Dynamic reserve site selection*

Christopher Costello† and Stephen Polasky‡

July 2, 2002

Abstract

Establishing biological reserves in key habitats is one method of preventing loss of biodiversity. We analyze a dynamic reserve site selection model in which a conservation planner receives a budget each period to purchase sites of heterogeneous cost and ecological value in order to maximize species conserved. Sites chosen as reserved are conserved, while sites left outside the reserve system are threatened by development. We formulate this problem as a stochastic dynamic integer programming problem. We solve several examples to illustrate how sequential choice compares with static choice in which all selections can be made at one time. We also compare optimal solutions with solutions obtained using simple heuristic choice algorithms. In general, an “informed” heuristic algorithm that accounts for both biodiversity values and development threats across sites results in a small percentage loss in the expected number of species conserved relative to the optimum. The performance of heuristic algorithms diminishes as the time horizon and number of sites increases. The theory is applied to vertebrate and development threat data from southwestern California.

JEL Classification: R52, Q20

Keywords: reserve network, biodiversity, threat, dynamic programming

---

*The authors acknowledge helpful insights and comments from David Stoms, Arthur Small, Geoff Heal, two anonymous referees, participants of the Economics of Biodiversity workshop held in Santa Barbara in May, 2001, participants of the World Congress Biodiversity session in June, 2002, and seminar participants from Georgia State University, University of Minnesota and the Stanford University center for conservation biology.

†University of California, Santa Barbara, corresponding author. Bren School of Environmental Science & Management, UC Santa Barbara, Santa Barbara, CA 93106. (Email) costello@bren.ucsb.edu. (Tel) 805-893-5802. (Fax) 805-893-7612.

‡University of Minnesota
1 Introduction

The conversion of large areas of undeveloped habitat to human dominated landscapes has been a major cause of the loss of biodiversity. Only about fifty percent of the forest area that existed at time of the rise of agriculture remains, of which less than half remains in large tracts capable of sustaining a full range of biological diversity (World Resources Institute 1998). Less than half of the wetlands in the United States that existed at the time of the coming of European settlers still remain, and in California, less than 10% remains (Viliesis 1997). Habitat conversion is likely to continue into the foreseeable future as the human population continues to grow, perhaps by another three billion or more over the next fifty years. Given present and projected future trends in land use and habitat conversion, the question of how best to conserve biological diversity is an urgent one. Which areas of habitat are the most important to protect in order to conserve biodiversity and how should conservation agencies target their actions so that they accomplish the most with limited budgets?

An important strategy to conserve biological diversity is to protect habitat through the establishment of a system of biological reserves. Examples of this approach include the system of National Wildlife Refuges in the U.S., national parks in many countries, and private conservation reserves established by such groups as the Nature Conservancy. It is not possible to protect all biologically important sites in a reserve network given limited resources. Therefore, it is important that a conservation agency choose wisely among potential reserve sites to allow for maximal protection given its limited budget.

Starting in the early 1980’s, a fairly substantial literature in conservation biology has
addressed the "reserve site selection problem" (see, for example, Kirkpatrick (1983), Margules et al. (1988), Pressey et al. (1993)). A simple formulation of the reserve site selection problem is to choose sites to include in a reserve network in order to conserve the greatest number of species possible given a budget constraint. A species is considered conserved if it is present in at least one site included in the reserve network. A species that is not present in any selected site is not considered conserved. This approach implicitly assumes that no species survives at sites outside not in the reserve network, which is like assuming that habitat will be destroyed in all sites not contained in the reserve network. If the presence or absence of each species at each site is known with certainty, the reserve site selection problem is what is called a "maximal coverage problem" in operations research (Church and ReVelle 1974). Branch-and-bound algorithms can be used to find optimal solutions. Applications of these methods to find optimal solutions to the reserve site selection problem include Cocks and Baird (1989), Saetersdal et al. (1993), Church et al. (1996), Kiester et al. (1996), Csuti et al. (1997), Pressey et al. (1997), Ando et al. (1998), Snyder et al. (1999), and Polasky et al. (2001).

All of these papers are static in the sense that they assume there is a one-time decision about which sites to protect. In reality, there are several important dynamic factors affecting conservation and land use decisions. Budgets for conservation agencies, public or private, are typically not set in a one-time lump sum amount. Rather, public agencies receive an allocation on a regular budget cycle. These agencies typically cannot borrow against future budget outlays to make land purchases in the present. Likewise, private conservation groups receive donations through time, and find it difficult to borrow against the promise of future donations. Conservation agencies therefore must sequence their selections, setting aside some land today knowing that they will set aside other lands in the future.
Some land that is desirable to protect, but cannot be set aside as a reserve at present due to budget restrictions, may be unavailable for protection in the future because it is developed in the meantime. Given this threat, a conservation agency may wish to take account of development pressures when it is deciding which parcels of land are the highest conservation priorities. In general, a conservation agency will know something about which areas face high or low development risk but will rarely know exactly when and where future development will occur.

The threat of habitat loss through development has been recognized by the conservation biology literature as an important element to consider in defining high priority locations for conservation. A number of papers have defined high priority sites for conservation, or “biodiversity hotspots,” as those sites containing both a high degree of threat and high biodiversity. Papers analyzing conservation priorities in this way include Myers 1988, Dinerstein and Wikramanayake 1993, Balmford and Long 1994, Sisk et al. 1994, Cole and Landres 1996, Mittermeir et al. 1998, Cowling et al. 1999, Lombard et al. 1999, Ricketts et al. 1999, Abbitt et al. 2000, Myers et al. 2000, and Pressey and Taffs 2001. This literature makes an advance over the reserve site selection literature in that it recognizes that some lands, but certainly not all, will remain natural at least for some period of time even if it is not formally protected. When there is some probability of species survival outside of reserves, minimizing the loss of biodiversity is not the same as maximizing the number of species conserved in a reserve network (Witting and Loeschcke 1993, 1995). The literature that places priority on sites with high biodiversity and facing high threat, however, does not attempt to analyze a dynamic problem. Combining uncertainty about the threat of development with sequential reserve site selection choice makes the conservation decision problem a dynamic problem. A similar problem, where future site availability is random,
is analyzed in a stochastic dynamic programming framework by Meir et al. 2002, who analy-
zye a 12 site multiple time-period model using data from Australia, and Meir et al. 2002, 
who compare the performance of various heuristic algorithms for very large problems. Im-
portantly, they find that when sites become available at random, opportunistic acquisition 
often outperforms a strategy of sequentially adding sites that are part of the static solution 
as they become available.

In this paper, we set up the reserve site selection problem as a stochastic dynamic integer 
programming problem. We assume the objective of a conservation agency is to maximize 
the number of species surviving at the end of the planning horizon. The agency affects 
the probability of survival by selecting sites as biological reserves each period given the 
budget constraint that period. Each period, all sites that have not been selected to date 
face some probability of conversion to developed land; conversion precludes future selection 
as a reserve. Given the threat of conversion and the per period budget constraints, what is 
the optimal sequence of reserve site selection? In section 2 we formally define the problem as 
a stochastic dynamic integer programming problem. In sections 3 and 4 we analyze several 
relatively simple examples that illustrate important general features of solutions to this 
problem. In section 5 we apply the theory to an empirical example of designing a reserve 
network to protect native vertebrate species in California. Section 6 contains concluding 
remarks.

2 A Model of Dynamic Reserve Site Selection

There are $J$ sites, indexed by $j = 1, 2, ..., J$, and $I$ species, indexed by $i = 1, 2, ..., I$. The 
$J \times I$ matrix $A$ has a typical element, $A_{ji}$ which equals 1 if site $j$ contains habitat suitable 
for species $i$, and equals 0 otherwise. At the start of each time period $t$, $t = 1, 2, ..., T$ every
site is either "developed", "reserved", or "unreserved". Unreserved status implies that the site has not been developed nor selected as a reserve. Species $i$ exists in site $j$ if and only if $A_{ij} = 1$ and site $j$ is not developed (it may be either reserved or unreserved). We assume that no species survives at a developed site. The development process is assumed to be irreversible. Similarly, once a site is selected as a reserve, it is assumed to remain protected forever. The cost of selecting site $j$ at time $t$ as a reserve is $C_{jt}$. The $t$ subscript allows $C_{jt}$ to embed both time dependent purchase costs and ongoing management or monitoring and enforcement costs associated with maintaining site $j$ in reserve. If site $j$, which starts period $t$ as unreserved, is not selected as a reserve during period $t$, then it is developed at the end of period $t$ with probability $P_{jt}$, and remains unreserved with probability $1 - P_{jt}$.

In general, the variables $C_{jt}$ and $P_{jt}$ will be correlated; sites that are likely to be developed may cost more to protect because the potential profit from development is capitalized into the land price, \textit{a priori} leaving ambiguous the optimality of selecting cheap sites (to allow more sites to be purchased) or costly sites (to place threatened sites in reserve).

Let $R_t$ be a $J \times 1$ vector where $R_{jt}$ equals 1 if site $j$ has been selected as a reserve prior to the beginning of period $t$, 0 otherwise. Let $X_t$ be a $J \times 1$ vector where $X_{jt}$ equals 1 if parcel $j$ is selected as a reserve in period $t$, 0 otherwise. Therefore, $R_{t+1} = R_t + X_t$. Let $N_t$ be a $J \times 1$ vector where $N_{jt}$ equals 1 if site $j$ is unreserved at the beginning of period $t$, 0 otherwise. Let $S_t$ be a $J \times 1$ random vector where element $S_{jt}$ equals 1 if site $j$ is converted from unreserved to developed in period $t$ (following the allocation decision in that period), 0 otherwise. To maintain model generality, the $J \times 1$ vector of probabilities of development, $P_t = (P_{1t}, P_{2t}, ... P_{jt})$, and the $J \times 1$ vector of acquisition cost, $C_t = (C_{1t}, C_{2t}, ... C_{jt})$, may depend on the current pattern of reserves and development. To account for this possibility, we can write the vectors of probability and cost as $P_t = P_t(N_t, R_t)$ and $C_t = C_t(N_t, R_t)$.
In each period, the planner faces a budget constraint. In period \( t \), the planner is given funds \( b_t \). We assume that the planner may not borrow. However, funds that are not spent during period \( t \) may be carried forward to period \( t + 1 \). Funds carried forward earn interest at rate \( \delta \). Let \( B_t \geq 0 \) represent the amount of money the planner begins the period with, prior to receiving the budget amount for period \( t \). Then at the start of period \( t + 1 \), the planner will have \( B_{t+1} = (B_t + b_t - X'_t C_t)(1 + \delta) \). Let \( B_0 \) be the initial budget at the beginning of the planning horizon.

We formulate the dynamic reserve selection problem as a stochastic dynamic integer programming problem. There are \( 2J + 1 \) period \( t \) state variables in this model: \( N_t, R_t \), and \( B_t \), and there are \( J \) period \( t \) control variables in this model, \( X_t \). The objective of the planner is to maximize the total number of species conserved at the end of the planning horizon (i.e., the beginning of period \( T + 1 \)). To do so, the planner chooses \( X_{jt} \leq N_{jt} \) for \( j = \{1, 2, ..., J\} \) and \( t = \{0, 1, ..., T\} \), henceforth denoted \( X_t \leq N_t \), which constrains reserve selections to those sites that are currently unreserved. A species is considered conserved at the end of the planning period if the species exists in at least one site that is not developed at the beginning of period \( T + 1 \). Given that species are sure to survive in undeveloped sites, selecting one site where the species occurs as a reserve is sufficient to ensure species survival. However, in the more realistic case where species survival at a site is less than one, there is value to conserving the species at multiple sites (Polasky et al. 2000). We briefly discuss issues raised by survival probabilities in section 6.

The timing of the model in each period is as follows. At the beginning of period \( t \), the planner observes \( N_t, R_t \), and \( B_t \). The planner receives budget outlay \( b_t \), and then chooses \( X_{jt} \leq N_{jt} \). Elements of \( N_t \) that have not been selected as reserve sites are then subject to possible development. With probability \( P_{jt} \) an unreserved site \( j \) is converted to developed.
Any remaining unspent budget is then carried forward to period $t + 1$ and earns interest.

Let $V(N_t, R_t, B_t)$ be the value of the optimal program given the state variables $(N_t, R_t,$ and $B_t)$ at the beginning of period $t$. Then we can write the stochastic dynamic integer programming equation as follows:

$$V(N_t, R_t, B_t) = \max_{X_t \leq N_t} \mathbb{E}_S V(N_{t+1}, R_{t+1}, B_{t+1})$$  \hspace{1cm} (1)

subject to \hspace{1cm} \begin{align*}
X_t' C_t &\leq B_t + b_t \hspace{1cm} (2) \\
N_{t+1} &= N_t - X_t - S_t, \quad S_t \leq N_t - X_t \hspace{1cm} (3) \\
R_{t+1} &= R_t + X_t \hspace{1cm} (4) \\
B_{t+1} &= (B_t + b_t - X_t' C_t)(1 + \delta) \hspace{1cm} (5)
\end{align*}

Equation (1) is the stochastic dynamic integer programming equation. $\mathbb{E}_S$ is the expectation operator over the vector $S_t$, the random vector that shows which unprotected sites are developed in period $t$. Equation (2) is the period $t$ budget constraint. Equations (3)-(5) provide the equations of motion that govern the transitions of the state variables from one time period to the next, given a reserve allocation in period $t$, $X_t$.

This dynamic program is solved by backward induction starting at the end of the planning period (the beginning of period $T + 1$). The value of the optimal program at the end of the planning period is:

$$V(N_{T+1}, R_{T+1}, B_{T+1}) = \min \left[ \epsilon_{1\times I}, (N_{T+1} + R_{T+1})' A \right] \epsilon_{I\times 1}$$  \hspace{1cm} (7)

where $\epsilon_{m\times n}$ is an $m\times n$ matrix of ones. In words, equation (7) says that the value of the program at the end of the planning horizon is equal to the number of species that remain conserved (i.e. those species that occur in reserves or unreserved lands) at the beginning of
period $T + 1$. The value of the program is independent of the money left over at the end of the planning horizon.

Stepping back one period to the beginning of period $T$ and taking advantage of the fact that in period $T$ we know the value of endowing the future ($T + 1$) with the levels of each state variable, we can write the stochastic dynamic programming problem (equation (1)) as:

$$V(N_T, R_T, B_T) = \max_{X_T \leq N_T} E_{S_T} V(N_{T+1}, R_{T+1}, B_{T+1})$$

subject to equations (2)-(5) for $t = T$. This problem is a stochastic integer programming problem that cannot be solved analytically, but standard computation techniques have been developed that solve problems of fairly high order. The solution simultaneously gives both the optimal $X_T$ (those unreserved sites in period $T$ to purchase and place in reserve) and the value function one period hence. The process is then repeated in period $T - 1$, and continues back to period 1.

Note that the complexity of solving this problem is determined by the number of sites, which determines the number of possible combinations of the state variables $N_t$ and $R_t$, and the number of values that the carryover budget, $B_t$, may assume. Solving a problem with a larger number of species or a larger number of time periods does not significantly complicate solving the stochastic dynamic integer programming problem.

### 3 A Simple Three Site Example

To help get a feel for the dynamic site selection problem, we begin with a simple example involving three sites and two time periods. In each period, one site may be selected. Table 1 lists the species that occur on each site and the probability of conversion for each site in
Table 1: Species configuration on the three sites for the "warm-up" exercise.

<table>
<thead>
<tr>
<th>Species Present</th>
<th>Site A</th>
<th>Site B</th>
<th>Site C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Species Present</td>
<td>1,2,3,4</td>
<td>5,6,7</td>
<td>1,2,5,6,8</td>
</tr>
<tr>
<td>Prob. of Conversion</td>
<td>0.4</td>
<td>0.5</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Insert Figure 1 approximately here.

each period. It is not immediately apparent from inspection of table 1 which site should be selected in the first period. Site A contains the most species that are not contained in any other site (species 3 and 4), site B faces the highest threat of conversion, and site C contains the greatest number of species.

Figure 1, shows the expected number of species conserved after two periods when one begins by choosing site A, site B or site C in period one, where developed and unreserved sites are labeled ‘D’ and ‘U’, respectively. After a site is selected there is conversion risk for each of the other two unprotected sites, which results in four possible outcomes for remaining unreserved areas in period two. If either one or zero unreserved areas remain in period two, the period two choice of what to select, if anything, is trivial. When two unreserved areas remain in period two, a choice must be made between the two remaining sites over which to select. The site that remains unprotected is then subject to another development risk. Figure 1 depicts the full decision tree. For example, if site A is chosen in period 1, there are four possible combinations of the status of sites B and C at the beginning of period 2: (1) only site B is available (prob = 0.1), so site B is selected in period 2, (2) only site C is available (prob = 0.4), so site C is selected in period 2, (3) sites B and C are available (prob = 0.4), so a decision must be made between sites B and C, and (4) neither site B nor site C is available (prob = 0.1). When both B and C are available at the
beginning of period 2 (case (3) above), then the decision proceeds as follows: if site B is
selected, there is an 0.8 probability that site C will remain unreserved at the end of period
2, in which case all 8 species survive, and a 0.2 probability that site C would be converted,
in which case only 7 species survive. Conditional on both sites B and C being unreserved
at the start of period two, choosing site B in period two yields an expected value of 7.8.
On the other hand, choosing site C in period 2 in this case yields an expected value of only
7.5. Therefore, it is optimal to choose B rather than C in the second period when both are
possibilities. The expected value when choosing site A in period 1 given that the optimal
choice will be made in period 2, is 7.02. Similar calculations give rise to the "Choose B"
and "Choose C" branches in figure 1, in expectation resulting in 6.74 and 6.55 species,
respectively.

As shown in figure 1, the optimal choice is to select site A in period one. Site A scores
well whenever it is combined with at least one other site. There is a fairly low probability
that both sites B and C will be converted in period one when not protected. Site C is
the most likely to remain in the unreserved category even without being protected so that
selecting site C is not a good choice. Since site C will likely remain, there is little value to
choosing site B because there is only one species (species 7) present at site B that is not
present at site C.

4 Larger examples

In this section, we develop examples with more sites, species and time periods to explore two
themes. In section 4.1, we investigate the importance of timing of site selection. We compare
outcomes under the dynamic reserve selection problem to the case where all sites may be
chosen prior to development and the case where development occurs prior to site selection.
Allowing the reserve planner to move prior to developers can be quite important to ensuring that high priority sites are conserved. In section 4.2, we compare the performance of several heuristic algorithms to the optimal dynamic reserve selection algorithm. Stochastic dynamic programming problems are often difficult to solve even for relatively small problems. Finding simple heuristics that are easy to solve and work reasonably well would be of great value in practical situations. We explore the implications adding more sites in section 4.3.

In these examples, we maintain the assumption that the planner chooses one site each period. We begin by considering examples with four periods \((T = 4)\), seven sites \((J = 7)\), and 28 species \((I = 28)\). We assume exactly one site is purchased each period. In order to explore a wide range of species distributions on the landscape \((A)\) and probabilities of development \((P)\), we use a Monte Carlo experiment with 100 realizations of \(A\) and \(P\), chosen from the following distributions:

\[
P_{jt} = y_j \quad \text{and} \quad \quad A_{ji} = \begin{cases} 
0 & \text{if } z_{ji} > .3, \\
1 & \text{if } z_{ji} \leq .3.
\end{cases} \tag{10}
\]

where \(y_j\) is a realization of a uniformly distributed random variable, \(Y \sim U[.2, .7]\), and \(z_{ji}\) is a realization of a uniformly distributed random variable, \(Z \sim U[0, 1]\).

### 4.1 Importance of Timing of Selections

In the base case where the conservation planner chooses one site each period, there is some risk that some desirable conservation sites will be developed prior to being selected. Suppose instead that the planner can select the same number of total sites as in the sequential choice but does so prior to any development. We will refer to this case as "front-end" selection. Alternatively, suppose that development occurs prior to reserve site selection. Note that
the case where the planner moves after development is really the same as the case with no site selection. After all development is determined no further sites will be developed regardless of whether they are reserved or not. Because of this, we will refer to this case as "no selection." In what follows we address how much difference the order of site selection versus development makes in terms of species conservation goals.

Table 2 shows results comparing the expected number of species conserved at the end of the four period time horizon for: a) the optimal sequential site selection assuming the planner chooses one site per period, b) the optimal front-end site selection assuming the planner can choose four sites immediately before any development occurs, c) the no selection alternative. The optimal front-end selection outperforms the optimal sequential allocation on all 100 Monte Carlo simulations. The expected percentage gain in species conserved from allowing the acquisition of four sites at the beginning—rather than four sites sequentially—ranges from 6.67% to 36.75%, with a mean of 19.25%. On the other hand, if no sites are reserved, on average only 6.73 species are conserved; a mean loss of 67.41%.

These results suggest the critical importance of timing. There exists a premium for making selections prior to development of prime conservation sites. Even having less budget overall, but getting it sooner in time, may make a conservation organization more effective. In some sense, our model can be viewed as a race between the conservation planner and development. Moving earlier, ahead of development, allows the conservation planner to win the race.

Our results differ from the option value literature on investment under uncertainty that demonstrates the value to avoiding making irreversible decisions prior to the resolution of uncertainty (Arrow and Fisher 1974, Henry 1974, Dixit and Pindyck 1994). However, in our case, it is always beneficial to make an irreversible decision to select a reserve site prior
to resolving uncertainty. This occurs because there are really two irreversible processes in our model: development and site selection. Uncertainty itself is over one of the irreversible processes: development. In our model, conservation objectives can only suffer with delay because some sites that would ideally be conserved may be lost to irreversible development. If on the other hand, there were uncertainty that resolved over time about which species were located in which sites, which species might survive on which sites, or about the costs associated with reserving particular sites, there would be option value to delaying site selection decisions.

### 4.2 Performance of heuristics

As pointed out in section 2, solving the stochastic dynamic integer programming problem with a large number of sites is computationally intensive. The dynamic programming approach is far more efficient than a brute-force forward optimization approach that analyzes all branches of a decision-tree. The difficulty of the brute-force method rises exponentially with both the time horizon ($T$) and the number of sites ($J$). The difficulty of the dynamic programming method rises exponentially only with $J$. Still, given the difficulty of finding optimal solutions in problems with a large number of sites, we wish to assess the performance of two simple heuristic choice algorithms. In the "naive myopic" algorithm, each selection adds the most species to the current pool of protected species, without regard for future allocation decisions or future development possibilities. In the "informed myopic"
algorithm, each selection adds the most species (in expectation) to the pool of protected species accounting for the possibility of development at the end of the current period. Both heuristics are easy to implement because they only look at one period not at the complete dynamic problem. Computing the expected value of species conserved under either heuristic remains a computationally intensive task, however.

The naive heuristic algorithm is optimal if there is only one selection to be made and if all unreserved sites will be developed. This latter assumption means that only species that occur in reserved sites count as reserved; species occurring in unreserved sites would be lost to development. This is an implicit assumption in much of the existing literature on reserve site selection. The algorithm is myopic for two reasons. First, it ignores the possibility that some unreserved sites may be spared from development. Second, it ignores the fact that other sites will be acquired in the future. The informed heuristic algorithm is similar to the naive algorithm, except that it relaxes the assumption that all unreserved sites will be developed at the end of the period. The informed myopic algorithm treats every period as if it were the final period of an optimal dynamic program (see equation (8)) so that some possibility that species in unreserved sites will survive is factored into the analysis. These myopic algorithms are similar to the ”expected greedy algorithm” where at each step the site that adds the greatest expected number of species to what is already conserved is selected (Polasky et al. 2000). The informed myopic rule is equivalent to a hotspot approach that looks to find the next site to protect based on current gaps in coverage and current threats of habitat loss.

In Table 3, we present the results of the Monte Carlo experiment comparing the performance of the heuristic algorithms with the optimal sequential algorithm. The table gives the mean and standard deviation (from 100 Monte Carlo simulations) of the expected num-
Table 3: Comparing optimal vs. heuristic algorithms: statistics for expected number of species conserved from 100 Monte Carlo simulations of $A$ and $P$.

<table>
<thead>
<tr>
<th></th>
<th>Optimal Sequential</th>
<th>Naive Myopic</th>
<th>Informed Myopic</th>
<th>No Selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>20.65</td>
<td>20.19</td>
<td>20.61</td>
<td>6.73</td>
</tr>
<tr>
<td>s.d.</td>
<td>1.73</td>
<td>1.79</td>
<td>1.75</td>
<td>2.07</td>
</tr>
</tbody>
</table>

Results from table 3 suggest that both myopic algorithms perform well compared to the optimal sequential algorithm and both significantly outperform the no-action alternative. The informed myopic algorithm outperforms the naive myopic algorithm and comes close to the performance of the optimal sequential algorithm. In percentage terms, using the naive myopic algorithm results in a species loss of between 0.00% and 10.21% with a mean of 2.23% when compared to the optimal sequential algorithm. The informed myopic algorithm - where the probabilities of development are taken into account - results in a small loss of between 0.00% and 2.75% with a mean loss of about 0.21%. These results are generated on a problem with a fairly small number of choice occasions ($T = 4$) and sites ($J = 7$). In the next subsection we explore the sensitivity of the above results to the number of choice occasions and sites.

4.3 Increasing the Number of Sites

Intuition suggests that when the number of sites ($J$) and the time horizon ($T$) are small, a good heuristic algorithm such as the informed myopic should perform reasonably well compared to the optimal sequential algorithm. This performance may deteriorate, however, as $J$ and $T$ increase because: (1) more sites imply more places that seem attractive through a myopic lens, and (2) more time periods suggest more opportunities to follow a misleading
Table 4: Mean of the expected number of species saved for the optimal sequential and informed myopic algorithms for $T = 1, 2, ..., 6$ and $J = 10$ and 7.

<table>
<thead>
<tr>
<th>Sites</th>
<th>Algorithm</th>
<th>$T = 1$</th>
<th>$T = 2$</th>
<th>$T = 3$</th>
<th>$T = 4$</th>
<th>$T = 5$</th>
<th>$T = 6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$J = 10$</td>
<td>Optimal Sequential</td>
<td>26.78</td>
<td>25.71</td>
<td>25.49</td>
<td>25.48</td>
<td>25.48</td>
<td>25.48</td>
</tr>
<tr>
<td></td>
<td>Informed Myopic</td>
<td>26.78</td>
<td>25.44</td>
<td>25.02</td>
<td>25.00</td>
<td>25.00</td>
<td>25.00</td>
</tr>
<tr>
<td></td>
<td>% Loss from Myopic</td>
<td>0.00%</td>
<td>1.05%</td>
<td>1.84%</td>
<td>1.89%</td>
<td>1.89%</td>
<td>1.89%</td>
</tr>
<tr>
<td>$J = 7$</td>
<td>Optimal Sequential</td>
<td>24.66</td>
<td>23.39</td>
<td>23.24</td>
<td>23.23</td>
<td>23.23</td>
<td>23.23</td>
</tr>
<tr>
<td></td>
<td>% Loss from Myopic</td>
<td>0.00%</td>
<td>0.30%</td>
<td>0.52%</td>
<td>0.47%</td>
<td>0.47%</td>
<td>0.47%</td>
</tr>
</tbody>
</table>

trajectory. To test this intuition, we run the following experiment. For $J = 10$ sites and $I = 28$ species we simulate $P$ and $A$ using the formulas in equations (9) and (10). From these, we compute the expected number of species conserved after $T = 1, 2, ..., 6$ time periods, for both the optimal sequential algorithm and the informed myopic algorithm. For comparison purposes, we next draw at random seven of the ten sites and solve for the expected number of species conserved for this $J = 7$ problem. This drawing procedure is repeated 100 times to give the expected number of species conserved for $J = 7$ site problems, where the sites are chosen randomly from the $A$ matrix for the 10 site problem above. This allows us to compare the results of a seven site example with results of a 10 site example, where the configuration of species on the landscape is similar in both cases.

Results of this experiment are given in table 4, which shows the expected number of species conserved under both the informed myopic and the optimal sequential algorithms for the $J = 10$ case, and the mean of the expected number of species conserved (over 100 Monte Carlo simulations) for the $J = 7$ case.

As shown in table 4, the expected loss from using the informed myopic algorithm rather than the optimal sequential algorithm grows as the number of sites increases in our example. The informed myopic algorithm results in an average loss of only about 0.47% in the Monte Carlo simulations.
Carlo experiment (for $J = 7$ and $T = 4$). By comparison, the expected loss was 0.21% in the Monte Carlo experiment from subsection 4.2 (this difference is caused by different methods of randomization of $A$ and $P$). The corresponding loss for the $J = 10$, $T = 4$ example is 1.89%. Thus, it appears as though the percentage loss of species from using the informed myopic algorithm instead of the optimal sequential algorithm increases with both $T$ and with $J$, a finding that should be tested in future work for much larger choices of the two parameters.

5 Example with Vertebrate and Probability of Development

Data from California

We now turn our attention to an empirical example of the dynamic reserve site selection model. An example using real data is useful both to illustrate how typical species and probability of development data integrate into an analysis of optimal dynamic reserve design and as an empirical trial of the theory presented above. We adapt and analyze the problem posed by Church et al. (1996), where Gap Analysis data are used to design a reserve for 333 native vertebrate species across 280 sites throughout southwestern California\(^1\). The sites are United States Geological Survey 7.5' quadrangles, and are typically 16000 hectares, though some coastal sites contain a smaller land area. The Church et al. paper shows that all 333 species can be represented by optimally selecting 12 sites.

The Church et al. analysis is static, and does not account for the chance of future species loss if a site is not immediately reserved. Our objective is to apply the dynamic reserve selection model to these data to determine the expected number of species covered.

\(^1\)The authors extend their gratitude to David Stoms for his assistance in identifying and compiling the data for this example.
Table 5: United States Geological Survey 7.5’ quadrangles in the 12 site solution covering 333 species in Church et al. (1996). The table gives the name of each site, the number of species present, and the probability of development.

<table>
<thead>
<tr>
<th>USGS Quad Name</th>
<th># Species Present</th>
<th>Probability of Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cuyama Peak</td>
<td>162</td>
<td>0</td>
</tr>
<tr>
<td>Liebre Mtn.</td>
<td>219</td>
<td>0</td>
</tr>
<tr>
<td>Newhall</td>
<td>199</td>
<td>0.496</td>
</tr>
<tr>
<td>Point Conception</td>
<td>238</td>
<td>0</td>
</tr>
<tr>
<td>Palm View Peak</td>
<td>222</td>
<td>0</td>
</tr>
<tr>
<td>Oxnard OE</td>
<td>232</td>
<td>0.521</td>
</tr>
<tr>
<td>Catclaw Flat</td>
<td>250</td>
<td>0.011</td>
</tr>
<tr>
<td>Forest Falls</td>
<td>167</td>
<td>0.197</td>
</tr>
<tr>
<td>Toro Peak</td>
<td>155</td>
<td>0</td>
</tr>
<tr>
<td>Morro Hill</td>
<td>196</td>
<td>0.180</td>
</tr>
<tr>
<td>Mount Laguna</td>
<td>208</td>
<td>0.122</td>
</tr>
<tr>
<td>National City</td>
<td>195</td>
<td>0.403</td>
</tr>
</tbody>
</table>

when reserve selection is sequential and species viability is threatened. As a measure of the probability of development we use 50 year projections of urban expansion developed for the “California Legacy Project”, a strategic statewide conservation planning process funded by the state of California (see Landis et al. 1998 for a description of the model used to generate urbanization projections). Although these projections were developed as estimates of the future level of urbanization throughout the state, for illustrative purposes, we interpret them as measures of the probability of development in each site. The name of each site, the number of species present, and the probability of development are given in table 5:

Note that five of the 12 sites face no threat of development (Cuyama Peak, Liebre Mountain, Point Conception, Palm View Peak, and Toro Peak). Under our model, species present on those sites will persist even if those sites are not selected. These five sites cover 282 species. There then remain 51 species that are in some need of protection. As in section 4, we assume that one site can be selected in each period (for up to 7 periods). Table 6
Table 6: Results for the California vertebrate example: the importance of timing.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Expected # Species (333 possible)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimal Sequential</td>
<td>325.6</td>
</tr>
<tr>
<td>Optimal Front-End</td>
<td>333</td>
</tr>
<tr>
<td>No Selection</td>
<td>306.3</td>
</tr>
</tbody>
</table>

Table 7: Results for the California vertebrate example: heuristic algorithm performance.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Expected # Species (333 possible)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimal Sequential</td>
<td>325.6</td>
</tr>
<tr>
<td>Informed Myopic</td>
<td>325.6</td>
</tr>
<tr>
<td>Naive Myopic</td>
<td>323.6</td>
</tr>
<tr>
<td>No Selection</td>
<td>306.3</td>
</tr>
</tbody>
</table>

compares the expected number of species conserved after 7 time periods under different assumptions about the timing of selection.

Even when optimal selections are made, there is a loss of about 7 species (2%) because only one site may be reserved each time period; non-selected sites are subject to development. In the case where no sites are selected, many species persist, but about 27 species (8%) are lost to development throughout the 7 year period. The relative performance of the optimal and myopic algorithms is illustrated in table 7. This problem is of similar size to those considered in section 4. However, the relative performance of the myopic algorithms is better using real data than it was using randomly generated data. At least in this example, a good heuristic algorithm can do extremely well in choosing sites to select. The informed myopic algorithm makes the same site selections and generates the same expected number of species conserved as does the optimal sequential algorithm. The naive myopic algorithm does nearly as well with an expected loss of only two species relative to the optimal solution. Because most of probabilities of development are low, the no selection alternative has an expected representation of 306.3 species after 7 periods.
6 Discussion

In this paper we analyzed a dynamic reserve site selection problem in which sites must be chosen sequentially because of budget constraints, with some sites being chosen after some of the threats of development are realized. A dynamic approach is realistic and generates potentially different recommendations about conservation priorities than does a static approach. Our results demonstrate the importance of considerations of timing. Choosing sites before conversion risk is realized yields much higher expected conservation payoffs than choosing sites afterwards. The results also show that standard static approaches do not do as well as solving the dynamic programming problem for an optimal solution. However, solving the dynamic programming problem may be quite difficult in large problems and there is value in looking for simple heuristics that perform reasonably well. For the examples considered here, certain types of static heuristic approaches generate small losses in the expected number of species conserved as compared to the optimal dynamic solution. In particular, the "informed myopic" algorithm, which accounts for both probability of conversion at all sites and species distributions across sites, typically results in losses on the order of between 0.21%-1.89% relative to the optimal solution; where the percentage loss appears to increase with the number of sites. These results are generated for relatively small problems. An open question is how well such heuristics would perform in problems with a large number of sites, time periods, and species.

Taking a dynamic approach opens up many interesting questions, only some of which are analyzed in this paper. The examples of the previous section used a simple budget constraint that did not incorporate heterogeneous land costs or the possibility of carrying budgets forward through time. The model described in section 2 included the ability to handle both of these complications. Incorporating realism about the budget constraint
adds to the computational difficulty of solving the model. Formally, doing so adds a state variable \((B_t)\). The degree of additional computational difficulty depends upon the number of different values \(B_t\) may take. Constraining \(B_t\) to a small number of potential values would only complicate the analysis slightly. On the other hand, allowing \(B_t\) to take any dollar amount between 0 and the maximum budget would significantly complicate numerical computations. We feel that the analysis of budget constraints in dollar terms, rather than in terms of number of sites, is an important next step in the analysis. Ando et al. (1998) and Polasky et al. (2001) have shown that budget constrained solutions that incorporate difference in land costs across sites often lead to distinctly different priorities than do site constraint solutions that ignore such differences.

In our examples, we assumed that the probability of conversion and the land price for any site at any time was independent of what other sites had been previously reserved or developed. It is highly likely that both the probability of conversion and land price for a particular site depends upon what has happened on nearby sites. For example, if a reserve is established on a neighboring site, the desirability of development for a given site might increase, which may increase both the probability of development and the price. Development also depends on the presence of infrastructure (sewer lines, roads, etc.), which introduces spatial correlation in the conversion probabilities and land prices. In principle, as illustrated in section 2, the stochastic dynamic programming approach can incorporate conversion probability functions and cost functions that depend upon the pattern of development and reserve selection. In this case, a planner choosing a particular reserve site would take account not only of the direct effect of choosing that site as a reserve but also the effect that doing so would have on altering probabilities of conversion and land prices on other potential reserve sites. Making land prices and probabilities endogenous in this
way would likely increase the value taking a dynamic approach.

Consideration of endogenous land prices and development probabilities raises further strategic issues. Because conservation designation often increases the property values of adjacent lands, conservation planners and developers can profit from information about future reserve designation. For the conservation planner, it might make sense to wait to purchase a large block all at once rather than trying to purchase small blocks sequentially only to see either the price of the remaining blocks bid up or sites developed before they are able to be reserved. In fact, the conservation planner may desire to purchase more land than they wish set aside in a reserve. By doing so, the conservation planner can sell land adjacent to the reserve and capture the rents that the biological reserve has created. If these rents are large, as they may be if there are valuable wildlife viewing or hunting opportunities, sales of adjacent land many pay a significant fraction of the cost of setting up a conservation reserve (Heal 1998). On the other hand, developers can also play this game. Developers may try to capture rents by purchasing land that they think will be of high priority to conservation groups or will be adjacent to reserves. Conservation organizations such as the Nature Conservancy are well aware of these strategic considerations and in some cases try to keep information private about the location of future acquisitions. We plan to analyze a strategic game between a conservation planner and developers in future work.

In this paper we assumed that a conservation agency could purchase any site in any period. In reality, there may be only limited windows of opportunity when a particular site is up for sale. We could incorporate this fact with a relatively minor modification of the existing model. The probability of conversion in our approach could be reinterpreted as the probability that the site will come up for sale in a given period. Then, a site cannot be reserved (or developed) unless it comes up for sale. When it does come up for sale, the
conservation agency faces the decision of whether or not to purchase the site knowing there is a high probability the site will be developed if it is not reserved. Conservation agencies recognize this formulation as being a close approximation to the real conservation problem they face. We plan on analyzing such a problem in future work.

It is also possible to consider restoration activities and the possibility of the conservation agency selling off previously acquired sites. In the approach considered in this paper, there were two types of irreversible decisions. First, a site that was developed was considered developed forever after. Second, the decision to reserve a site meant it was reserved forever after. In principle, both types of decisions could be reversed at some cost. To ”unreserve” a site is as simple as allowing a conservation agency to sell back a previously purchased site. A conservation agency may wish to do so if the pattern of development has made a particular site of lesser conservation value or if it has learned over time that a combination of other sites can protect biodiversity more effectively. To ”undevelop” a site is potentially more difficult. The relatively new field of restoration ecology focuses on this issue: how to restore ecosystems that have been heavily modified by previous human activities. Allowing reversibility of both the development and reserve activities could potentially change the strategy of the conservation agency in important ways. Analysis of the problem with reversible development and reserve decisions is an important topic for future research.

Considering whether species are likely to persist in reserves and unreserved areas is a vitally important consideration in conservation planning, and one that is often not adequately handled. Different species have different home range requirements. Some sites are perfectly suitable for supporting populations of some species but may be inadequate for supporting viable populations of other species. Further, the probability of survival for a species may depend on the entire pattern of reserves and unreserved areas rather than on whether a
particular site is developed or not. Stochastic environmental or demographic events may cause a local population of a species to die out. However, if the population of the species inhabiting a site is part of a larger meta-population, then a site may be recolonized from other sites as long as there is sufficient ability of populations to move between sites. In addition, there are predator-prey or competitive relationships among species that make the probabilities of survival among species non-independent. While introducing dependence in survival probabilities across sites or species complicates the analytics of the model, the more difficult challenge at present is our lack of ecological understanding that would allow modeling non-independent probabilities. Incorporating species survival probabilities into reserve site selection and other large-scale conservation planning exercises is an important topic on which more research is needed.

A further source of uncertainty is that there is often incomplete species range information. In our notation, this would make the elements of the matrix $A$ uncertain. Polasky et al. (2000), Camm et al. (2001), and Haight et al. (2000) analyze a static reserve site selection problem with incomplete information about species occurrences.

Finally, we have not considered other ecological benefits of conservation beyond the benefits of species conservation, such as the continued provision of ecosystem services (water purification, flood mitigation, nutrient cycling, climate regulation). With consideration of the benefits ecological services, the ”cost” of conserving land would be amended to be the land price minus the value of ecosystem services provided by the site. Each site can be viewed as giving some contribution to the value of ecosystem services. However, ecosystem functions typically depend the pattern of development over a region. Changing land use on any particular site will then have effects that spread more broadly through the ecosystem, which would change the contribution of other sites not only of the site on which the land
use changed. Just like the consideration of endogenous land prices, ecosystem service values would typically depend on the entire pattern of development.

References


26


at the end of the horizon. Choosing A in period 1 gives the highest expected number of species conserved (7.02 < 6.74 < 6.22).

Figure 1: Expected payoffs and probabilities for the three species exercise. In the first period, A, B, or C can be played into reserve.