ECOLOGY LETTERS

Ecology Letters, (2010) **13**: 1325–1337 doi: 10.1111/j.1461-0248.2010.01514.x

IDEA AND PERSPECTIVE

Allocating monitoring effort in the face of unknown unknowns

Brendan A. Wintle,[1]* Michael C. Runge[2] and Sarah A. Bekessy[3]

[1]*School of Botany, University of Melbourne, Melbourne, Vic. 3010, Australia*

[2]*US Geological Survey, Patuxent Wildlife Research Center, Laurel, MD 20708, USA*

[3]*RMIT, School of Global Studies, Social Science and Planning, Melbourne, Vic. 3001, Australia*

Correspondence: E-mail: brendanw@unimelb.edu.au

Abstract

There is a growing view that to make efficient use of resources, ecological monitoring should be hypothesis-driven and targeted to address specific management questions. 'Targeted' monitoring has been contrasted with other approaches in which a range of quantities are monitored in case they exhibit an alarming trend or provide *ad hoc* ecological insights. The second form of monitoring, described as surveillance, has been criticized because it does not usually aim to discern between competing hypotheses, and its benefits are harder to identify *a priori*. The alternative view is that the existence of surveillance data may enable rapid corroboration of emerging hypotheses or help to detect important 'unknown unknowns' that, if undetected, could lead to catastrophic outcomes or missed opportunities. We derive a model to evaluate and compare the efficiency of investments in surveillance and targeted monitoring. We find that a decision to invest in surveillance monitoring may be defensible if: (1) the surveillance design is more likely to discover or corroborate previously unknown phenomena than a targeted design and (2) the expected benefits (or avoided costs) arising from discovery are substantially higher than those arising from a well-planned targeted design. Our examination highlights the importance of being explicit about the objectives, costs and expected benefits of monitoring in a decision analytic framework.

Keywords

Adaptive management, cost efficiency, decision theory, Knightian uncertainty, surprise, surveillance.

Ecology Letters (2010) 13: 1325–1337

INTRODUCTION

Monitoring is central to the study of ecology and the management of natural resources because it provides the primary mechanism by which we discover things that were previously not known, discern among competing hypotheses, gauge the state of biodiversity resources and learn about the effectiveness of conservation investments. Monitoring provides the critical feedback loop in adaptive management, an important paradigm for conservation management in the face of uncertainty (Walters 1986). The past decade has seen significant advances in optimal monitoring (Johnson *et al.* 1997; Shea & The NCEAS Working Group on Population Management 1998; Yoccoz *et al.* 2001; Hauser *et al.* 2006; Nichols & Williams 2006; McCarthy & Possingham 2007; Chades *et al.* 2008; McDonald-Madden *et al.* 2010). Monitoring that is designed to be optimal with respect to a clearly stated objective and that aims to discern among clearly stated *a priori* hypotheses has been described as *targeted* (or focused) monitoring (Nichols & Williams 2006). The U.S. Fish and Wildlife Service's Adaptive Harvest Management Program for waterfowl (USFWS 2009) is an emblematic example of targeted monitoring because it utilizes information gained from monitoring to discern among competing hypotheses about the effectiveness of waterfowl management options. The attributes monitored in the program are specifically chosen to optimize learning about the relative credibility of competing hypotheses. Targeted monitoring has been contrasted with *surveillance* monitoring programs, that tend to lack clearly stated *a priori* hypotheses, and in which a range of quantities are monitored in case they exhibit an alarming trend or provide *ad hoc* ecological insights (Nichols & Williams 2006). Examples include the North American Breeding Bird survey, the Pan-European Common Bird Monitoring Scheme and the British Breeding Bird Survey (Greenwood 2002; Sauer *et al.* 2005; PECBMS

2009). Targeted monitoring stands at one end of a continuum, with classical surveillance monitoring at the other, and mixed strategies such as the US LTER program (Hobbie *et al.* 2003) in between.

One of the benefits of targeted or management-focused monitoring (*sensu* Nichols & Williams 2006) is that managers (and accountants) can see the value of investing in monitoring when it resolves competing hypotheses of cause-and-effect and improves management decisions and efficiency (e.g. Hauser *et al.* 2006; Nichols *et al.* 2007; McDonald-Madden *et al.* 2010). Monitoring programs at the surveillance end of the continuum, such as the breeding bird surveys (BBS), have been criticized because they do not have an obvious management focus, their benefits are harder to identify *a priori*, they generally suffer from sampling bias and low statistical power, and there are no actions triggered by any particular observed trend (Nichols & Williams 2006). Using arguments from both the philosophy of science (Chamberlin 1897; Platt 1964) and a cost-efficiency framework, Nichols and Williams make a compelling case that investment in such programs is an inefficient use of scarce resources. Should we, therefore, do away with monitoring programs that are not focused on a particular (set of) management action(s) within a decision theoretic framework? Or is there a coherent cost-efficiency logic that supports the case for investment in monitoring that is not explicitly focused on specific management or scientific goals?

Nichols and Williams invoke the work of John Platt (1964) to argue that monitoring programs which are not explicitly designed to discern efficiently between *a priori* hypotheses about the state or functioning of a system provide 'weak inference' (*sensu* Platt 1964). Platt defers to the work of Chamberlin (1897) when making the case that knowledge most rapidly advances if science focuses on the collection of information that efficiently discriminates between multiple competing hypotheses. Nichols and Williams (2006) interpret this, in the modern context of urgent and under-resourced conservation management, as meaning that 'management targeted' monitoring, aimed at discerning among hypotheses about management effectiveness, should be prioritized for limited conservation funding over non-targeted-surveillance programs, and that many existing monitoring programs that lack *a priori* hypotheses are comparatively poor investments. But how are these hypotheses generated, and is there a role for surveillance monitoring therein?

Everyday reasoning depends predominantly on patterns of repeated experience rather than deductively valid arguments: 'We believe that bread will nourish us today because it has done so in the past, but this is not a guarantee that it will always do so: Someone who insists on sound deductive justifications for everything would starve to death' (Hume 1748). In ecology, inductive reasoning via experience and observation is a predominant means of hypothesis generation. Conversely, it may be argued that after more than a hundred years of ecological study we already have many untested hypotheses and do not need to go looking for more, or that targeted monitoring is just as likely to lead to *ad hoc* insights and new hypotheses as surveillance. Should we, therefore, pause, and invest our efforts in testing existing ideas with targeted monitoring and experiments, or should we continue to invest in long-term, surveillance observation programs in the hope that they may generate new hypotheses that change the way we think, or save us from unforeseen peril?

Humans have long recognized that we live in a state of partial knowledge, and that it might be advantageous to be open to, perhaps even seek, surprise. Thoreau (1854) quotes Confucius (*Analects*, book 2, verse 17); 'To know that we know what we know, and that we do not know what we do not know, that is true knowledge.' In the past century, typologies and calculus' of uncertainty have been developed (Halpern 2003). Knight's (1921) three part typology has been enormously influential to modern decision theory: *certainty* governs the case when the deterministic outcomes of all alternative actions are known; *risk* forms the basis of normative decision theory, and describes a decision context in which the outcome is stochastic, but the contingencies and their probabilities are known; and *uncertainty* concerns the case of severe uncertainty under which the probabilities (or probability distributions) of a set of known contingencies are unknown (e.g. Ben-Haim 2006). Genuine *surprise* (Hilborn 1987) goes beyond the conventional interpretation of Knightian *uncertainty* to encompass the case in which both the contingencies and (by definition their probabilities) are unknown. Such uncertainties are called 'unknown unknowns' or 'black swans' in popular parlance (Furlong 1984; Rumsfeld 2002; Taleb 2007) and are considered by some to be the most important uncertainties in our lives (Taleb 2007). Particularly in the face of global climate change, there is considerable concern about how to be open to surprise and seek discovery of unforeseen phenomena (e.g. Schneider *et al.* 1998).

Thus, an argument in favour of surveillance monitoring programs is that they *may*, due to their tendency to have a broader geographic, temporal and biological scope, be better designed to discover Rumsfeldian 'unknown unknowns' than the purely management-targeted monitoring activities. Nichols & Williams (2006; p. 671) pre-empt this defense of surveillance monitoring activities, and argue that 'no monitoring program, whether targeted or surveillance, can be *assured* of consistently registering unanticipated events'. Given the 'large number of extant conservation issues, and the finite resources available to address them', they argue 'against designing monitoring programs solely to recognize unanticipated problems, even if it were clear how to do so.'

That no monitoring program can be assured of consistently registering unanticipated events is not central to a debate about the relative merits of the two monitoring paradigms. The more salient question is: (1) how much greater would the probability of detecting unforeseen (and possibly costly) events need to be under a surveillance monitoring strategy to consider investing resources in such an approach and (2) how much more benefit (or avoided disbenefit) arises from the discovery of unknown unknowns compared with what can gained from learning more about solutions to problems that have already been identified? It is reasonable to question whether a portion of available resources should be spent on surveillance, especially if it were not particularly costly and its collection substantially increased the probability of new ecological insights, key discoveries, averting disasters or providing windfalls. The question is, for a given set of circumstances and assumptions, what proportion of a limited budget should be allocated to the two approaches? How should the trade-off between learning about *known* unknowns and *unknown* unknowns be handled?

In this article, we first explore the case in favour of surveillance monitoring approaches by describing some real and hypothetical examples of where investment in surveillance monitoring has brought windfalls or helped to avert disasters. We then ask whether there could be a monitoring design that is more amenable to fostering discovery of unforeseen phenomena, and exploring what such a design might look like. Because neither of these explorations can provide proof of the worth of surveillance monitoring, we then set up a general analytical model for the value of ecological monitoring to more rigorously explore the circumstances under which investment in surveillance monitoring may be defensible. We conclude with a discussion about the implications of our findings for existing and future monitoring programs.

THE CASE FOR SURVEILLANCE MONITORING

There are several steps to coping with emergent, unpredicted patterns: discovery, hypothesis generation, corroboration and response. There have been several instances in which surveillance monitoring has either led to the discovery of ecological change or has been extremely useful in corroborating concerns arising from *ad hoc* observations about a particular phenomenon, leading to appropriate responses. Hawkins *et al.* (2006) used annual statewide spotlighting surveys of mammal distributions and abundance in Tasmania, Australia to infer a relationship between Devil Facial Tumour Disease (DFTD) emergence and a decline in the abundance of the Tasmanian Devil (*Sarcophilus harrisii*). With this evidence DFTD researchers were able to rapidly instigate action to address the problem without having to wait for the results of a separate targeted study to confirm the cause and magnitude of the population decline (Clare E. Hawkins, personal communication). Given the speed with which DFTD has infected the population, and the difficulties associated with finding control (uninfected) populations, it may have been extremely difficult to establish a study into the population effects of this disease. Without the existing monitoring data, which was certainly not collected with disease mapping in mind (let alone DFTD), there would have been a serious delay in measuring the population decline, potentially delaying remedial action. The statewide roadside spotlight survey data were collected since 1975 for the purposes of monitoring to 'keep an eye on' threatened species, assisting with harvest regulation, and gauging the performance of threat management strategies (Driessen & Hocking 1992, 2008), making it an example of an omnibus design at the surveillance end of the targeted-surveillance continuum.

Surveillance monitoring can also play a role in fundamental ecological research. In 1979, 20 000 trees (> 10 cm d.b.h.) were identified, mapped and measured in a 50-ha plot on Barro Colorado Island (BCI) in Panama. This study has evolved into a long-term, large-scale forest monitoring site, which has led to the development of 34 other such sites across 20 countries. The ongoing inventory of species has become extremely important for generating and testing emerging ecological theories and for understanding the ecological implications of global change. In 2001, Stephen Hubbell published the Unified Neutral Theory of Biodiversity and Biogeography (Hubbell 2001). Hubbell utilized the BCI data as an example of where the 'zero sum model', central to the neutral theory, works well. The corroboration of the zero sum model was not part of the original design requirements for the BCI data set and could not have been foreseen in 1979 when the program was established to gain a better understanding of tropical forest dynamics (Hubbell & Foster 1986).

Similar stories about the role of surveillance monitoring in discovery or corroboration of hypotheses exist for a range of ecological phenomena and conservation issues. For example, white-nose syndrome, an emergent disease that is devastating bat populations in the eastern United States, was discovered through anecdotal accounts from property owners (Blehert *et al.* 2009). Its effects were corroborated by inspection of an ongoing surveillance-style monitoring program, which is undertaken to track the status of Federal- and State-listed threatened bat species and which includes ancillary monitoring of non-listed species (Jeremy T.H. Coleman, USFWS, personal communication). Without baseline and longitudinal data, the scope, magnitude and rapidity of the decline in bat populations would not have been evident, and swift responses to manage the disease might not have been undertaken

(Szymanski *et al.* 2009). This example highlights the point that, while surveillance monitoring may not necessarily lead directly to the discovery of a new ecological phenomenon, it may serve the purpose of rapidly corroborating a new idea or anecdotal report of ecological change.

Sauer *et al.* (2000; 2005a) utilized long-term BBS data to confirm precipitous declines in grassland-dependent bird populations throughout the United States. While the problem had been anecdotally identified, affirming these declines stimulated and motivated ideas about how to arrest declines before the risk of species loss became unacceptably high. As the declining distributions of grassland birds have been confirmed, a number of important steps have been taken to buffer the impacts of agricultural intensification and other factors linked to decline (e.g. Trocki & Paton 2005). In this case, the existence of the BBS data did not bring about the discovery of the decline. However, had the geographically extensive, long-term data set not been available to confirm the existing reports of decline, it is questionable whether there would have been sufficient scientific effort and political will to invest in mitigation measures (Wildlife Habitat Management Institute 1999) and a detailed study of their effectiveness (Vickery & Herkert 2001).

These four examples highlight the potential role of surveillance monitoring in generating and corroborating new ideas, and discovering unforeseen patterns. Similar data sets have been used in other disciplines to corroborate new ideas or generate hypotheses. The International Panel on Climate Change (IPCC) Third Assessment Report (IPCC 2001) presents a 140-year data record of thermometer readings to provide compelling evidence that a steady warming has occurred since the industrial revolution that does not accord with a natural fluctuation hypothesis. Temperatures and various other climatic variables have been routinely measured by governments and private individuals for hundreds of years, in some cases with a specific purpose in mind, but in other cases, just because it seems like a useful thing to track. Scientists, governments and 'private enthusiasts' appreciate the value of 'keeping a finger on the pulse' of various aspects of the environment, because, from time to time, this activity has proven to be extremely useful. Sociology, demography, biosecurity, public health and defense are disciplines in which the routine collection of surveillance data is commonplace and often leads to interesting and important findings (e.g. Buckles & Hungerman 2008).

What are the characteristics of these data sets that led to important discoveries and hypotheses, and could these benefits have arisen with existing targeted-monitoring designs? There are some general attributes of surveillance and management-targeted monitoring programs that may help address these questions (Table 1). Targeted and surveillance monitoring designs tend to differ in their

Table 1 General characteristics of targeted and surveillance monitoring designs. These two approaches represent the extremes of a continuum which may include intermediate or mixed strategies

Targeted monitoring	Surveillance monitoring
Targeted to improve management by learning about pre-specified processes or knowing the system state. May have a specific scientific purpose (i.e. optimized to discern between competing hypotheses)	Generally not based on a particular management problem or scientific question
Sampling optimal to address specific hypotheses or to estimate state. High statistical power to differentiate between specific hypotheses or to achieve precise estimates of state variables	Sampling not optimized to a particular purpose, though trend detection is often cited as a rationale. Generally low power to differentiate between hypotheses or to estimate trends
Deductive reasoning	Inductive reasoning
A priori hypotheses articulated	Poorly specified, vague or non-existent *a priori* hypotheses
Typically has a narrower scope; fewer species, few state variables, fewer covariates	Often broad in geographic scope; many species, many state variables, many covariates
Generally well stratified, replicated and exhibiting low bias for the pre-specified purpose	Often poorly stratified, not replicated, and having a biased sampling frame
Variable sample sizes, depending on the problem at hand	Often very large sample sizes
Generally collected by professional scientists (expensive per data point)	Generally collected by a mix of professional scientists and volunteer (casual) observers (cheap per data point)
Less amenable as a tool for community engagement	Community engagement commonly cited as one of the primary objectives

temporal and geographic scope and the number of variables measured. Targeted-monitoring designs tend to be optimized to address a particular question about system processes or management, meaning that only the most relevant proximal variables tend to be measured [e.g. the abundance of a particular duck species and covariates selected to discern among competing *a priori* hypotheses about causes of fluctuations in abundance (Johnson *et al.* 1997)]. In contrast, surveillance monitoring programs often measure a broader array of covariates reflecting underlying or nascent hypotheses.

It is possible that discovery or corroboration of unforeseen patterns or processes could arise as a by-product of targeted monitoring. In 1970, Likens *et al.* (1970) published a classic paper describing the results of a targeted study into the effects of forest cutting and herbicide treatment on nutrient budgets in the Hubbard Brook watershed. A short time later, Likens *et al.* (1972) and Likens & Bormann (1974) were able to use the same case study to draw a causal link between SO_2 emissions and the concentration of SO_4^{2-} in precipitation within the Hubbard Brook watershed; thereby having a profound, if serendipitous impact on public policy of the time (Gene E. Likens, personal communication). If the frequency with which targeted-monitoring designs lead to unforeseen discoveries is as high as the frequency with which surveillance studies lead to such discoveries, then it would make sense to achieve the benefits of a higher powered study to address a pressing conservation problem and enjoy the same chance of discovering the unknown unknowns. So, is the probability of an unforeseen discovery arising from a management-targeted study as high as the probability of serendipitous discovery arising from surveillance?

The number of variables monitored, the breadth of the environmental and geographic space sampled, the length of time over which the monitoring takes place and the number of samples taken in each time period will determine the probability of detecting unforeseeable events or phenomena. However, the probability that a particular monitoring design will discover or corroborate something novel of great importance is difficult to determine. Standard approaches in estimating the statistical power of a monitoring design cannot be applied when the hypothesis in question does not yet exist. Consequently, proponents of surveillance monitoring are in the difficult position of being unable to demonstrate the value of their monitoring programs in a traditional statistical way. Similarly, critics currently have no way to prove that the heuristic appeal of surveillance monitoring for 'keeping a finger on the pulse' or 'stumbling' on an important answer is illogical or inefficient. Here, we attempt to bring clarity to arguments about the relative utility of surveillance and targeted monitoring by couching the argument in a decision theoretic framework. In the following section, we derive a model to calculate the expected benefit of any monitoring strategy on the continuum between targeted and surveillance monitoring. We then use the model to search for the parameter space that legitimizes surveillance approaches and ask whether such a space could exist in real life.

A FRAMEWORK FOR ALLOCATION OF MONITORING EFFORT

Model of net benefit

Suppose there is a total budget (B) that can be allocated between two monitoring programs: one that is focused specifically on some pre-defined management objectives, and one that is designed to be open to serendipitous discovery of unforeseen emergent patterns and the generation of hypotheses. The amount spent on the targeted monitoring is v, and the amount spent on surveillance monitoring is $B - v$. Each of these programs has a probability p_i of discovering an unforeseen pattern, given the pattern occurs, and probability q_i of achieving a specific management benefit. All these probabilities (the p_i's and q_i's) depend on the amount of effort (and capital) invested. If either program is successful in achieving the management benefit, the reward is R. If both programs fail to detect any unforeseen emergent pattern, the cost is C. In general terms, R is defined as the reward gained from improvements to management arising from better knowledge or understanding of a *known* unknown. C is the cost averted by discovering (and ameliorating) a novel threat (an *unknown* unknown). The decision makers may view these as separate fundamental objectives. Note that in this model, R and C need to be expressed in the same units that reflect the trade-offs among these objectives. These units might be monetary, but they need not be; they could be measured, say, in tons of fish harvested, or the number of endemic species preserved, or the average score that emerged from a multi-criteria decision analysis. Finally, we have an *a priori* belief that an unforeseen pattern will emerge with probability f, the background frequency with which unforeseen disasters occur.

Formally, we can express the expected net benefit, conditional on the allocation of monitoring effort, as

$$\begin{aligned} V(v) = R[1-(1-q_1(v))(1-q_2(B-v))] \\ -fC[(1-p_1(v))(1-p_2(B-v))]. \end{aligned} \tag{1}$$

The first term captures the expected management benefit, where the term in square brackets represents the probability that at least one of the monitoring programs (targeted or surveillance) will provide the information to achieve the desired management benefit, R. The second term captures the expected loss from an undetected emergent problem.

Here, the term in square brackets represents the probability that both monitoring programs will fail to detect the emergent pattern. If either program detects the pattern, the disaster is averted and no cost is incurred. With this formulation, we can ask a number of questions, such as when is investment in surveillance monitoring warranted? How much of the budget should be allocated to each of the monitoring approaches?

The properties of the efficiency functions (that is, the equations that govern the probability of management benefit or emergent pattern detection as a function of monitoring investment) are important for understanding the optimal allocation of effort. First, as both sets of functions are probabilities, the $p_i(x)$ and $q_i(x)$ are constrained to [0,1] for any levels of investment, x. Second, without any investment, the probability of achieving management benefit (q_i) from a particular monitoring program is 0, and the probability of discovering an unforeseen pattern (p_i) is likewise 0. Third, we expect that increased investment in a monitoring program is never detrimental. Therefore, the performance functions are monotonically increasing in x (derivatives constrained to be ≥ 0 for all values of x). We might also expect diminishing marginal returns on investment, thus, the second derivatives of these two functions are likely negative, at least for larger values of x. Fourth, we assume that targeted monitoring is always better than surveillance monitoring at achieving management benefit, for equal investment. Thus,

$$q_1(x) \geq q_2(x) \tag{2}$$

for all x, where q_1 is the probability of achieving management benefit from the targeted-monitoring program, and q_2 is the same probability for the surveillance monitoring program. The justification is that the targeted-monitoring program is designed specifically for the purpose of maximizing the likelihood of management benefit, so it should be at least as good as any other monitoring program and, most of the time, strictly better.

The final assumption is the crux of the matter. Are surveillance monitoring programs better for serendipitous discovery than targeted-monitoring programs? That is, is it the case that

$$p_2(x) \geq p_1(x) \tag{3}$$

for any level of investment, x? This is the underlying justification for surveillance monitoring programs, but is it true? We return to this question in the Discussion, below.

Analysis

We can begin by looking at the endpoints of investment, that is, investing in only one or the other monitoring program. The expected net benefit for investing the entire budget in targeted monitoring is

$$E_1 = V(B) = Rq_1(B) - fC(1 - p_1(B)), \tag{4}$$

and the corresponding net benefit if the entire budget is invested in surveillance monitoring is

$$E_2 = V(0) = Rq_2(B) - fC(1 - p_2(B)). \tag{5}$$

When E_2 exceeds E_1, it is better to invest the entire budget in surveillance monitoring than in targeted monitoring. This occurs when

$$p_2(B) - p_1(B) > \frac{R}{fC}[q_1(B) - q_2(B)]. \tag{6}$$

This critical result says that in order to justify full investment in surveillance monitoring, the probability of serendipitous discovery under the surveillance monitoring program must exceed that under the targeted-monitoring program ($p_2 - p_1$), by an amount that is at least equal to the degree to which targeted monitoring is more likely to achieve management benefit than surveillance monitoring ($q_1 - q_2$), attenuated by the ratio of the expected benefits and costs (R/fC). The fourth assumption (eqn 2) assures us that the right-hand side of this criterion is non-negative. If the fifth assumption (eqn 3) is not warranted, then this criterion is not met, and full investment in surveillance

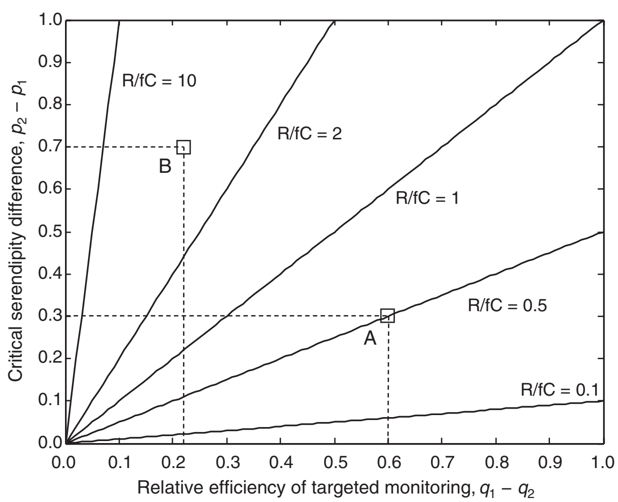

Figure 1 Critical values for full investment in surveillance monitoring vs. full investment in targeted monitoring, as a function of the relative efficiencies of the two programs to achieve management benefit ($q_1 - q_2$) or detect unknown unknowns ($p_2 - p_1$), and as a function of the expected benefit to cost ratio (R/fC). The contours shown are critical values for the relative efficiency of surveillance monitoring for serendipitous discovery, for various benefit-cost ratios, as derived from eqn 6. Values above a particular contour favour surveillance monitoring. Points A and B are examples explained in the text.

monitoring cannot be justified. If the fifth assumption is warranted, then there is at least the possibility that surveillance monitoring could be rationally justified.

Equation 6 helps us to understand the conditions that warrant full investment in surveillance monitoring (Fig. 1). For example, if the targeted-monitoring program is substantially more likely to achieve the management benefit than the surveillance monitoring program (i.e. $q_1 - q_2 = 0.6$), and the benefit to cost ratio is 0.5 (Fig. 1, point A), then the absolute difference in the probability of detecting an unknown unknown under surveillance and targeted monitoring ($p_2 - p_1$) needs to be at least 0.3 to warrant full investment in surveillance monitoring. Likewise, if the relative efficiency of surveillance and targeted monitoring for serendipitous discovery is $p_2 - p_1 = 0.7$, and the targeted-monitoring program more effectively brings management benefits ($q_1 - q_2 = 0.22$), then full investment in surveillance monitoring is favoured, if the expected benefit:cost ratio is < 3.2 : 1 (Fig. 1, point B). In the extremes, if the benefit:cost ratio is very small (if the expected loss from some unknown is huge compared to the management benefits), then surveillance monitoring is warranted, even when the targeted monitoring is much more efficient at achieving management benefit. On the other hand, if the benefit:cost ratio is high and $q_1 - q_2$ is even just moderate, surveillance monitoring cannot be warranted even if it is guaranteed to detect unknown unknowns.

These results so far just govern the full investment of the budget in one of the monitoring programs. Taking this a step further, under what conditions is a mixed portfolio of investment optimal? That is, when is it best to allocate part of the budget to targeted monitoring and part to surveillance monitoring? The maximum net benefit will be greater than either E_1 or E_2 and occur with targeted-monitoring investment $0 < v < B$ when a critical maximum occurs between 0 and B. A necessary, but not sufficient, condition is that the derivative of net benefit is 0, that is, when

$$\frac{\mathrm{d}V}{\mathrm{d}v} = 0, \tag{7}$$

which occurs when

$$(1-p_1(v))\frac{\mathrm{d}p_2(B-v)}{\mathrm{d}v} - (1-p_2(B-v))\frac{\mathrm{d}p_1(v)}{\mathrm{d}v} = \frac{R}{fC}\left[(1-q_2(B-v))\frac{\mathrm{d}q_1(v)}{\mathrm{d}v} - (1-q_1(v))\frac{\mathrm{d}q_2(B-v)}{\mathrm{d}v}\right]. \tag{8}$$

Equation 8 is not readily solved, owing to the potential complexities of the four efficiency functions, but several features of the solution are evident. First, the answer depends in part on the ratio of the expected benefits to the expected costs. Second, the differences in the *failure* rates (e.g. $1 - p_1$) play a role. And third, the slopes of the efficiency functions are important; that is, the solution depends on how quickly the efficiencies of the various monitoring programs improve with increasing investment.

Consider the following heuristic example. Suppose we are managing a threatened species of colonial insect. There are about 200 extant colonies and management is directed at reducing known threats and making modest additions to the number of colonies. Effective management requires monitoring to survey the status of the colonies, identify where known threats are occurring to focus intervention, and identify gaps in territories where new colonies can be placed. Let's say that the potential management benefit, R, is 10 added colonies, the potential catastrophic loss, C, from an unforeseen event is 100 lost colonies, and the frequency with which we expect the catastrophic loss to occur, f, is 0.01. Consider the efficiency curves given in Fig. 2. Targeted monitoring is better at achieving management benefit than surveillance monitoring (Fig. 2a), capping at a probability of 0.8 once the investment in it is somewhere around \$60 000. The probability of surveillance monitoring achieving management benefit caps at about 0.3. On the other hand, surveillance monitoring is better than targeted monitoring at detecting consequential unknown unknowns over all levels of investment (Fig. 2b).

With these parameters, and with a fixed monitoring budget of \$100 000, $E_1 > E_2$ (Fig. 3). That is, if you have to invest the entire budget in just one monitoring program, the targeted-monitoring program provides a greater net benefit (7.0 added colonies, on average) than the surveillance monitoring program (2.05 added colonies). But, the optimal allocation of monitoring effort is to invest about 57% of the budget in targeted monitoring and the rest in surveillance monitoring.

This optimal allocation is itself a function of the total budget (Fig. 4). At very low budget levels, the best strategy is to invest only in targeted monitoring. As the budget increases, it becomes optimal to start investing in surveillance monitoring as well. This transition occurs when the budget is about \$70 000. Note that at this point, the slopes of the targeted-monitoring curves (q_1 and p_1) have become quite small, and the initial slopes of the surveillance monitoring curves (q_2 and p_2) now provide benefit to investing in that strategy. The pattern of investing in surveillance monitoring only when the monitoring budget is generous, while not a general result, is similar to the argument made by Nichols & Williams (2006) and others – that when budgets are quite limited, it is better to invest in targeted monitoring. When there are strongly diminishing marginal returns on q_1 (the probability that a management benefit is obtained for a given investment in targeted monitoring), mixed investment strategies, involving some investment in surveillance monitoring, are likely to arise,

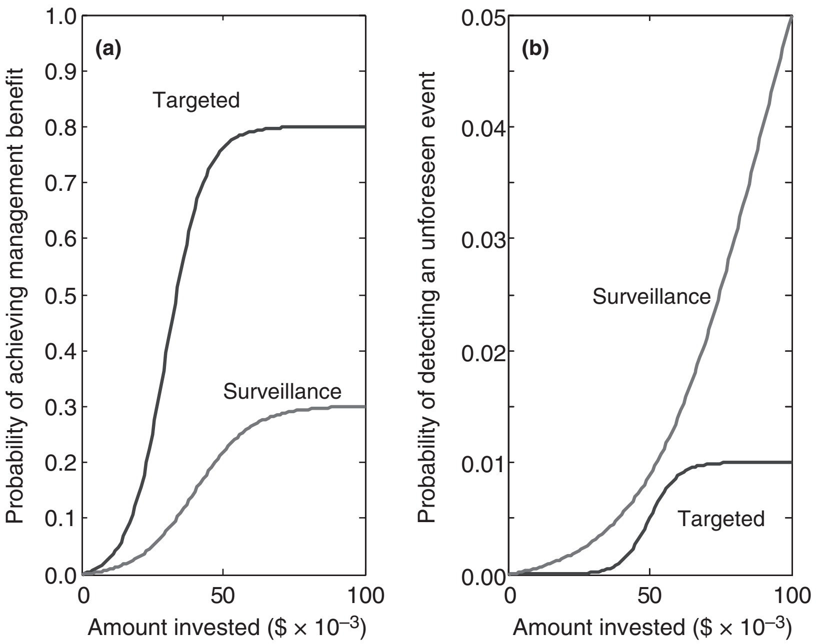

Figure 2 Monitoring efficiency functions for an heuristic example. (a) The probability of achieving management benefit, q_i, as a function of the amount invested in any one type of monitoring program. (b) The probability of serendipitous discovery of an unforeseen event, p_i, as a function of the amount invested in any one type of monitoring program.

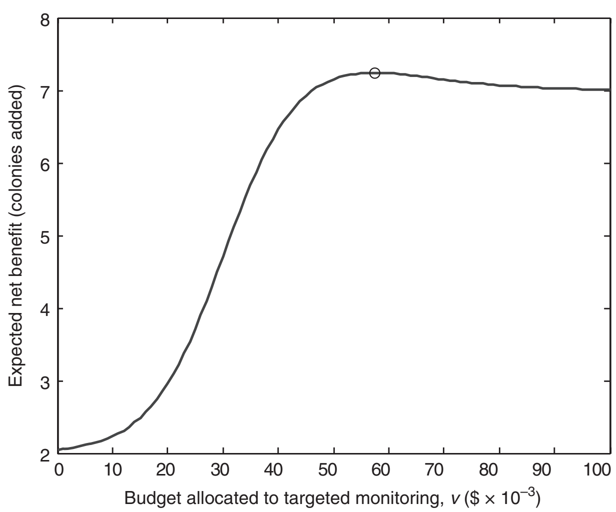

Figure 3 Expected net benefit (in colonies added) as a function of the amount allocated to targeted monitoring from a total budget of $100 000. For this heuristic example, the optimal allocation occurs when 57% of the budget is allocated to targeted monitoring, and the remainder to surveillance monitoring.

especially if the slopes of the p_2 and q_2 curves are fairly steep at low levels of investment.

Model parameter estimation and implementation

We see three broad styles of application of our model. First, the model may be used as a heuristic tool, to increase clarity of thought and discussion around monitoring planning and design. Without estimating a single parameter, the model can be used as a checklist of considerations necessary for sensible monitoring investment decisions. The model makes explicit many considerations and trade-offs that tend to be implicit in decisions about monitoring investment.

The second broad approach to using our model is a tool for sensitivity analysis, where model parameters are estimated as precisely as is defensible given available data and experience among the designers. Estimates of utility or expected benefit $V(v)$ may then be used in an informal sensitivity analysis setting to ask 'what-if' questions about the particular circumstances of the monitoring design. For example, managers of a national park might use the model to argue that a surveillance program for novel disease or pest incursion can be justified only if the cost of failing to rapidly detect such an incursion is more than 1000 times as great as the management benefits arising from improved knowledge about the best way to limit the impact of a known fungal pathogen, and if a surveillance monitoring strategy is twice as likely as the targeted program to detect such an outbreak. The robustness of the choice of monitoring investment to deviations from assumed parameter values can be explored by individually or simultaneously varying model parameters and observing the subsequent change in the expected benefit values $V(v)$. Mixed investments in surveillance and targeted monitoring could be evaluated using the same approach.

The third type of model application would involve a formal uncertainty analysis that explores the full space of monitoring investment options and parameter uncertainties to identify the most robust monitoring investment (*sensu* Wald 1945, Ben-Haim 2006). A formal uncertainty analysis

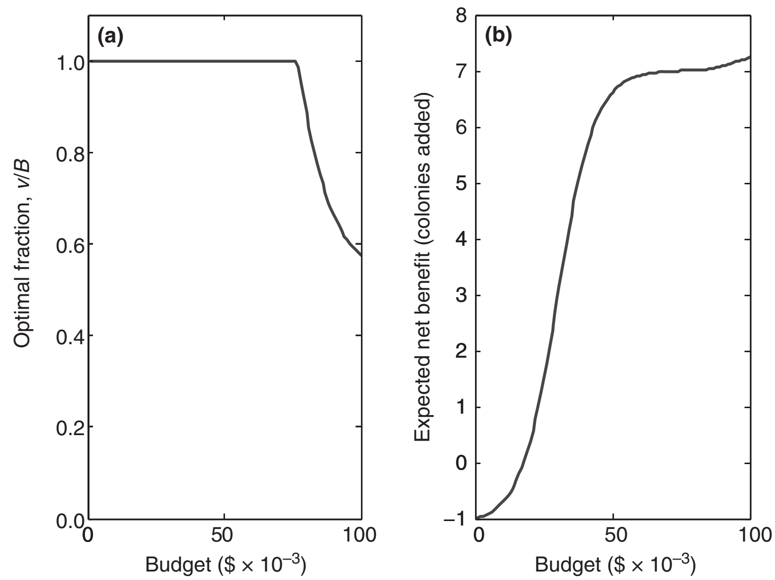

Figure 4 Optimal allocation of monitoring effort as a function of the total budget, for an heuristic example. (a) Optimal fraction of budget to allocate to targeted monitoring, as a function of the total budget. (b) Expected net benefit of the optimal monitoring allocation (in expected number of colonies added), as a function of the total budget.

would identify the robust-optimal monitoring investment strategy that achieves some minimum performance criteria under the most extreme scenario of parameter estimation error, or the widest range of possible parameter values, depending on the preferred definition of robustness.

Estimation of the parameters of our model will require a blend of standard statistical approaches and expert-elicited estimation or bounding. The key parameter q_i, the probability that a management benefit will be achieved given a particular monitoring design and investment, can be thought of as the statistical power of a given monitoring design, computed as a function of the background (or 'natural') variability of the state variable, the magnitude of the effect of interest, the sample size possible given a particular budget, the time period over which measurement occurs, and the rate at which false-alarms are tolerated (e.g. the alpha-level in classical hypothesis testing). The q_i in our model can be obtained with little additional effort beyond power estimation, which is considered a basic requirement of monitoring design (Taylor & Gerrodette 1993).

The magnitude of management return arising from the findings of a given monitoring design (R), should also be considered under standard monitoring design conditions. The parameter R addresses the question; 'What do we expect to gain from doing this monitoring well?' In targeted-monitoring designs, this parameter is central to determining whether or not it is worth spending the money on the monitoring at all (Chades *et al.* 2008). This parameter would usually be determined by expert consideration of how management efficiency would be expected to improve in light of a positive monitoring result. For example, a manager may expect to save \$$R$/year of their weed control budget if their monitoring program confirms that burning is a more cost efficient means of controlling an annual weed than herbicide application.

The most difficult parameters in our model to estimate are the parameters of surprise: the expected frequency of unforeseen ecological patterns or phenomena, the costs of those surprises, and the probability of discovering them with a given monitoring design (f, C, p_i). By definition, surprises about ecological phenomena are not predictable and their impacts cannot be known *a priori*. However, that does not mean that the historical rate at which surprises occur cannot be observed and estimated, especially given some careful bounding of the problem. For example, it may be possible to explore historical data within a domain of interest (e.g. novel diseases such as DFTD) over a fixed period to estimate a frequency distribution of costly disease surprises. More generally, it may be possible to estimate how often non-budgeted, unplanned remediation actions were required to urgently address a novel threat. We argue that within an administrative jurisdiction or a discipline, this frequency could be estimated to within an order of magnitude, making it possible to commence a sensitivity or robustness analysis. In short, these sorts of unpredictable events occur regularly enough that it should be possible to put bounds around the frequency with which they occur, even if the exact nature of the events/phenomena/threats cannot be predicted. The same logic can be applied to estimating the distribution of costs (C) associated with unforeseen threats.

Finally, the most challenging of the surprise parameters to estimate is p_i, the probability of detecting (and averting) an emerging threat or new ecological pattern under a given monitoring strategy or design. In some domains, it may be possible to use historical data on the emergence of novel threats or phenomena and the amount of surveillance and targeted monitoring that was invested over the period. This information could be used to estimate a relationship between the effort (e.g. money or hours) expended on monitoring and the probability of detecting the phenomenon. The simplest example arises from medical epidemiology. With the availability of good epidemic data, it is possible to estimate the number of novel disease outbreaks in a given time period (say, 20 years) and the approximate time between first emergence and detection. It is also possible to track how much money was spent over that period on both disease surveillance monitoring and targeted human health monitoring programs (e.g. monitoring of the effectiveness of HIV treatments). With those data at hand, it is possible to estimate, for a given monitoring effort and approach, the probability that a novel disease would be detected within a certain period since emergence. In this example, the parameter p_i would increase as a function of the time since disease emergence because more and more people would be showing novel symptoms. This sort of non-constant hazard function (or time dependence) is commonly estimated using standard survival analysis techniques (Harrell 2001).

Time dependence is an important consideration in the estimation of p_i and q_i in ecology. The length of time the monitoring program is undertaken relative to the time scales over which change occurs determines the probability (p_i or q_i) that a given monitoring design will provide powerful and timely information about unforeseen ecological change or the effectiveness of management interventions. The relationship between time, p_i and q_i, the magnitude of ecological change, and the potential costs of those changes (C) is complex, but not intractable within the logical structure that we have provided.

DISCUSSION

The model we have developed provides a framework for analysing the trade-off between improved management arising from the resolution of *known* unknowns, and avoided costs (or windfalls) arising from the timely discovery of *unknown* unknowns. Under our model, novel discoveries or resolution of *a priori* hypotheses can arise from any monitoring design, ranging from highly targeted through to classical surveillance and any number of mixed strategies along a continuum between these two extremes. When evaluating different monitoring designs, our model incorporates the probability that improvements to management will arise (q_i), that surprises will be uncovered (p_i), and the costs and benefits of these outcomes. We might expect q_i to be higher for targeted-monitoring designs, and p_i to be higher for surveillance monitoring designs. But these are contested assumptions (Nichols & Williams 2006). The point of the model is to provide a transparent way to evaluate the importance of these assumptions (and assumptions about f, C and R) in determining the best use of limited monitoring budgets. We define the best use of a limited monitoring budget as the one that brings the greatest benefit for the lowest investment.

The terms in our model can be thought of as a high-level set of design parameters. For example, the probability of detecting an effect of interest to managers under targeted monitoring (q_1) is derived from the parameters of statistical power analysis (i.e. n, α, Δ; Taylor & Gerrodette 1993). Any monitoring design consideration can be couched in terms of its influence on one or more of the parameters of our model. We have deliberately avoided a full explication of all of the 'sub-models' that could be developed to describe influences on the parameters of our model. We recognize that the task of constructing these models for each unique situation is not trivial. We have attempted to provide guidance on how our model parameters could be estimated to use the model analytically. The model brings greater clarity and structure to arguments about the purpose of monitoring and the relative merits of competing approaches.

Recent reviews of ecological monitoring have argued that surveillance monitoring is a poor investment compared with monitoring targeted to inform specific management options. So why do many managers invest in surveillance monitoring programs? Do they feel that 'keeping a finger on the pulse' is somehow worthwhile even if the benefits are difficult to quantify and, therefore, to include in a cost-benefit analysis? Do they intuitively consider the advantages that surveillance can provide to hypothesis generation or to the discovery of unforeseen ecological phenomena? Or are they being lazy in failing to clearly articulate the precise purpose of their monitoring and optimizing the design for that purpose?

We argue that the decision to invest in a particular monitoring design should rest firmly on a rational evaluation of how the design influences probabilities of achieving clearly articulated outcomes. Our model allows us to explore the relative merits of competing monitoring investment strategies. We found that a decision to invest in surveillance monitoring may be rational (Figs 1 and 3), if there is a belief that (1) the surveillance design is more likely to discover something of great importance than a more targeted, management-oriented design (Fig. 2b) and (2) the expected benefits (or avoided costs) arising from a new discovery are substantially higher than the benefits arising from improvements to management, resulting from targeted monitoring findings.

How often has this logic been clearly spelt out in the decision to invest in a surveillance monitoring study? We were not able to find an example of where this was the case (see also Yoccoz *et al.* 2001). A possible explanation for this is that it is difficult to estimate the parameters of our decision model. We have described three general approaches to using our model. Perhaps the most tractable use of the model is as a communication tool for increasing clarity of thought and discussion about the purpose and design of monitoring programs. An adjunct benefit of using the model as the basis for discussion is that it can work as a high-level inventory of the things that must be considered during monitoring design, ranging from the objectives of the monitoring through to the environmental and social influences on p_i and q_i. The use of the model as an analytical tool for quantifying and comparing the benefits of competing monitoring investment options is a more challenging prospect. The magnitude of uncertainty surrounding most of the model parameters demands that the model be utilized within a sound uncertainty analysis framework. Relevant approaches might include (but are not limited to) robust optimization, maxi–min or info-gap decision theories (Wald 1945; Ben-Tal & Nemirovski 2002; Ben-Haim 2006).

There is a demonstrable advantage in being specific about the objectives of a monitoring program because it is then possible to choose a monitoring design that is optimal with respect to those objectives (e.g. Johnson *et al.* 1997). For these reasons, we encourage conservation managers to make a serious attempt at clearly articulating the objectives of monitoring and state the actions that will be triggered by a given data signal (*sensu* Nichols & Williams 2006). This paper is not intended to provide a motivation for people to underplay the management-oriented objectives of their monitoring program in favour of emphasizing the value of the program for detecting or corroborating unforeseen ecological phenomena, or to encourage managers to be vague about the specific objectives of their monitoring program. Rather, we wanted to point out that while there *may* be a rational basis for a decision to invest in a surveillance monitoring program, the conditions under which surveillance is a defensible strategy are not easily met (eqn 6). If surveillance monitoring is chosen, the rationale for choosing it should be clearly spelt out prior to making such an investment. Our model provides the means for exploring and supporting that rationale.

When is surveillance monitoring just poorly designed management monitoring?

The importance of clearly articulating the objective(s) of a monitoring program cannot be over-emphasized. If the intention of a surveillance program is discovery of new ecological phenomena, development of new ideas, or early detection of novel threats, the design of such a program requires some consideration of the *type* of surprise that is anticipated. A surveillance monitoring program aimed at cave-hibernating bats has, in its definition, made choices about the resources of importance and the types of unknown unknowns that might be relevant or interesting. This suggests that there *may* be a nascent management objective or hypothesis underlying surveillance. We believe it is important to articulate this objective as clearly as possible, and ask whether the proposed monitoring design is the best way to satisfy it.

Arguments in favour of surveillance monitoring have highlighted the role of those programs in the identification of declining species, resulting in their listing under threatened species legislation (Carter *et al.* 2000; Bart *et al.* 2004). This implies the fundamental objective of conserving all species considered by the monitoring program. It also conveys an assumption that early recognition of a decline can lead to identification of the process(es) causing the decline, and initiation of mitigation efforts. Using standard statistical approaches, it should be possible to: (1) evaluate the likelihood that the monitoring design will lead to sufficiently early warning of decline and (2) to focus monitoring on learning about the most likely agents of decline. In short, managers of such monitoring programs could treat their program as management-targeted rather than classical surveillance, even if the number of species monitored is high and the spatial and temporal scope is broad.

Maximizing the value of management-targeted monitoring

If conservation budgets are so limited that there is little scope for investment in surveillance monitoring, what can be done to maximize the chances that the combined targeted programs have the greatest possible chance of leading to important discovery of unknown unknowns? Some simple answers lie in the use and availability of existing data. First, analysing the data collected is an important step towards hypothesis generation and discovery; vast amounts of monitoring data are never actually analysed, representing a waste of (usually) public money. Second, the availability of data for analysis by other analysts is also likely to increase the probability that important surprises will be uncovered. Greater value arises by making data publicly available on the web so that it is exposed to a world of modellers, web-crawlers and data-dredgers (Sutherland *et al.* 2004; Galaz *et al.* 2009). Global science databases and search engines could be used as a form of ecological monitoring for detection of novel phenomena (Galaz *et al.* 2009 and see http://search.driver.research-infrastructures.eu; Ginsberg *et al.* 2009). The requirement that all data collected under publicly funded monitoring and research programs be made available on the web (as is the case for the US LTER program) may increase our ability to cheaply discover unknown unknowns.

CONCLUSION

Increasing pressure on science and conservation budgets demands that investments in monitoring be as rigorously justified as possible. We support the argument that the design of a monitoring program should follow a rational, structured process that involves a clear articulation of the purpose of the program. We have provided a framework for such an examination of any monitoring program, and encourage those advocating surveillance monitoring to ground their justifications in the framework we have presented.

ACKNOWLEDGEMENTS

This work arose in conversation with participants of the Applied Environmental Decision Analysis (AEDA) working group on optimal monitoring. We make particular acknowledgement of the contributions of Yohay Carmel, Yakov Ben-Haim, Clare Hawkins, Adrian Manning, Mark Antos, Mark Burgman and Hugh Possingham. We are grateful to Jim Nichols and Georgia Garrard for comments on an earlier draft. Reviews by Nigel Yoccoz, Jonathan Rhodes and two anonymous referees substantially improved the manuscript. BW and SB were supported by ARC grants DP0774288 and LP0882780, and the Commonwealth Department of Environment, Water, Heritage and the Arts via AEDA. MR was supported by AEDA, the Australian Centre for Excellence in Risk Analysis and the US Geological Survey.

REFERENCES

Bart, J., Burnham, K.P., Dunn, E.H., Francis, C.M. & Ralph, C.J. (2004). Goals and strategies for estimating trends in landbird abundance. *J. Wildl. Manage.*, 68, 611–626.

Ben-Haim, Y. (2006). *Info-gap Decision Theory*, 2nd edn. Elsevier Ltd., Oxford, UK.

Ben-Tal, A. & Nemirovski, A. (2002). Robust optimization – methodology and applications. *Math. Programming Ser. B*, 92, 453–480.

Blehert, D.S., Hicks, A.C., Behr, M., Meteyer, C.U., Berlowski-Zier, B.M., Buckles, E.L. *et al.* (2009). Bat white-nose syndrome: an emerging fungal pathogen? *Science*, 323, 227.

Buckles, K.S. & Hungerman, D.M. (2008). Season of birth and later outcomes: old questions, new answers. Working paper 14573. National Bureau of Economic Research, Cambridge, USA.

Carter, M.F., Hunter, W.C., Pashley, D.N. & Rosenberg, K.V. (2000). Setting conservation priorities for landbirds in the United States: the Partners in Flight approach. *Auk*, 117, 541–548.

Chades, I., McDonald-Madden, E., McCarthy, M.A., Linkie, M., Wintle, B.A. & Possingham, H.P. (2008). Save, survey or surrender: optimal management of a cryptic threatened species. *Proc. Natl. Acad. Sci. USA*, 105, 13936–13940.

Chamberlin, T.C. (1897). The method of multiple working hypotheses. *J. Geol.*, 5, 837–848.

Driessen, M.M. & Hocking, G. (2008). Review of wildlife monitoring priorities. Resource Management and Conservation Division, Department of Primary Industries and Water Hobart, Tasmania, Australia, p. 50.

Driessen, M.M. & Hocking, G.J. (1992). *Review and Analysis of Spotlight Surveys in Tasmania: 1975–1990*. Department of Parks, Wildlife and Heritage Hobart, Tasmania, Australia.

Furlong, R.B. (1984). Clausewitz and Modern War Gaming. *Air Univ. Rev.*, 35, 4–7.

Galaz, V., Crona, B., Daw, T., Bodin, O., Nystrom, M. & Olsson, P. (2009). Can web crawlers revolutionize ecological monitoring. *Front. Ecol. Environ.*, 8, 99–104.

Ginsberg, J., Mohebbi, M.H., Patel, R.S., Brammer, L., Smolinksi, M.S. & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, 457, 1012–1015.

Greenwood, J.D. (2002). The monitoring of British breeding birds: a success story for conservation science? *Sci. Total Environ.*, 310, 221–230.

Halpern, J.Y. (2003). *Reasoning About Uncertainty*. MIT Press, Cambridge, MA.

Harrell, F.E. (2001). *Regression Modeling Strategies: With Application to Linear Models, Logistic Regression, and Survival Analysis*. 1st edn. Springer, New York, NY.

Hauser, C.E., Pople, A.R. & Possingham, H.P. (2006). Should managed populations be monitored every year? *Ecol. Appl.*, 16, 807–819.

Hawkins, C.E., Baars, C., Hesterman, C., Hocking, G.J., Jones, M.E., Lazenby, B. *et al.* (2006). Emerging disease and population decline of an island endemic, the Tasmanian devil *Sarcophilus harrisii*. *Biol. Conserv.*, 131, 307–324.

Hilborn, R. (1987). Living with uncertainty in resource management. *North Am. J. Fish. Manag.*, 7, 1–5.

Hobbie, J.E., Carpenter, S.T., Grimm, N.B., Gosz, J.R. & Seastedt, T.R. (2003). The US Long Term Ecological Research program. *Bioscience*, 53, 21–32.

Hubbell, S.P. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press, Princeton, NJ.

Hubbell, S.P. & Foster, R.B. (1986). Biology, chance, and history and the structure of tropical rain forest tree communities. In: *Community Ecology* (eds Diamond, J. & Case, T.J.). Harper and Row, New York, pp. 314–329.

Hume, D. (1748). *An Enquiry Concerning Human Understanding*. Clarendon Press, Oxford, UK.

IPCC (2001). *Mitigation. Contribution of Working Group III to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK.

Johnson, F.A., Moore, C.T., Kendall, W.T., Dubovsky, J.A., Caithamer, D.F., Kelley, J.R. *et al.* (1997). Uncertainty and the management of mallard harvests. *J. Wildl. Manage.*, 61, 202–216.

Knight, F.H. (1921). *Risk, Uncertainty, and Profit*. Houghton Mifflin Co, Boston, USA.

Likens, G.E. & Bormann, F.H. (1974). Acid rain: a serious regional environmental problem. *Science*, 184, 1176–1179.

Likens, G.E., Bormann, F.H., Johnson, N.M., Fisher, D.W. & Pierce, R.S. (1970). Effects of forest cutting and herbicide treatment on nutrient budgets in the hubbard brook watershed-ecosystem. *Ecol. Monogr.*, 40, 23–47.

Likens, G.E., Bormann, F.H. & Johnson, N.M. (1972). Acid rain. *Environment*, 14, 3–40.

McCarthy, M.A. & Possingham, H.P. (2007). Active adaptive management for conservation. *Conserv. Biol.*, 21, 956–963.
McDonald-Madden, E., Probert, W., Hauser, C., Runge, M., Jones, M., Moore, J. *et al.* (2010). Learning latency: the use of active adaptive management to conserve an iconic threatened species. *Ecol. Appl.*, 20, 1476–1489.
Nichols, J.D. & Williams, B.K. (2006). Monitoring for conservation. *Trends Ecol. Evol.*, 21, 668–673.
Nichols, J.D., Runge, M.C., Johnson, F.A. & Williams, B.K. (2007). Adaptive harvest management of North American waterfowl populations: a brief history and future prospects. *J. Ornithol.*, 148, 343–349.
PECBMS (2009). *The State of Europe*'s Common Birds 2008. CSO/RSPB. Prague, Czech Republic.
Platt, J.R. (1964). Strong inference. *Science*, 146, 347–353.
Rumsfeld, D.H. (2002). Donald Rumsfeld press conference as US secretary of Defense. [WWW Document]. Available at: http://www.defense.gov/transcripts/transcript.aspx?transcriptid=2636.
Sauer, J.R., Hines, J.E., Thomas, I., Fallon, J. & Gough, G. (2000). *The North American Breeding Bird Survey, Results and Analysis 1966–1999*. USGS Patuxent Wildlife Research Center, MD, USA.
Sauer, J.R., Hines, J.E. & Fallon, J. (2005). *The North American Breeding Bird Survey, Results and Analysis 1966–2005*. USGS Patuxent Wildlife Research Center, MD, USA.
Schneider, S.H., Turner, B.L. & Garriga, H.M. (1998). Imaginable surprise in global change science. *J. Risk Res.*, 1, 165–185.
Shea, K. & The NCEAS Working Group on Population Management (1998). Management of populations in conservation, harvesting and control. *Trends Ecol. Evol.*, 13, 371–375.
Sutherland, W.J., Pullin, A.S., Dolman, P.M. & Knight, T.M. (2004). The need for evidence-based conservation. *Trends Ecol. Evol.*, 19, 305–308.
Szymanski, J.A., Runge, M.C., Parkin, M.J. & Armstrong, M. (2009). *White-Nose Syndrome Management: report on Structured Decision Making Initiative*. Department of Interior, U.S. Fish and Wildlife Service Fort Snelling, MN, USA.
Taleb, N.N. (2007). *The Black Swan: The Impact of the Highly Improbable*. Random House, New York, NY.
Taylor, B.L. & Gerrodette, T. (1993). The uses of statistical power in conservation biology: the Vaquita and Northern Spotted Owl. *Conserv. Biol.*, 7, 489–500.
Thoreau, H.D. (1854). *Walden; or, Life in the Woods*. Ticknor and Fields, Boston, MA.
Trocki, C.L. & Paton, P.W.C. (2005). *Developing a Conservation Strategy for Grassland Birds at Saratoga National Historical Park*. Department of Natural Resources Science, University of Rhode Island, RI, USA.
USFWS (2009). *Adaptive Harvest Management: 2009 Hunting Season*. U.S. Fish and Wildlife Service, U.S. Dept. Interior. Washington, DC, p. 52.
Vickery, P.D. & Herkert, J.R. (2001). Recent advances in grassland bird research: where do we go from here? *Auk*, 118, 11–15.
Wald, A. (1945). Statistical decision functions which minimize the maximum risk. *Ann. Math.*, 46, 265–280.
Walters, C.J. (1986). *Adaptive Management of Renewable Resources*. MacMillan, New York, NY.
Wildlife Habitat Management Institute (1999). Grassland birds. *Fish and Wildlife Habitat Management Leaflet 8*. Natural Resources Conservation Service, U.S. Department of Agriculture, Madison, MO.
Yoccoz, N.G., Nichols, J.D. & Boulinier, T. (2001). Monitoring of biological diversity in space and time. *Trends Ecol. Evol.*, 16, 446–453.

Editor, Stephen Polasky
Manuscript received 21 January 2010
First decision made 28 February 2010
Second decision made 3 June 2010
Manuscript accepted 22 June 2010