University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Papers in Natural Resources | Natural Resources, School of

7-3-2006

Planning for robust reserve networks using uncertainty analysis

Atte Moilanen
Metapopulation Research Group, Department of Biological and Environmental Sciences, P.O. Box 65, FI-00014 University of Helsinki, Finland

Michael Runge
USGS Patuxent Wildlife Research Center, 12100 Beech Forest Road, Laurel, MD 20708, United States

Jane Elith
School of Botany, University of Melbourne, Parkville, Vic. 3010, Australia

Andrew Tyre
University of Nebraska - Lincoln

Yohay Carmel
Israel Institute of Technology, Haifa 32000, Israel

See next page for additional authors

Follow this and additional works at: http://digitalcommons.unl.edu/natrespapers

Part of the Natural Resources and Conservation Commons

Moilanen, Atte; Runge, Michael; Elith, Jane; Tyre, Andrew; Carmel, Yohay; Fegraus, Eric; Wintle, Brendan A.; Burgman, Mark A.; and Ben-Haim, Yakov, "Planning for robust reserve networks using uncertainty analysis" (2006). *Papers in Natural Resources.* Paper 102.
http://digitalcommons.unl.edu/natrespapers/102

This Article is brought to you for free and open access by the Natural Resources, School of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Papers in Natural Resources by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Authors
Atte Moilanen, Michael Runge, Jane Elith, Andrew Tyre, Yohay Carmel, Eric Fegraus, Brendan A. Wintle, Mark A. Burgman, and Yakov Ben-Haim

This article is available at DigitalCommons@University of Nebraska - Lincoln: http://digitalcommons.unl.edu/natrespapers/102

ECOLOGICAL MODELLING 199 (2006) 115–124

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/ecolmodel

Planning for robust reserve networks using uncertainty analysis

Atte Moilanen[a,*], *Michael C. Runge*[b], *Jane Elith*[c], *Andrew Tyre*[d], *Yohay Carmel*[e], *Eric Fegraus*[f], *Brendan A. Wintle*[c], *Mark Burgman*[c], *Yakov Ben-Haim*[e]

[a] *Metapopulation Research Group, Department of Biological and Environmental Sciences, P.O. Box 65, FI-00014 University of Helsinki, Finland*
[b] *USGS Patuxent Wildlife Research Center, 12100 Beech Forest Road, Laurel, MD 20708, United States*
[c] *School of Botany, University of Melbourne, Parkville, Vic. 3010, Australia*
[d] *School of Natural Resource Sciences, University of Nebraska-Lincoln, 202 Natural Resources Hall, East Campus Lincoln, NE 68583-0819, United States*
[e] *Faculty of Mechanical Engineering, Technion, Israel Institute of Technology, Haifa 32000, Israel*
[f] *NCEAS National Center for Ecological Analysis and Synthesis, 735 State Street, Suite 300, Santa Barbara, CA 93101, United States*

ARTICLE INFO

Article history:
Received 21 September 2005
Received in revised form 19 June 2006
Accepted 3 July 2006
Published on line 17 August 2006

Keywords:
Reserve selection
Site selection algorithm
Conservation planning
Uncertainty analysis
Information-gap decision theory

ABSTRACT

Planning land-use for biodiversity conservation frequently involves computer-assisted reserve selection algorithms. Typically such algorithms operate on matrices of species presence–absence in sites, or on species-specific distributions of model predicted probabilities of occurrence in grid cells. There are practically always errors in input data—erroneous species presence–absence data, structural and parametric uncertainty in predictive habitat models, and lack of correspondence between temporal presence and long-run persistence. Despite these uncertainties, typical reserve selection methods proceed as if there is no uncertainty in the data or models. Having two conservation options of apparently equal biological value, one would prefer the option whose value is relatively insensitive to errors in planning inputs. In this work we show how uncertainty analysis for reserve planning can be implemented within a framework of information-gap decision theory, generating reserve designs that are robust to uncertainty. Consideration of uncertainty involves modifications to the typical objective functions used in reserve selection. Search for robust-optimal reserve structures can still be implemented via typical reserve selection optimization techniques, including stepwise heuristics, integer-programming and stochastic global search.

© 2006 Elsevier B.V. All rights reserved.

1. Introduction

In the last decade, much scientific effort attempting to cope with biodiversity loss has been directed into reserve selection studies (see Pressey, 1999; Margules and Pressey, 2000; Cabeza and Moilanen, 2001; Moore et al., 2003; Noss, 2003; Cabeza et al., 2004; Williams et al., 2004 for reviews). A common feature of these studies is the search for the set of candidate sites that optimizes a performance function (e.g., number of species adequately covered by reserves) while observing con-

* *Corresponding author*. Tel.: +358 9 191 57753; fax: +358 9 191 57694.
E-mail addresses: atte.moilanen@helsinki.fi (A. Moilanen), michael_runge@usgs.gov (M.C. Runge), j.elith@unimelb.edu.au (J. Elith), atyre2@unl.edu (A. Tyre), yohay@tx.technion.ac.il (Y. Carmel), fegraus@nceas.ucsb.edu (E. Fegraus), brendanw@unimelb.edu.au (B.A. Wintle), markab@unimelb.edu.au (M. Burgman), yakov@techunix.technion.ac.il (Y. Ben-Haim).
0304-3800/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecolmodel.2006.07.004

straints (typically total area or cost of protected land). Many variants of site selection algorithms have been produced to solve such optimization problems (Csuti et al., 1997; Pressey et al., 1997; Cowling et al., 2003; Moilanen, 2005a,b; Moilanen et al., 2005). However, a common feature of all of these methods is that optimizations are performed and evaluated as if there is no uncertainty in the species distributions that underlie them (Gaston and Rodrigues, 2003; Rodrigues et al., 2004; Wilson et al., 2005).

For example, consider the much-studied single-representation problem, which seeks "the least expensive set of sites that includes at least one occurrence of each species" (Csuti et al., 1997; Pressey et al., 1997). This formulation implicitly assumes that (i) observations or predictions of species presence are correct, (ii) species distributions do not change and (iii) present occurrence indicates long-term persistence in sites. If there are errors in observation, prediction or characterization of spatial population dynamics, then these assumptions are no longer valid. As a consequence, using the single-representation formulation for reserve planning may maximize within-reserve extinction rates, especially if selection units are small (Cabeza and Moilanen, 2003). Thus, optimizing performance (minimizing reserve cost in the single-representation problem) leads to solutions that have no robustness to uncertainty in the species distribution information.

Uncertainty analysis provides a consistent framework for understanding potential consequences of errors in inputs. Here we apply one particular branch of decision theory, information-gap theory (Ben-Haim, 2006) to conservation planning. Insights gained from this theory indicate that under severe uncertainty, the solution that maximizes a performance function is in general unlikely to be a tenable solution. In particular, a basic theorem of info-gap theory asserts that utility (or performance) optimization is equivalent to minimization of immunity to uncertainty. When maximizing performance one employs best-estimates, the so-called nominal models (which may be probabilistic) and ignores the potential structural or factual errors. The optimal solution will be prone to failure due to errors in the underlying distribution models because the optimization process is completely reliant on these models. This is the case with the single-representation problem.

A reserve solution that is optimal with respect to the nominal estimate is often not robust to uncertainty, whereas another solution may be sub-optimal when evaluated using the nominal distributions but more robust to uncertainty in these distributions. In fact, a robust solution that is apparently sub-optimal may be preferable to an optimal solution that has no robustness to potential error. We call such solutions *robust-optimal*. Robustness is a property of the solution that expresses immunity to uncertainty in the underlying database and fitted models. More concretely, we may not be sure that an entire list of species that are to be protected in a specific set of conservation areas will actually persist within the areas to be conserved. However, we want to be reasonably certain that at least, say, 90% of those species will actually persist within the region. Furthermore, we would like to be confident that this outcome will be achieved even if the models upon which our decision is based are substantially flawed. This approach to reserve planning is entirely different from the common approach, which chooses the optimal set of conservation areas based on best-estimated models, while risking the possibility of large failure due to model error.

Decision analytic methods, including uncertainty analysis, are often employed in population viability analyses (see e.g., Possingham et al., 2001a,b; Drechsler and Burgman, 2004; Regan et al., 2005). However, PVAs typically concern detailed population dynamical analyses for a couple of conservation scenarios and one or a few species. In contrast, reserve selection may involve large optimization problems with large-scale spatial planning and many species. It is not surprising that reserve design has so far largely ignored uncertainty, because in most cases uncertainty is not even acknowledged let alone quantified in the habitat models that are used for predicting underlying species distributions. Methods do exist for producing uncertainty estimates (e.g., Elith et al., 2002; MacKenzie et al., 2003; Van Niel et al., 2004; Wilson et al., 2005), but they are rarely implemented. We believe that part of the problem is that even if methods are available, there are few tools for incorporting uncertainty into decision-making (Burgman et al., 2006).

Fig. 1 shows a categorization of the kinds of sites that may be available for conservation. Areas that are certain to have high biological value are most important for conservation. Areas that are certain to have low biological value should be avoided. The robustness analysis applies to areas that have high estimated value, but this value is uncertain. These areas have potential for negative surprises for conservation, and this potential for negative surprises should be controlled via uncertainty analysis. The fourth category includes areas that have low value, but we are unsure of this. These could be areas that, for example, have been poorly surveyed. These areas have potential for positive surprises, which can be handled via the concept of *opportuneness* in the context of the information-gap theory (Ben-Haim, 2006).

The goal of this article is to provide terminology and a methodological basis for robust reserve planning. We describe in detail a simple application of information-gap theory to reserve planning based on probabilities of occurrence. We also describe plausible uncertainty models for some more complex conservation planning applications.

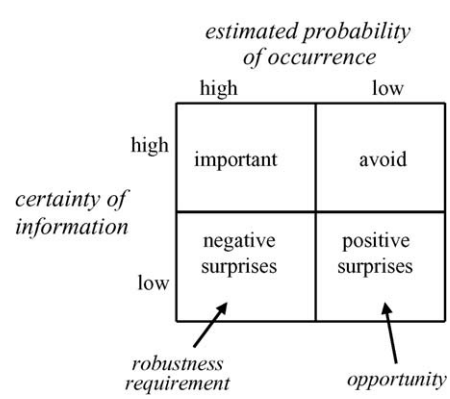

Fig. 1 – A categorization of four kinds or sites with different implications for conservation decision making.

Table 1 – Symbols and notation

Symbol	Explanation
s	Index for species
c	Index for cell (i.e., selection unit, site)
m	Index for model
C	Set of indexes for cells, which together define a reserve selection
p_{scm}	True probability of occurrence for species s in cell c according to model m
$\tilde{p}_{scm}$	Nominal distribution: our best estimate for p_{scm}
w_{scm}	Uncertainty weight corresponding to $\tilde{p}_{scm}$. It could be, for example, the standard error or length of the confidence interval for $\tilde{p}_{scm}$
$\tilde{p}$	The vector of $\tilde{p}_{scm}$ values
α	Horizon of uncertainty
$U(\alpha, \tilde{p})$	Uncertainty model centered around $\tilde{p}$ with horizon of uncertainty α
T_s	Target level of representation for species s
$R_{sm}(C, p)$	Representation level of species s in reserve C according to model m and probability vector p
$\hat{\alpha} = \hat{\alpha}(C, T)$	Overall robustness of reserve design C over all models and species with targets T
$\hat{\beta} = (C, T)$	Opportuneness value of reserve design C over all models and species with targets T
C^*	Robust-optimal solution

2. Methods

2.1. Overview

Below we define three formulations (performance functions) for the reserve selection problem. We then describe the information-gap decision theoretic approach to uncertainty analysis and its application to reserve planning. The concepts of uncertainty model, robustness, opportuneness, and a robust-optimal solution are central to this work. Symbols and notation are summarized in Table 1.

2.2. Notation and objective

Indices s, c and m identify species, cells in the spatial domain (often referred to as sites, patches or areas), and models. Sets of cells, which constitute potential reserves (or reserve networks), are identified by sets of indices denoted C. Thus C_1, C_2, etc. represent alternative reserve selections. Not all sets of cells are feasible choices due to economic, social or political constraints. The collection of all feasible reserve selections is denoted C.

Species s occurs in spatial cell c with probability p_{scm} under model m. The nominal, i.e., best available estimate of p_{scm} is given by $\tilde{p}_{scm}$. $U(\alpha, \tilde{p})$ is an info-gap model for the uncertain deviation between the true probability distribution and the best estimate (Ben-Haim, 2006).

$R_{sm}(C, p)$, referred to as a representation, is a performance function (e.g., abundance, population count, number of occurrences), for species s based on model m in reserve network C, where the probability in each cell is given by p_{scm}. These performance functions are defined so that a large value is preferable to a small value. Furthermore, we will express our performance requirements by stating minimal acceptable target values of performance. That is, reserve C has acceptable performance, based on the probability distribution p_{scm}, if $R_{sm}(C, p)$ exceeds the target value T_s.

Before worrying about uncertainty, our reserve-planning task is to choose a reserve structure (including N cells) whose biodiversity performance value is greatest. Intuitively, when planning a robust reserve, we begin to worry about uncertainty and deal with it by choosing a reserve structure whose biodiversity performance value is reasonably certain to be good enough, and whose immunity to uncertainty is as large as possible, given certain minimal required outcomes.

We next describe some objectives (performance functions) used commonly in reserve selection. Then we describe an info-gap model for uncertainty in the probability of occurrence distribution of a species. Finally, we derive the robustness functions for potential reserve designs.

2.3. Performance functions for reserve selection

The following discussion is, for convenience, in terms of species, which is our shorthand for any biodiversity feature (species, land cover type, vegetation community type) of conservation interest.

Expected number of occurrences (e.g., ReVelle et al., 2002; Cabeza, 2003). The total statistical weight of each species s in cells indexed by C, based on model m, is:

$$R_{sm}(C, p) = \sum_{c \in C} p_{scm}, \tag{1a}$$

In this formulation p_{scm} could also denote, for instance, area coverage of a vegetation community type. A typical objective would be to obtain with least cost a given proportion of occurrences for each species.

Probability of having at least one occurrence (Haight et al., 2000; Polasky et al., 2000; Araujo and Williams, 2000). The probability of each species s occurring at least once in reserve C is, under model m:

$$R_{sm}(C, p) = \left[1 - \prod_{c \in C}(1 - p_{scm})\right]. \tag{1b}$$

In this formulation a probability target could be set for each species and the question would be which set of sites achieves this target with least cost.

Reserve value via benefit functions (Arponen et al., 2005). Reserve value $V(C)$ is a sum over values of individual species:

$$V(C) = \sum_s V_s(C) = \sum_s v_s f(R_{sm}(C, p)), \tag{1c}$$

where v_s is the weight of species s, and $f()$ is an increasing function of representation, which can be defined for example via Eq. (1a). The objective is to achieve highest possible value for the reserve network given a constraint on the total value of land that can be obtained. In the benefit function approach, the target T_s is actually a reference level at which the species gets value v_s, i.e., $f(T_s) = 1$. Importantly, both over- and under-representation (with respect to T_s) are valued—a higher representation always implies higher value. A reasonable form for $f()$ would be a monotonically increasing function,

which approaches an asymptote with high R_{sm} and which has $f(0) = 0$. Note that a typical reserve selection formulation with a target level set T_s for the expected number of occurrences (Eq. (1a)), $R_{sm}(C, p) > T_s$, actually is a special case of (1c) with $v_s = 1$ and $f()$ being a unit step function with the step located at T_s.

2.4. *Info-gap uncertainty models*

Info-gap models can be used to represent a wide variety of prior information about uncertain probabilities. In this section we will discuss three different info-gap models, addressing absolute errors, fractional errors, and asymmetrical error intervals.

Our best estimate of the probability of occurrence of species s, in cell c, based on an ecological model m, is $\tilde{p}_{scm}$. That we use probabilities implies that the actual occurrence of species will be a stochastic event. Additionally, these probabilities themselves are highly uncertain—we do not know their true values, and an info-gap model is used to represent this uncertainty. An info-gap model $U(\alpha, \tilde{p})$ for the probabilities of model m is a family of nested sets of probability distributions p_{scm} centered on the best estimate $\tilde{p}_{scm}$, where α is a parameter describing a *horizon of uncertainty*. Three different structures will be suggested for the uncertainty model, but we first discuss the intuitive idea.

Even though $\tilde{p}_{scm}$ is our best estimate, it could err by some quantity, say by α. Letting p_{scm} be the correct probability, then $|\tilde{p}_{scm} - p_{scm}| \le \alpha$. Under severe uncertainty (Ben-Haim, 2006), we do not know the horizon of uncertainty, that is, we do not have a meaningful estimate of the error, but info-gap theory still allows us to specify an uncertainty model. This uncertainty model is a family of nested intervals of possible p_{scm} values:

$$|\tilde{p}_{scm} - p_{scm}| \le \alpha, \quad \alpha \ge 0. \tag{2}$$

These ideas can be formalized in various ways. The following envelope-bound model will often be useful:

$$U(\alpha, \tilde{p}) = \{0 \le p_{scm} \le 1 : |\tilde{p}_{scm} - p_{scm}| \le \alpha w_{scm}, \text{ for all } s, c, m\}, \quad \alpha \ge 0. \tag{3}$$

The set $U(\alpha, \tilde{p})$ contains all probabilities $p_{scm} \in [0, 1]$ whose deviation from the best estimate, $\tilde{p}_{scm}$, is bounded. Importantly, the "uncertainty weights" w_{scm} modulate the fractional errors. A large value of w_{scm} is chosen for particular indices, scm, if the corresponding probability p_{scm} tends to vary more, or be less certain than others. Conversely, small uncertainty weights indicate relatively reliable probability estimates. For example, when confidence intervals are available for the probability estimates from model m, one can choose each uncertainty weight proportional to the size of the (lower half of the) respective confidence interval. The bound is determined by the horizon of uncertainty, α, which is unknown. Thus the info-gap model is not a single set, but rather a family of nested sets.

Other info-gap models can be useful. In Eq. (3), the terms αw_{scm} represent absolute error in the probability. A typical alternative is to consider uncertain fractional error:

$$U(\alpha, \tilde{p}) = \left\{ p_{scm} \ge 0 : \left| \frac{p_{scm} - \tilde{p}_{scm}}{\tilde{p}_{scm}} \right| \le \alpha w'_{scm}, \text{ for all } s, c, m \right\}, \quad \alpha \ge 0, \tag{4}$$

which is equivalent to Eq. (3) with $w_{scm} = w'_{scm}\tilde{p}_{scm}$.

The info-gap models in Eqs. (3) and (4) entail symmetric uncertainty intervals (except for the non-negativity requirement). Asymmetric uncertainty intervals can be introduced with

$$U(\alpha, \tilde{p}) = \left\{ 0 \le p_{scm} \le 1 : \tilde{p}_{scm}(1 - \alpha w_{L,scm}) \le p_{scm} \le \tilde{p}_{scm}(1 + \alpha w_{U,scm}), \text{ for all } s, c, m \right\}, \quad \alpha \ge 0 \tag{5}$$

in which w_L and w_U correspond to lengths of lower and upper halves of confidence intervals, respectively. An asymmetric interval may be appropriate when working with bounded quantities, such as probabilities. Other more complex models deal with dependencies between parameters and other features of underlying models (Ben-Haim, 2006).

2.5. *Robustness*

The only probability distribution that we know is that of the best estimate, $\tilde{p}_{scm}$. Based on this distribution, we would choose the reserve whose robustness is greatest for a specific performance requirement.

Consider first a performance function ((1a) or (1b)), $R_s(C, p)$, for a given species s. We wish to evaluate a potential reserve, defined by its set of indices C. The reserve C is acceptable for species s (based on the best-estimated probability distribution $\tilde{p}_{sc}$), if the performance function exceeds the target value, $R_s(C, \tilde{p}) - T_s \ge 0$.

The *robustness question* for species s in this reserve is: how wrong can $\tilde{p}$ be, without jeopardizing the required performance of reserve C? The robustness, $\hat{\alpha}_s(C, T)$, of species s in reserve C given performance target T_s, is the greatest horizon of uncertainty up to which C has adequate performance:

$$\hat{\alpha}_s(C, T) = \max \left\{ \alpha : \left(\min_{p \in U(\alpha, \tilde{p})} [R_s(C, p) - T_s] \right) \ge 0 \right\}. \tag{6}$$

$\hat{\alpha}_s(C, T)$ trades robustness to uncertainty against the demanded performance: as the target value T_s gets larger (implying stricter ecological requirements) the robustness $\hat{\alpha}_s(C, T)$ gets smaller. High performance is more vulnerable to epistemic limitations (e.g., model uncertainty or data error) than low performance. Stated differently, the attainment of demanding conservation goals stretches the probability data $\tilde{p}_{sc}$ to the limit; the vulnerability to errors increases as the demands on the data-base grow.

In practice this means that in our reserve selection formulations (Eqs. (1a) and (1c)) the optimal reserve structure based on nominal distributions is likely to have zero robustness for at least one species. Assume, for example, a reserve selection problem where a target number of populations has been set for each species, and the aim is to achieve this representation

with minimal cost (area). The optimal solution for this problem is by definition such that no sites can be removed from the solution without causing a failure to meet the target for at least one species. If the solution consists of many small selection units (as it would in real landscape-level planning), some species will be just above the representation target level. Thus, this solution would be robust to little uncertainty. Any error in the distribution information for a species represented just at the required target level may cause the target to be failed for that species.

We next extend the robustness defined in Eq. (6), by considering multiple species s and alternative (species-specific) models m. The reserve indexed by C is acceptable, based on the best estimated probability distribution $\tilde{p}$, if the performance function exceeds the target values, $R_{sm}(C, \tilde{p}_{sm}) - T_{sm} \geq 0$, for each of species s and each model m. Use of multiple models may be reasonable for example when AIC-based model fitting (Burnham and Anderson, 2002) suggests multiple plausible, yet structurally different models.Now the robustness of reserve C, given performance targets T_{sm}, is the greatest horizon of uncertainty up to which C has adequate performance on all species and with all models:

$$\hat{\alpha} = \hat{\alpha}(C, T) = \max\left\{\alpha : \left(\min_{s,m}\left[\min_{p \in U(\alpha, \tilde{p}_{sm})} R_{sm}(C, p) - T_{sm}\right]\right) \geq 0\right\}. \tag{7}$$

The overall robustness of the reserve plan $\hat{\alpha}$ trades off against the demanded performance targets T_{sm} in the same way as $\hat{\alpha}_s(C, T)$: robustness goes down as the target values go up. Examination of Eq. (7) reveals that $\hat{\alpha}$ equals the robustness of the least robust species-model combination:

$$\hat{\alpha}(C, T) = \min_{s,m} \hat{\alpha}_{sm}(C, T) \tag{8}$$

where $\hat{\alpha}_{sm}$ denotes the robustness of species s using model m. We will prefer reserve C over reserve C', if the former has greater robustness to uncertainty in the probabilities, at the same target levels T:

$$C \succ C' \quad \text{if } \hat{\alpha}(C, T) > \hat{\alpha}(C', T). \tag{9}$$

The strength of the preference for C increases as the magnitude of the difference between the robustness values increases. The info-gap robust-optimal reserve, for performance targets T, is the set of cells whose robustness is greatest:

$$C^*(T) = \arg\max_{C} \hat{\alpha}(C, T). \tag{10}$$

We can "read" this relation as: the robust-optimal reserve is the reserve structure C^*, which maximizes overall robustness $\hat{\alpha}(C, T)$.

Computationally, it is relevant to observe that the performance of our reserve is lowest when the probabilities are at the lower bounds of the intervals defined by any of the Eqs. (3)–(5). Thus, for any given target T and reserve structure C, $\hat{\alpha}_{sm}(C, T)$ may be calculated by using binary search to pinpoint the exact value of α, for which $R_{sm}(C, p) - T_{sm} = 0$ for at least one species. Then, the robustness of the reserve is a minimum over these species-specific values (Eq. (8)). If the robustness is <0, then reserve C fails to achieve the given target for at least one species.

2.6. *Opportuneness*

In info-gap terminology, "opportuneness" is defined as the lowest level of uncertainty that must be present for a sweeping success to be possible. Intuitively, locations with very uncertain biological value could be very bad but also possibly very good. Opportuneness aims at exploiting the potential for better than expected biological outcome latent in such locations. This is expressed by the opportuneness function (the dual of the robustness function):

$$\hat{\beta}(C, T) = \min\left\{\alpha : \left(\min_{s,m}\left[\max_{p \in U(\alpha, \tilde{p}_{sm})} R_{sm}(C, p) - A_s\right]\right) \geq 0\right\}, \tag{11}$$

where $\hat{\beta}$ is the opportuneness, and A_s is an aspiration level for species s. The aspiration level would be set to a value that indicates a very good conservation outcome.

Low opportuneness means that the nominal estimates must be severely wrong before substantial success is possible, which is presented by a large value of $\hat{\beta}(C, T)$. In contrast, high opportuneness means that only small errors in nominal estimates are required for success to be possible, represented by a small value of $\hat{\beta}(C, T)$. Different reserve structures will have different robustness and opportuneness characteristics. In reserve planning, it may be a good strategy to sacrifice some biological performance or some robustness in return for high opportuneness. For example, an area that is remote and poorly surveyed may have high uncertainty about the biological value residing there. On the other hand, high uncertainty implies the possibility of the area proving much more valuable than what is presently assumed.

3. Results

3.1. *Simple reserve planning for robustness using the expected number of occurrences*

For the purpose of demonstration, consider the following reserve design problem. Fifteen sites are available for inclusion in a reserve, and the political and economic exigencies allow for the selection of three of those sites. The decision-makers, after careful deliberation, have indicated that the goal of planning is to maximize the sum of the probability of occurrence (expected number of populations) across three species in the reserve sites (performance function of Eq. (1a)). Estimates of the probability of occurrence for each species at each site are available, as are measures of uncertainty (Table 2).

The optimal solution to this problem can be found by inspection, by first considering the situation where the target level for all species is set to 1.0. Species Z is found only at sites 10–12, and the solution requires taking any two of these sites to achieve the target. Since species X is not well represented in these sites, the third site in the reserve should have a

Table 2 – Data for reserve selection example

Site (c)	Species X			Species Y			Species Z		
	p_{Xc}	$w_{Xc,L}$	$w_{Xc,U}$	p_{Yc}	$w_{Yc,L}$	$w_{Yc,U}$	p_{Zc}	$w_{Zc,L}$	$w_{Zc,U}$
1	0.9	0.1	0.05	0.5	0.1	0.1	0	0	0.5
2	0.95	0.2	0.03	0.5	0.5	0.3	0	0	0
3	0.85	0.1	0.04	0.5	0.5	0.5	0	0	0
4	0.8	0.1	0.08	0.4	0.1	0.3	0	0	0
5	0.7	0.3	0.1	0.4	0.3	0.1	0	0	0
6	0.6	0.3	0.2	0.5	0.5	0.3	0	0	0
7	0.5	0.3	0.3	0.5	0.1	0.4	0	0	0
8	0.4	0.5	0.5	0.4	0.5	0.5	0	0	0
9	0.4	0.1	0.1	0.4	0.1	0.2	0	0	0
10	0.3	0.1	0.15	0.5	0	0.1	0.4	0.1	0.4
11	0.1	0	0.05	0.6	0	0.2	0.7	0.3	0.2
12	0.1	0	0.1	0.3	0	0.1	0.9	0.5	0.1
13	0	0	0.2	0.4	0	0.05	0	0	0
14	0	0	0.4	0.4	0	0.1	0	0	0
15	0	0.2	0.2	0.4	0	0.3	0	0	0

Nominal probabilities of occurrence (p), and measures of uncertainty for three species at fifteen candidate sites. Non-symmetric error intervals are given with w_L and w_U giving hypothetical relative error rates related to lower and upper halves of confidence intervals, respectively.

high probability for species X, which suggests one of sites 1–4. Further inspection reveals that site 4 should never be chosen because site 1 has higher nominal probabilities and both lower and upper bounds of site 1 exceed those of site 4. Tentatively, one could pick site 2, which has highest nominal probability for species X. The reserve design [2, 10, 11] has a minimum aggregate sum of probability of 1.1 (it is limited by species Z). We can actually improve the performance by selecting site 12 instead of 11: reserve design [2, 10, 12] achieves a minimum aggregate sum of 1.3, which turns out to be the optimal reserve design, for this particular performance function, under perfect certainty. Note, however, that this solution is quite sensitive to uncertainty; in particular, the uncertainty associated with species Y at site 2 is quite high. If uncertainty is a concern, it might be better to choose reserve design [1, 10, 12]. This entails a small sacrifice in the probability of occurrence of species X, but affords some protection from uncertainty about the probability of occurrence of species Y. Thus, considering the uncertainty in the probabilities of occurrence motivates the need for info-gap decision theory.

The solution to this reserve design problem using info-gap theory can be found by embedding the performance function (Eq. (1a)) into Eq. (6), considering the uncertainty model (Eq. (3)), and calculating the robustness, $\hat{\alpha}$, for a range of target values T for the performance criterion. (Note that we are using only one model here, and the target value is not species-specific). The full info-gap solution as a function of target level is shown in Fig. 2. The outer boundary of this figure, shown in bold, includes the Pareto-optimal solutions (e.g., Chankong and Haimes, 1968)—the solutions that cannot be improved on without either decreasing the reservation targets of the most vulnerable species or the robustness of the solution.

Under perfect certainty, $\alpha = 0$, the optimal design is the set of sites [2, 10, 12] and it achieves a performance target of 1.3, as noted above. However, as the desire for robustness to uncertainty increases, two things happen: the guaranteed achievable performance T decreases, and different reserve designs become favored. For example, at $\hat{\alpha} = 0.5$, design [1, 11, 12] achieves the highest assured performance (of 1.05). With a low requirement for robustness, reserve design is primarily based on high nominal probabilities, whereas when high robustness (high $\hat{\alpha}$) is required, the solutions consist of sites having low standard errors (high certainty) for the nominal probabilities. Note that if the uncertainty weights, w_{sc}, represent standard errors in the estimates of p_{sc}, $\hat{\alpha} = 0.5$ is roughly equivalent to the uncertainty represented by a 40% confidence interval.

For various desired target levels, a number of the reserve designs are tied for maximum robustness, but additional information could be used to resolve those ties. For example, at

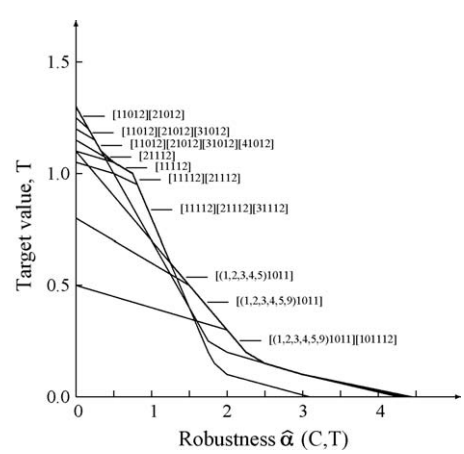

Fig. 2 – Robustness for different solutions for the simple reserve planning example (Table 2). The lines correspond to different solutions. At different robustness requirements (α) different solutions are robust-optimal. Alternative equal solutions are in brackets, and equally exchangeable sites in solutions are listed within parentheses. With low values for α high nominal probabilities are characteristic for the selected sites. With high α sites with moderate probabilities but low uncertainty (small standard errors) dominate the solution.

target level 1.0, highest robustness is achieved by two designs, [1, 11, 12] and [2, 11, 12], which each have $\hat{\alpha} = 0.75$. Yet further inspection reveals that these two reserve designs are equal only with respect to robustness of the most vulnerable species: the designs have robustness vectors (1, 4, 0.75) and (0.75, 0.8, 0.75) for species X, Y and Z, respectively. If we use the robustness of the next most vulnerable species to resolve the tie, this indicates that [1, 11, 12] is a better design than [2, 11, 12].

3.2. Opportuneness

Designs can be compared in terms of their nominal performance, as well as the robustness and opportuneness offered by them. Of the designs compared in Fig. 3, [1, 10, 12] is a clear overall winner: First, it has highest nominal performance, 1.3 (at $\alpha = 0$). Then in terms of robustness, [1, 10, 12] is preferred with high targets and the design is never much less robust than the alternatives. Furthermore, [1, 10, 12] has highest opportuneness across all targets (lowest uncertainty required to allow the possibility of meeting the aspiration).

The choice between [6, 10, 11] and [4, 11, 12] would not be so clear (Fig. 3). At high targets design [4, 11, 12] is more robust than [6, 10, 11] whereas with low targets the situation is reversed. In terms of opportuneness [6, 10, 11] is overall superior to [4, 11, 12]. Thus the choice between [4, 11, 12] and [6, 10, 11] would depend on planning priorities. Design [4, 11, 12] is preferred when assuming robust planning and low to moderate errors in the models. Design [6, 10, 11] is preferred if robustness to high error is required, or if the planning objective is to gamble for the best possible outcome (i.e., to take highest opportuneness).

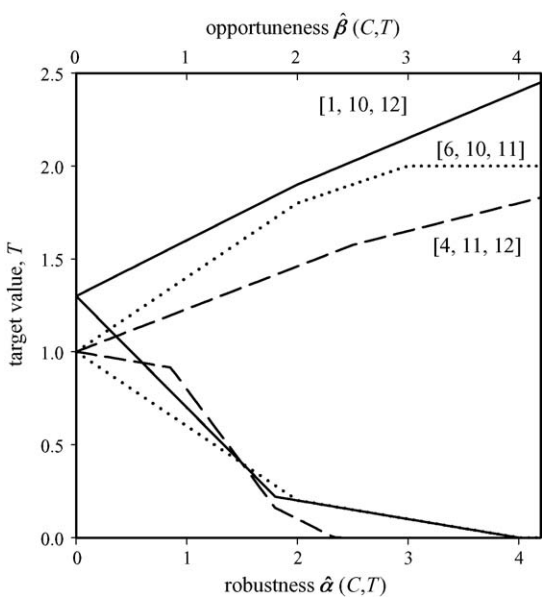

Fig. 3 – Comparison of three reserve designs in terms of robustness and opportuneness. Results for designs [1, 10, 12], [6, 10, 11] and [4, 11, 12] are marked by solid, dotted and dashed lines, respectively. For each design, the lower half of the line shows robustness ($\hat{\alpha}$, Eq. (7)) of the design at the respective target value. The upper half of the line shows opportuneness ($\hat{\beta}$, Eq. (11)), the minimum uncertainty required to allow the possibility of meeting the aspiration.

In general, both robustness and opportuneness functions for different designs can either cross or not. If they do not cross, then one design is always preferred over the other, with the distance between the functions indicating the strength of preference. If the functions cross, then there is a reversal of preference at the crossing point.

3.3. Probability of one occurrence

An alternative formulation of the performance objective (Eq. (1b)) asks for a set of sites that contains at least one population for each species with a very high probability. With the case of $T_s = 0.9999$, this target could be achieved, e.g., by two sites with probabilities 0.99 for each species or four sites with probabilities of 0.9. It is worth noticing that this formulation is extraordinarily sensitive to uncertainty. Because probabilities accumulate in a multiplicative manner in Eq. (1b), the achievable target will be very sensitive to errors in cell-specific probabilities.

Going back to the example of Table 2 and looking at species X, it appears that choosing sites [1, 2, 3] would give a probability of $1-(1-0.9)(1-0.95)(1-0.85) = 0.99925$ of having at least one occurrence of the species. For conservation decision making this probability appears close enough to absolute certainty. The situation changes if uncertainty is acknowledged. At α-level 1.0 (representing a maximum uncertainty of one standard error), we robustly achieve a probability of $1-(1-0.8)(1-0.75)(1-0.75) = 0.9875$. At $\alpha = 2.0$ we achieve a probability of 0.9425, which may no longer be satisfactory. As in Fig. 2, higher uncertainty implies that a lower target can be achieved reliably.

Furthermore, Eq. (1b) implicitly relies on the assumption that population dynamics (and probabilities of occurrences) are independent between selection units. With selection units of the scale relevant for land-use planning (~hectares), such independence cannot generally be expected for sites that are close to each other. Ovaskainen and Hanski (2003) find that if the degree of correlation between the occupancy states of two patches is ρ, then the effective number of independent populations in a system of N patches is $N_e = N/[(N-1)\rho + 1]$ instead of N. In our example, if sites 1–3 are close to each other and have fully correlated dynamics, then at $\alpha = 1.0$, we only achieve a probability of 0.8 of having an occurrence for the species. This simple analysis suggests that other sites farther away with lower correlation would be needed to achieve a higher target probability of occurrence. Thus, considering uncertainty may change our perception of the safety of our conservation plan, from the perspective of maintaining populations of species.

3.4. Robust reserve planning using a benefit function

The final reserve planning formulation we consider differs from the two previous ones in that no strict target (or aspiration) levels for individual species representations are given; rather, performance is measured by a continuous benefit function that is a weighted sum over species (Eq. (1c)). The optimal reserve structure based on nominal distributions is the one that gives the highest value of the summed continuous ben-

efit functions; let us assume this optimal value is V^*. (Finding V^* is an optimization task in itself.)

The optimal value, V^*, is obtainable with zero robustness ($\hat{\alpha} = 0$), that is, with no error in the nominal distributions $\tilde{p}$. However, we may be willing to forgo some value in favor of robustness to uncertainty in the design. That is, while the maximum performance is V^*, there may be some lower level, say $(1-\gamma)V^*$, that is satisfactory. Our task, then, is to find the reserve design that maximizes robustness while achieving this satisfactory level of performance. We use info-gap methods to do this, replacing Eq. (7) with

$$\hat{\alpha}(C, T) = \max\left\{\alpha : \left(\min_{s,m}\left[\min_{p \in U(\alpha, \tilde{p}_{sm})} V(C) - (1-\gamma)V^*\right]\right) \geq 0\right\}, \tag{12}$$

where $\gamma \in (0, 1)$ is the parameter specifying how much value we are willing to give up in return for robustness.

4. Discussion

The presence or persistence of biological value in reserve sites is uncertain for many reasons, including observation error, inaccuracies of distribution modeling, spatial population dynamics, the extinction debt, anthropogenic threat, succession, land-use changes and climate change. In general, most of these factors are ignored in computational reserve design, which may be based directly on presence–absence data or on model-predicted probabilities of occurrence. It follows that even a simple uncertainty analysis might reveal important weaknesses in reserve candidates. Even so, uncertainty analysis is all but missing from reserve selection methodology. The aim of the present work is to lay out a methodological basis that can be used as a framework for implementing uncertainty analyses in reserve planning. We point out that traditional reserve selection objective functions need to be modified to include effects of uncertainty. Actual search for the robust-optimal reserve structure is still a site selection optimization problem, which can be solved using typical optimization and approximation techniques, including stepwise heuristics, integer-programming and stochastic global search.

Uncertainty analysis has the potential to change conservation decisions. Do we think the conservation outcome is good enough based on best estimates? Are we sure the conservation outcome is good enough? These are different questions. Of two options with equal or near equal nominal value, the one which is more certain is better: this is old news in economics, but the message applies equally to conservation. Essentially, our uncertainty analysis is about understanding and quantifying tradeoffs between apparent biological value and the certainty of that information (see Table 1). To reiterate, best places are sure to have high value, robustness needs attention when both value and uncertainty are high, and opportuneness exists when apparent value is low but uncertain.

The present work concentrates on a couple of relatively simple reserve selection formulations, and we focus our uncertainty analysis on errors in predicted probabilities of occurrence from habitat models. However, we emphasize that the utility of the info-gap theory is not limited to these cases.

There are many other quantities that may be used to improve reserve planning algorithms and that are uncertain. For instance, Moilanen et al. (2005) outlined a method to aggregate reserves that avoids the use of an arbitrary boundary length penalty. Instead, it produces aggregated reserves based on distributions that have been smoothed according to species-specific dispersal capabilities. However, dispersal distance distributions are notoriously difficult to estimate correctly (e.g., Thomas et al., 2002). Dispersal is often handled via normalized dispersal kernels (e.g., Kot et al., 1996), which specify the distribution of dispersal distances for individuals leaving a focal site. One common dispersal kernel is the negative exponential, $\exp(-\alpha_s d_{ij})$, where α_s is a species-specific parameter describing the width of the dispersal kernel. There are two ways to info-gap this. In the simplest case, one could say that our estimate of α_s uncertain, and the question would be, is our conservation decision robust to this uncertainty. Alternatively, one could specify that both the parameter and the function itself are uncertain. So, the negative-exponential with a given parameter would be the nominal estimate of the dispersal kernel, but we could specify that the shape of the function is uncertain. Published info-gap formulations include cases where the uncertainty model is, for example, a family of monotonously decreasing normalized functions, which have a bounded deviation from a given nominal estimate (see Ben-Haim, 2006). Such a formulation would be appropriate for analyzing the effects of uncertainty in the (tails of the) dispersal kernel.

Distance-dependent correlation in population dynamics (or environmental noise) is another factor which may have great impact in PVA analyses or reserve planning (e.g., Moilanen and Cabeza, 2005). A decaying-by-distance correlation function would be a very similar, although even more uncertain, target for info-gap than a dispersal kernel. These functions may be troublesome to estimate because obtaining long time series of reliable large-scale data for model fitting is very hard. Furthermore, spatial population dynamics may be most severely affected by rare events (very good or very bad years), which are strong and correlated at large scales (e.g., Thomas et al., 2002). Data of such events is rarely available and thus the correlation scale and strength of environmental noise would be a very uncertain quantity that often gets ignored or underestimated in spatial population models or reserve planning.

Info-gap decision theory provides a way to guard against epistemic uncertainty, but it does not relieve the decision-maker from the challenge of setting clear objectives. The choice of the performance function is not trivial, nor is it the purview of the consulting scientist. Rather, it is an expression of the political and social values governing the reserve design. There are multiple elements of this expression of values: the choice of the performance function, $R_{sm}(C, p)$; the choice of the target values, T_{sm}; the choice of species (or biodiversity elements) to consider in the evaluation; and the method used to balance species against each other (a minimum function in Eq. (7), a fixed weighting in Eq. (12), and alternatives exist). Indeed, one of the central lessons of info-gap theory—that satisficing is more robust than optimization—forces consideration of the ultimate objectives of reserve design. In this endeavor, it is important to identify which elements are the jurisdiction of

scientific assessment, and which are the jurisdiction of political discourse.

In summary, we recommend info-gap uncertainty analysis as a standard practice in computational reserve planning. The need for robust reserve plans may change the way biological data are interpreted. It also may change the way reserve selection results are evaluated, interpreted and communicated. Information-gap decision theory provides a standardized methodological framework in which implementing reserve selection uncertainty analyses is relatively straightforward. We believe that alternative planning methods that consider robustness to model and data error should be preferred whenever models are based on uncertain data, which is probably the case with nearly all data sets used in reserve planning.

Acknowledgements

This work was conducted as a part of the Working Group on "Setting priorities and making decisions for conservation risk management", supported by the National Center for Ecological Analysis and Synthesis, a Center funded by NSF (Grant # DEB-94-21535), the University of California at Santa Barbara, and the State of California. A.M. acknowledges support from the Academy of Finland project #1206883 and The Finnish Center of Excellence Programme 2000-2005, grant #44887.

REFERENCES

Araujo, M.B., Williams, P.H., 2000. Selecting areas for species persistence using occurrence data. Biol. Cons. 96, 331–345.

Arponen, A., Heikkinen, R., Thomas, C.D., Moilanen, A., 2005. The value of biodiversity in reserve selection: representation, species weighting and benefit functions. Cons. Biol. 19, 2009–2014.

Ben-Haim, Y., 2006. Information-gap Decision Theory: Decisions Under Severe Uncertainty, 2nd ed. Academic Press, London, UK.

Burgman, M., Lindenmayer, D.B., Elith, J., 2006. Managing landscapes for conservation under uncertainty. Ecology 86, 2007–2017.

Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multi-model inference: A Practical Information-Theoretic Approach. Springer-Verlag, New York.

Cabeza, M., 2003. Habitat loss and connectivity of reserve networks in probability approaches to reserve design. Ecol. Lett. 6, 665–672.

Cabeza, M., Moilanen, A., 2001. Design of reserve networks and the persistence of biodiversity. TREE 16, 242–248.

Cabeza, M., Moilanen, A., 2003. Site-selection algorithms and habitat loss. Cons. Biol. 17, 1402–1413.

Cabeza, M., Moilanen, A., Possingham, H.P., 2004. Metapopulaton dynamics and reserve network design. In: Hanski, I., Gaggiotti, O. (Eds.), Metapopulation Ecology, Genetics, Evolution. Academic press, pp. 541–564.

Chankong, V., Haimes, Y.Y., 1968. Multiobjective Decision Making: Theory and Methodology. North-Holland, New York.

Cowling, R.M., Pressey, R.L., Sims-Castley, R., le Roux, A., et al., 2003. The expert or the algorithm?—comparison of priority conservation areas in the Cape Floristic Region identified by park managers and reserve selection software. Biol. Cons. 112, 147–167.

Csuti, B., Polasky, S., Williams, P.H., Pressey, R.L., et al., 1997. A comparison of reserve selection algorithms using data on terrestrial vertebrates in Oregon. Biol. Cons. 80, 83–97.

Drechsler, M., Burgman, M.A., 2004. Combining population viability analysis with decision analysis. Biodiv. Conserv. 13, 141–164.

Elith, J., Burgman, M.A., Regan, H.M., 2002. Mapping epistemic uncertainties and vague concepts in predictions of species distribution. Ecol. Model. 157, 313–329.

Gaston, K.J., Rodrigues, A.S.L., 2003. Reserve selection in regions with poor biological data. Cons. Biol. 17, 188–195.

Haight, R.G., Revelle, C.S., Snyder, S.A., 2000. An integer optimization approach to a probabilistic reserve site selection problem. Oper. Res. 48, 697–708.

Kot, M., Lewis, M.A., van den Driessche, P., 1996. Dispersal data and the spread of invading organisms. Ecology 77, 2027–2042.

MacKenzie, D.I., Nichols, J.D., Hines, J.E., et al., 2003. Estimating site occupancy, colonization and local extinction probabilities when a species is not detected with certainty. Ecology 84, 2200–2207.

Margules, C.R., Pressey, R.L., 2000. Systematic conservation planning. Nature 405, 243–253.

Moilanen, A., Cabeza, M., 2005. Variance and uncertainty in reserve selection. Cons. Biol. 19, 1663–1667.

Moilanen, A., 2005a. Reserve selection using nonlinear species distribution models. Am. Nat. 165, 695–706.

Moilanen, A., 2005b. Methods for reserve selection: interior point search. Biol. Cons. 124, 485–492.

Moilanen, A., Franco, A.M.A., Regan, I.E., et al., 2005. Prioritising multiple-use landscapes for conservation: methods for large multi-species planning problems. Proc. R. Soc. London B 272, 1885–1891.

Moore, J.L., Folkman, M., Balmford, A., et al., 2003. Heuristic and optimal solutions for set-covering problems in conservation biology. Ecography 26, 595–601.

Noss, R., 2003. A checklist for wildlands network designs. Cons. Biol. 17, 1270.

Ovaskainen, O., Hanski, I., 2003. Extinction threshold in metapopulation models. Ann. Zool. Fenn. 40, 81–97.

Polasky, S., Camm, J.D., Solow, A.R., et al., 2000. Choosing reserve networks with incomplete species information. Biol. Cons. 94, 1–10.

Possingham, H.P., Lindenmayer, D.B., McCarthy, M.A., 2001a. Population viability analysis. Encyclopedia of Biodiversity 4, 831–843.

Possingham, H.P., Andelman, S.J., Noon, B.R., Pulliam, H.R., 2001b. Making smart conservation decisions. In: Orians, G.E., Soulé, M.E. (Eds.), Research Priorities for Conservation Biology. Island Press, Washington, DC, pp. 225–244.

Pressey, R.L., 1999. Systematic conservation planning for the real world. Parks 9, 1–6.

Pressey, R.L., Possingham, H.P., Day, J.R., 1997. Effectiveness of alternative heuristic algorithms for identifying indicative minimum requirements for conservation reserves. Biol. Cons. 80, 207–219.

Regan, H., Ben-Haim, Y., Langford, B., Wilson, W.G., Lundberg, P., Andelman, S.J., 2005. Robust decision making under severe uncertainty for conservation management. Ecol. Appl. 15, 1471–1477.

ReVelle, C.S., Williams, J.C., Boland, J.J., 2002. Counterpart models in facility location science and reserve selection science. Env. Model. Assess. 7, 71–80.

Rodrigues, A.S.L., Andelman, S.J., Bakarr, M.I., et al., 2004. Effectiveness of the global protected area network in representing species diversity. Nature 428, 640–643.

Thomas, C.D., Wilson, R.J., Lewis, O.T., 2002. Short-term studies under-estimate 30-generation changes in a butterfly metapopulation. Proc. R. Soc. London B 268, 1791–1796.

Van Niel, K.P., Laffan, S.W., Lees, B.G., 2004. Error and uncertainty in environmental variables for predictive vegetation modelling. J. Veg. Sci. 15, 747–756.

Williams, J., ReVelle, C.S., Levin, S.A., 2004. Using mathematical optimization models to design nature reserves. Front. Ecol. Environ. 2, 98–105.

Wilson, K.A., Westphal, M.I., Possingham, H.P., Elith, J., 2005. Sensitivity of conservation planning to uncertainty associated with predicted species distribution data. Biol. Cons. 122, 99–112.