) \) using Eq. 1.
5. Return into the GA the better of solutions \( S^R \) and \( S^L \) and the associated fitness value.

The solution \( S \) is generated by the GA as a result of stochastic operations selection, recombination, and mutation. Therefore there is no guarantee that \( S \) satisfies the resource constraint. Thus the first step of fitness evaluation is resource limitation, in which randomly selected sites are taken out of the solution until cost(\( S \)) \( \leq \) \( R \). Also, the optimal solution will be such that all or almost all resource has been used. Therefore, if the cost of a solution is less than the limit, randomly selected sites are added to the solution until no more sites can be added without exceeding the resource limit. Resource limitation produces solution \( S^R \) from \( S \). Next solution \( S^R \) is evaluated, during which statistics about the occupancy of patches are gathered. These statistics provide information used in local search. Information is gathered also for patches not included in \( S^R \). Dynamics of these patches are iterated (they can go extinct and be recolonized) along with patches in \( S \), with the difference that populations on patches that are not in \( S \) cannot colonize other patches, and therefore do not affect the dynamics of selected sites.

**Local search**

Local search improves on a candidate solution by examining its neighborhood and choosing a better solution if one with a lower value (when minimizing) for the objective function can be found. In practice this is done by doing small directed changes to the site selection, and reevaluating \( F(S) \) in the hope of finding a more persistent site selection. Here local search is implemented based on information gathered during the fitness evaluation of the solution \( S \) that has been produced by the GA. First we define some quantities that were alternatively utilized by local search. The following quantities (Eqs. 3–5) are increasingly complex problem-specific measures for the dynamic importance of individual habitat patches:

\[
J_i = \frac{1}{NT} \sum_{n=1}^{N} \sum_{t=1}^{T} p_i(n, t). \tag{3}
\]

\( J_i \) is the incidence of patch \( i \), in other words the frequency with which patch \( i \) was occupied in the course of the simulations. \( J_i \) tells something about the importance of a patch. Obviously a patch that is almost never occupied cannot be very important for the dynamics of the metapopulation. However, a patch with high \( J_i \) does not have to be important for persistence. Consider a small patch that is located just outside a dense cluster of patches. This patch may have a relatively high \( J_i \) value, but only due to the rescue effect, which means that the patch is not persistent by itself and is unlikely to be occupied during “bad” years when occupancy is low.

It is obvious that years with low patch occupancy are most critical for metapopulation persistence. Thus patches that are often occupied during bad years are likely to be important for persistence, as these patches are the ones that recolonize the rest of the metapopulation. We next introduce another quantity, the weighted incidence of patch \( i \), which takes into account when the patch is occupied by weighting the occupancy by the yearly metapopulation extinction probability:

\[
J_i^p = \frac{1}{\sum_{n=1}^{N} \sum_{t=1}^{T} \mu(X_{n,i}) p_i(n, t)} \sum_{n=1}^{N} \sum_{t=1}^{T} \mu(X_{n,i}) p_i(n, t). \tag{4}
\]

\( J_i^p \) does not take into account whether a patch is in a position to effectively recolonize other patches. Therefore a large but isolated patch can have a moderately high \( J_i^p \) value. We further define the colonization capacity of patch \( i \) as follows:

\[
\gamma_i = J_i \sum_{j \neq i, j < k} C(i, j)(1 - E_j) \tag{5}
\]

where \( C(i, j) \) is the probability that patch \( i \) colonizes patch \( j \), and \( E_j \) is the intrinsic extinction probability of patch \( j \). The functions \( C(i, j) \) and \( E_j \) are specific for the metapopulation model. The weighted colonization capacity \( \gamma_i^p \) is obtained by replacing \( J_i \) in Eq. 5 by \( J_i^p \).
Measures given in Eqs. 3–5 ignore the price of the patch, which obviously can make a difference to the optimal site selection. We denote measures that have been divided by patch cost by superscript “c”, for example $J^c = J/cost(i)$. One further quantity is used in local search:

$$\pi(S, \lambda) = \begin{cases} \lambda_i / \sum_{i \in S} \lambda_i & \text{for } i \in S \\ 0 & \text{for } i \notin S. \end{cases} \quad (6)$$

In Eq. 6, $\lambda$ stands for any local search measure (e.g., $\gamma^{nc}$); $\pi(S, \lambda)$ is a normalized probability vector (elements sum to one) that has nonzero elements for patches in set $S$. These probabilities are used when adding/removing patches from the site selection during local search.

Basically the local search procedure throws out a number of patches that have a low value for the local search measure $\lambda$. Next patches that have a high value for $\lambda$ are added until no patches can be added without exceeding the resource limit. In detail, the local search algorithm is as follows:

**ALGORITHM** local search
1. Set $S_{out} = \emptyset$, $S_{all} =$ set of all patches, $S =$ the solution that local search attempts to improve.
2. Remove from $S$ all patches $i$ that have $J_i = 0$, add these patches to $S_{out}$.
3. Set $N_{out} = 0.2(\text{count}(S))$, where count($S$) is the number of selected patches in $S$. Set “out” $= 0$.
4. **WHILE** out $< N_{out}$ **DO**
   4.1 Compute probabilities $\pi(S, 1/\lambda)$.
   4.2 Randomly select patch $i \in S$ based on probabilities $\pi(S, 1/\lambda)$.
   4.3 Remove $i$ from $S$. Add $i$ to $S_{out}$, $N_{out} = N_{out} + 1$.
5. **DO**
   5.1 Set $S_{in} = S_{all} \cap (S_{out} \cup S)$.
   5.2 Remove from $S_{in}$ all patches $i$ having $\text{cost}(i) > R\text{-cost}(S)$.
   5.3 **IF** $S_{in} = \emptyset$, **STOP**.
   5.4 Compute probabilities $\pi(S_{in}, \lambda)$.
   5.5 Randomly select patch $i \in S_{in}$ based on probabilities $\pi(S_{in}, \lambda)$. Add patch $i$ to $S$.

*The incidence function model (IFM)*

The metapopulation model used in this study, the incidence function model (Hanski 1994), has been extensively discussed elsewhere (Hanski 1994, 1997, 1999, Hanski et al. 1996, Moilanen et al. 1998, ter Braak et al. 1998, Moilanen 1999, 2000). Therefore only the essentials of the model will be given here. The incidence function model is a stochastic patch occupancy model. It relates the basic processes of classical metapopulation dynamics, local extinction, and colonization, to the basic measures of highly fragmented landscapes, habitat patch area, and isolation. Separate functions are used to describe extinction of local populations and colonization of empty habitat patches. Local extinction depends on the area of the habitat patch; the intrinsic extinction probability of patch $i$ is given by $E_i = eA_i^\gamma$, where $e$ and $x$ are parameters, where $A_i$ is the area of patch $i$. The colonization probability of an empty habitat patch depends on connectivity: $C(i) = S(i)(S(t)^\gamma + y^2)$, where $S(i)$ is the connectivity of patch $i$ at time $t$, and $y$ is a parameter. The way of defining connectivity is an important feature of the model (see Moilanen and Nieminen 2002):

$$S_i(t) = \sum_j p_j(t)\exp(-\alpha d_{ij})A_j^\gamma \quad j \neq i \quad (7)$$

where $p_j(t)$ is the state of occupancy of patch $j$ at time $t$, $d_{ij}$ is the distance between patches $i$ and $j$, and $\alpha$ is a constant that sets the effect of distance on migration. Parameter $b$ scales the relationship between the migration rate and patch area. The final important ingredient of the model is the rescue effect (Brown and Kodric-Brown 1977), which is implemented in the model by assuming that colonization pressure decreases the effective extinction probability of a local population. Thus $E_i$ is replaced by $E_i(t) = (1 - C_i(t))E_i$ in model parameterization and simulation (for elaboration see Hanski [1997, 1999]).

Regional stochasticity (Hanski 1991), that is, spatially correlated environmental stochasticity, may be an important component in metapopulation dynamics, because the increased correlation between dynamics of local populations strongly increases fluctuations in metapopulation occupancy. In other studies using the IFM, regional stochasticity with log-normally distributed yearly variation in effective patch areas has been used (see Moilanen et al. 1998). In this, the areas of all patches are multiplied each time unit by a number that is distributed $10^{\ln(s)\eta}$, with the standard deviation ($s$) of the normal distribution ($N$) determining the strength of regional stochasticity. Unrealistically, regional stochasticity is synchronous across the entire network in this implementation. Here we implement regional stochasticity in a one-step more complex manner.

We start from the fact that the M. diamina metapopulation consists of relatively dense patch aggregates, which are separated from each other by distances that are long compared to the diameter of a single patch aggregate or to the typical dispersal distances of the species (thus these aggregates are almost independent in terms of population dynamics). It is reasonable to assume that in such a system the correlation in local dynamics can be strong within a cluster, but less so between clusters. We thus divide the patch network into separate parts, and use two components for regional stochasticity, one that is synchronous over the network, and one that is synchronous within clusters but independent between clusters. The yearly area multiplier for each network is $10^{\eta_{\text{cn}}}w$, where $\eta \sim N(0, \sigma_\eta)$ is the same for all networks, and $w_{\text{cn}} \sim N(0, \sigma_w)$ is subnetwork-specific. Note that the patch class division does not
Fig. 1. The *M. diamina* patch system. The sizes of the circles are scaled according to the area of the patch. The shading indicates cost category: black patches are 10 times more expensive than unshaded patches, and gray patches have twice the cost. The dotted lines separate three dynamically semi-independent regions of the metapopulation. The resized rectangle at the bottom right shows an enlargement of the most important patch cluster in the metapopulation. Small squares within circles indicate patches that were occupied in summer 1999.

Melitaea diamina metapopulation and patch costs

The procedures described in Methods are applied to a metapopulation of the false heath fritillary butterfly, *Melitaea diamina*, which is an endangered species in Finland. The butterfly has declined greatly during the latter half of the twentieth century, and currently it is known to occur only in two areas in Finland. In one of these areas a network of 125 habitat patches is scattered around an area of $\sim 20 \times 30$ km near Tampere in southern Finland (Wahlberg et al. 1996; Fig. 1). This metapopulation of *M. diamina* is endangered because many of the meadows are rapidly being overgrown with willow or planted spruce trees.

*Melitaea diamina* lives on moist meadows growing the obligate larval host plant, *Valeriana sambucifolia* (Wahlberg 1997). In 1999, 49 of the meadows were occupied by the butterfly, with the minimum, average, and maximum areas of the occupied sites being 0.046, 0.69, and 3.97 ha, respectively. The patch network satisfies the requirements for the species to have metapopulation persistence as habitat patches are discrete, none of the patches is large enough to be a “mainland,” and, considering the spatial scale of the patch network, migration between patches is necessarily affected by distance.

The conservation of the *M. diamina* metapopulation involves maintaining the quality of the meadows for grazing or by clearing away high vegetation. Maintaining all 125 meadows is not feasible without substantial voluntary help, and therefore it is desirable to direct the conservation measures to those parts of the patch network that best facilitate the persistence of the metapopulation. Wahlberg et al. (1996) were able to successfully predict the distribution of *M. diamina* using the incidence functional model (IFM) parameterized for an ecologically similar species, *Melitaea cinxia* (Hanski et al. 1996). Since this analysis, new information from the metapopulation has been gathered, and it is now possible to produce detailed suggestions about the conservation of the metapopulation.

Patch cost is a key factor in our problem formulation. We assume that the base cost of protecting a patch is directly proportional to patch area. Good quality patches, that is, patches in early successional stages, are assigned this cost. The second cost category includes patches in the late successional stage. These patches are assigned two times cost, which consists of two maintenance operations, one immediately and one in the middle of the
M. diamina parameter estimation

Parameter estimation for the M. diamina metapopulation presented some difficulties due to certain properties of the empirical data. The metapopulation was first surveyed in 1994. Then 94 patches, 35 of them occupied, were found. The second survey was conducted in 1999. However, between 1994 and 1999, 14 patches had been almost or completely destroyed due to construction, agriculture, and forestry, whereas 36 new patches had been found. This number of new patches is so large that they cannot be omitted from parameter estimation without potentially biasing the estimates. We assumed that the occupancy values in the 1994 data for patches unknown at that time would be the same as in the 1999 data. This causes a slight underestimation of metapopulation turnover rate as some (probably less than five) turnover events are missed, but the parameters most affecting the actual occupancy pattern (x and α) should not be much affected. Another problem in parameter estimation has to do with patch quality. The patches were delineated in the field based on the presence of the larval host plant. The standard approach to parameter estimation assumes that all patches are of equal quality, which however turns out not to be the case here. Biologically implausible parameter estimates are obtained when parameters are estimated using all patches. For example, local extinction risk was estimated to be almost independent from patch area (the value of parameter x was estimated to be as low as 0.3). Erroneous estimates were obtained because the data contains several large and well connected but empty patches. All these patches are in the late successional stage, with high vegetation shading the ground. In 1994 and 1999 there were 15 and 34 patches in the late successional stage, respectively, but only one and seven of them were occupied. The average occupancy of these patches is very low compared to the ~40% occupancy of early successional patches. The numbers of butterflies observed in late successional patches are low, and field experts assessed that it is doubtful whether the species is actually able to successfully reproduce in these patches (J. Helioäla, personal communication). In brief, the late successional patches seem to act as sink habitats, and we deleted them from the data set before parameter estimation. This leaves 91 patches with occupancy information for two years. Values missing for 30 patches in 1994 were assumed to be identical to those in the 1999 data. Parameter estimates were then obtained using the method of Moilanen (1999). Parameter estimation was repeated three times using 5500 parameter set evaluations for each estimation (for details see Moilanen [1999]). Combining data from the three estimations produced the following parameter values (with 95% confidence intervals computed based on likelihood ratio tests): $x = 1.08 \ (0.89, 1.27)$, $e' = 1.62 \ (1.15, 1.75)$, $\alpha = 0.71 \ (0.69, 0.73)$, and $A_c = 0.078 \ (0.044, 0.098)$. The values of IFM parameters $e$ and $y$ are calculated as $e = A_5$ and $y = \sqrt{e'/e}$, through which $e'$ is defined.

**Results**

**The optimization method**

Fig. 2 illustrates the convergence of the optimization algorithm. All variants of genetic algorithm plus local search (GA + LS) converge faster than the GA itself, but there are differences between LS variants. All LS variants based on the colonization index γ converge faster than LS variants based on the predicted incidence $J$. This demonstrates that search results are improved by the use of extra information about the properties of the problem. Overall, the γ index variants converge quickly and consistently to almost identical site selections and $F(S)$ values (a few small patches may occasionally be selected differently).

Table 1 illustrates other aspects of the search process. In the beginning of search, solutions are not yet very good and they are relatively easy to improve upon, and consequently the success rate of the local search op-
frequent and in the simulations used to determine population) and the number of extinctions observed score is lower for the variant using the $J$ rate of local search using the $\bar{w}ic$ with the results in Fig. 2. For example, the success shows that local search is often able to improve a site over the entire optimization run (45%) is high and than what was achieved with the combination of GA lemm, which is a good result considering that the com-

The dispersal ability of the species (here given by parameter $\alpha$) is possibly the single most important factor in a metapopulation model. When $\alpha$ is large (Fig. 3D, smaller dispersal range), the selected patches are consistently closer together, and more of them are initially occupied, than when $\alpha$ is small (Fig. 3A). Note the patches in the southeast region of the network at Siitama. These patches are never selected in the runs with $\alpha = 1.5$, but they are always at least partially selected when $\alpha = 0.4$. A comparison between Fig. 3A and Fig. 3C demonstrates potential effects of the time frame on optimization. When the time frame is short (Fig. 3C), large and initially occupied patches tend to be selected. When the time frame is long (Fig. 3A), metapopulation dynamics become important and the significance of connectivity increases in comparison with initial oc-

Table 2 demonstrates the difference between $F(S)$ (the mean yearly extinction probability of the metapopulation) and the number of extinctions observed in the simulations used to determine $F(S)$. When $F(S)$ is small, no extinctions are observed. When regional stochasticity is increased, extinctions become more frequent and $F(S)$ increases sharply. The number of extinctions would have been a bad criterion to use in optimization, because it cannot separate between two different but always persistent site selections. Note also the binomial variation in the numbers of extinctions. It is interesting to observe the effect of regional stochasticity on $F(S)$. When there is no regional stochasticity, the $F(S)$ values are very low because of the rescue effect and independent patch dynamics. When regional stochasticity is increased, the degree of correlation in patch dynamics increases, and the $F(S)$ values increase rapidly. Unfortunately, the true mechanism(s) and the strength of regional stochasticity usually remain unknown, but for butterflies, and insects in general, values like $v = w_i = 0.25$ do not seem unreasonably high. In practical terms this means that a low $F(S)$ value in the absence of regional stochasticity does not guarantee the persistence of the metapopulation, but a high $F(S) (>0.01)$ is an almost certain indicator of the metapopulation being in dan-

Table 2. Extinction risk, $F(S)$, and the number of global extinctions for each value of regional stochasticity, $v = w_i$.

<table>
<thead>
<tr>
<th>$v$</th>
<th>$R = 50,000$</th>
<th>$R = 100,000$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$F(S)$</td>
<td>Ext.</td>
</tr>
<tr>
<td>0.00</td>
<td>0.0024</td>
<td>0</td>
</tr>
<tr>
<td>0.05</td>
<td>0.0008</td>
<td>0</td>
</tr>
<tr>
<td>0.10</td>
<td>0.0033</td>
<td>0</td>
</tr>
<tr>
<td>0.15</td>
<td>0.0071</td>
<td>1</td>
</tr>
<tr>
<td>0.20</td>
<td>0.0209</td>
<td>6</td>
</tr>
<tr>
<td>0.25</td>
<td>0.0303</td>
<td>3</td>
</tr>
<tr>
<td>0.30</td>
<td>0.0845</td>
<td>10</td>
</tr>
</tbody>
</table>

Notes: The $F(S)$ measure detects differences between the persistence of two different site selections even when no extinctions occur during the simulations (here $N = 50$ was used). A site selection that is apparently highly persistent when no regional stochasticity is present can turn out not to be persistent when strong regional stochasticity is present. Note binomial variation in the numbers of extinctions when $R = 50,000$. 

---

Table 1. Comparison of the performance of local search (LS) variants based on the same optimization runs used to generate Fig. 2, reporting the success rates of different local search operators in the beginning of optimization (first three genetic algorithm (GA) generations) and over the entire optimization run (25 GA generations), along with $F(S)$ values at the end of optimization.

<table>
<thead>
<tr>
<th>LS variant</th>
<th>No. GA generations</th>
<th>Final $F(S)$ (1 SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$J$</td>
<td>3 25</td>
<td>$2.71 \times 10^{-31}$ ($2.0 \times 10^{-31}$)</td>
</tr>
<tr>
<td>$J^c$</td>
<td>0.83 0.22</td>
<td>$7.41 \times 10^{-32}$ ($5.7 \times 10^{-32}$)</td>
</tr>
<tr>
<td>$J^e$</td>
<td>0.80 0.39</td>
<td>$9.33 \times 10^{-32}$ ($9.4 \times 10^{-32}$)</td>
</tr>
<tr>
<td>$J^\gamma$</td>
<td>0.90 0.42</td>
<td>$9.27 \times 10^{-32}$ ($5.8 \times 10^{-32}$)</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.76 0.29</td>
<td>$1.26 \times 10^{-32}$ ($7.9 \times 10^{-32}$)</td>
</tr>
<tr>
<td>$\gamma^w$</td>
<td>0.62 0.25</td>
<td>$1.27 \times 10^{-32}$ ($1.1 \times 10^{-32}$)</td>
</tr>
<tr>
<td>$\gamma^w$</td>
<td>0.78 0.43</td>
<td>$1.42 \times 10^{-32}$ ($9.8 \times 10^{-32}$)</td>
</tr>
<tr>
<td>$\gamma^w$</td>
<td>0.78 0.45</td>
<td>$1.27 \times 10^{-32}$ ($7.8 \times 10^{-32}$)</td>
</tr>
<tr>
<td>GA only</td>
<td>... ...</td>
<td>$1.55 \times 10^{-29}$ ($4.0 \times 10^{-29}$)</td>
</tr>
</tbody>
</table>
Fig. 3. Demonstration of typical effects of parameter changes, time frame, and cost inclusion with the Siitama patch network: (A) $\alpha = 0.4$, $T = 100$, no cost; (B) $\alpha = 0.4$, $T = 100$, cost included; (C) $\alpha = 0.4$, $T = 5$, no cost; and (D) $\alpha = 1.5$, $T = 100$, no cost. Dark color indicates a patch selected into the optimal site selection. Patches occupied in summer 1999 have been marked with dots (white or black). When $T$ is short, comparatively many initially occupied patches are selected. When $\alpha$ is high (low dispersal ability), a dense cluster of patches is selected. This region of the metapopulation includes two patches with 2× cost. These patches belong to the optimal site selection when cost is not considered, but not when cost is considered.

Melitaea diamina conservation

Fig. 4 shows how the optimal site selection changes as resource ($R$) is added. With $R = 50,000$ m$^2$ and $R = 100,000$ m$^2$, it is optimal to select patches only from the Siitama cluster. At $R = 300,000$ m$^2$, it is optimal to select the core areas from all three patch clusters. Fig. 5 shows the effect of $R$ on $F(S)$. $F(S)$ drops rapidly (persistence increases) when the resource is increased up to $R = 500,000$ m$^2$ [$F(S) = 1.3 \times 10^{-21}$]. At this point 54 patches (34 occupied) are selected with the combined area of 43.7 ha. Subsequently $F(S)$ decreases less rapidly up to $R = 900,000$ m$^2$ [$F(S) = 3.7 \times 10^{-28}$, 81/46 patches, 61.7 ha], after which $F(S)$ levels off. This indicates that adding resource up to $R = 500,000$ is very useful for improving persistence, but thereafter less and less is achieved by adding resource. After $R = 900,000$ m$^2$ practically no improvement can be gained by adding resource; this is the point where only isolated and often expensive patches remain unselected. The cost of the entire patch network would be ~$1,900,000$ resource units, hence it appears that protecting all known habitat in this network would not be cost effective.

Figs. 6–8 show results of sensitivity analysis of optimal site selection, based on information gathered during parameter estimation. Parameter estimation found 450 parameter combinations that were within the joint four-parameter 95% confidence interval (likelihood ra-
Fig. 4. Optimal site selection at different resource levels, $R$, assuming $M. \text{diamina}$ standard parameter values and regional stochasticity $\psi = w_i = 0.20$.

tio test with critical point taken from the $\chi^2$ distribution with four degrees of freedom). We randomly sampled 100 of these points, and the optimal site selection was found for each of these parameter sets. We calculated for each patch an inclusion score ($I_i$), that is, the proportion of replicates in which the patch was included in the optimal site selection. This analysis was done separately for the three clusters (Fig. 1) because of regional stochasticity. While it is quite certain that regional stochasticity exists in this system, there is little information that could be used to figure out its strength and parameterize it. Therefore, considering that the patch clusters are relatively small, we would not feel comfortable about suggesting that one cluster would be good enough for preserving the species. It is desirable to maintain at least two patch clusters to allow for a managed reintroduction in case the metapopulation disappears from one cluster. Therefore, we produced optimal site selections for the three metapopulation regions (see Fig. 1) separately. The amount of resource used in the analysis ($R = 50\,000–150\,000$) was chosen so that the results would be relevant for a situation in which only a small part of the habitat ($\sim 35\%$) can be preserved.

Fig. 5 shows the results of the sensitivity analysis for the Siitama region. Only patches from the main cluster are selected when $R = 100\,000$ m$^2$ [respective mean $F(S) = 1.6 \times 10^{-10}$]. With $R = 50\,000$ m$^2$ [$F(S) = 8.0 \times 10^{-13}$], selected patches aggregate into the northern part of the main cluster. Fig. 7 shows similar results for the Viitapohja region. With $R = 50\,000$ m$^2$
We now summarize our recommendations for the conservation of *M. diamina* based on the metapopulation dynamical analysis. Due to the possibility of regional stochasticity, we do not feel confident that any one of the three semi-independent patch networks is large enough to maintain a *M. diamina* metapopulation through two or three very bad years. Also, apart from metapopulation dynamical issues, it is doubtful whether a small well-connected patch network can maintain adequate genetic diversity for long-term persistence (see e.g., Saccheri et al. 1998). Therefore, we conclude that it is desirable to protect the core areas of all three subregions of the *M. diamina* metapopulation. The Siitama region is in our estimation clearly the most persistent part of the metapopulation. We recommend that the northern part of the Siitama patch network (box in Fig. 6) be protected entirely. In Viitapohja there is another patch network of eight patches (box in Fig. 7) that is predicted to be quite persistent considering its small size. We recommend the protection of this region. The Sorila region is more problematic than the other two regions, because the area is larger, with more patches spaced further apart from each other. Our recommendation is to protect the central parts of both the eastern and the western parts of the Sorila patch network (larger box in Fig. 8). The total number of patches protected in our suggestion is 40, requiring that ~35% of the *M. diamina* habitat area be protected.

**Discussion**

Site selection algorithms have traditionally used information about the occurrence of species in a group of sites at one point in time. Simplifying, the goal of
site selection has been to provide a minimum set of patches that may include only one (or a few) populations for each species (e.g., Pressey et al. 1993, 1997). This approach is focused on representation, and it ignores the spatial dynamics of species; species are assumed to persist on the patches where they have been observed. In contrast, in the context of metapopulation dynamics, a species living in a set of a few small habitat patches does not necessarily survive for long. There are many examples of species living in metapopulations, and it appears that some species may require tens of small patches to persist (Hanski 1999:17–19).

In this study we present site selection methods for a single species living in a metapopulation. It is appropriate to reiterate the critical points in doing metapopulation site selection. The first step is finding a metapopulation model that adequately corresponds to the biology of the species. This step requires knowledge of the habitat requirements and biology of the species. Here we used a patch occupancy model, which is best suited for metapopulations living in highly fragmented landscapes (Hanski 1999). The next step is model parameterization (see Moilanen 1999, 2000). This is probably the most critical step of the study, as all subsequent predictions can be strongly affected by the parameter estimates. In some cases parameter values may be available from the literature (Wahlberg et al. 1996), but in most cases empirical data are needed. It is recommended that at least two complete surveys of the metapopulation be available for parameter estimation, so that there is direct information about population turnover. It is important that the empirical data corresponds to the assumptions made in parameter estimation. For example, in this study there was a large quality difference between patches in early and late successional stages, which had major consequences for parameter estimation. The final step is to produce predictions, in which variation in parameter estimates should be considered by some form of sensitivity analysis. Quantitative estimates of metapopulation persistence should be interpreted cautiously, because they are likely to be significantly dependent on the assumptions made about regional stochasticity.

To avoid problems associated with the use of point estimates, we found optimal site selections for 100 parameter combinations sampled from within the joint confidence limits for the four parameters of our metapopulation model. Results of these replicate estimations were then combined to provide each patch a probability of belonging to the optimal site selection. This idea is analogous to the irreplaceability concept of Pressey et al. (1994), which can be used to establish priorities for the conservation order of sites. Irreplaceable sites are such that have no flexible alternatives if the conservation goal is to be achieved (see, e.g., Williams [2000] for a review). Based on this analysis, we identified three patch aggregates that are responsible for the viability of the *M. diamina* metapopulation. It appears that protecting a core set of ~40 patches that corresponds to ~35% of the pooled habitat area in the metapopulation would be a cost-effective way for protecting the species. This result agrees with the intuition of field experts (J. HelioÈla È and N. Wahlberg, personal communication). The greatest difference between field expert recommendations and metapopulation dynamical analysis is that field experts would consider protecting a few isolated but large and good-quality patches. These patches evidently are not important for metapopulation dynamics, as they are never included in the optimal site selection found by the metapopulation site selection algorithm. However, this conclusion has to be interpreted carefully. It might be that large and good-quality patches are extremely important as refuges during bad years (these patches could then be used as a source for reintroductions into other patches). Also, if model parameterization is wrong so that the dispersal capability is badly underestimated or extinction probabilities of large populations overestimated, then the importance of large but isolated patches becomes underestimated. However, considering that the meadows are being overgrown, and good habitat quality will thus not be maintained without management, we conclude that it is probably best to maintain habitat in clusters so as to facilitate recolonization of empty meadows. Finally, we point out that there are cases where identifying the optimal site selection is practically impossible without a computational aid. For example, if patch costs vary greatly, figuring out an optimal site selection by just looking at a map will be very difficult.

The algorithm we have introduced is quite successful technically. It converged fast and reliably to (almost) identical site selections in our examples. In our opinion,
the success of the algorithm was due to two factors: the definition of the object function, and the inclusion of local search into the optimization algorithm, which significantly improved both optimization speed and consistency when compared to using only the genetic algorithm. Unfortunately, we have no way of ensuring that the solution we find is the global optimum. However, if an intuitively reasonable result is consistently found with a stochastic search method, it is likely that the solution is good.

We conclude by summarizing the correct application and risks associated with the present approach. The greatest risk in doing metapopulation site selection is the misinterpretation of $F(S)$ values, leading to the underestimation of the habitat needed for persistence. This may happen, for example, if regional stochasticity is ignored. Overall, giving any recommendations about the “minimum viable metapopulation size” or any similar quantity is quite risky, because unknown factors may be affecting metapopulation persistence. There are other things that may be done with less risk. If the amount of resource available for conservation is known, it is valid to figure out the site selection that maximizes persistence given the resource limit. Even if the resource limit is not known, it is possible to find a resource level beyond which $F(S)$ does not decrease much. This happens when all dense habitat patch clusters have been selected and the remaining habitat consists of isolated patches that have little effect on metapopulation persistence.

This study is related to metapopulation viability analysis, but instead of focusing on the time to extinction or the probability of metapopulation extinction in a small number of given scenarios, we focus on selecting the combination of sites that maximizes metapopulation persistence. It is possible to expand the present approach to multiple independent or interacting species in larger spatial scales, given that information about species interactions and metapopulation model parameters exist.

Acknowledgments

We thank Janne Helioä and Niklas Wahlberg for information about the *M. diamina* metapopulation. Mark Burgman, Ilkka Hanski, Hugh Possingham, and Niklas Wahlberg are thanked for helpful comments on the manuscript. This study was supported by The Academy of Finland grant no. 71516 to A. Moilanen and grant no. 45125 to M. Cabeza.

Literature Cited


