Biological Conservation 144 (2011) 1214–1223

Contents lists available at ScienceDirect

Biological Conservation

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

Special Issue Article: Adaptive management for biodiversity conservation in an uncertain world

Which uncertainty? Using expert elicitation and expected value of information to design an adaptive program

Michael C. Runge [a,b,*], Sarah J. Converse [a], James E. Lyons [c]

[a] US Geological Survey, Patuxent Wildlife Research Center, 12100 Beech Forest Rd, Laurel, MD 20708, United States
[b] Australian Centre of Excellence for Risk Analysis and the Applied Environmental Decision Analysis Hub, University of Melbourne, Victoria 3010, Australia
[c] US Fish and Wildlife Service, Division of Migratory Bird Management, 11510 America Holly Dr, Laurel, MD 20708, United States

ARTICLE INFO

Article history:
Received 16 December 2009
Received in revised form 21 December 2010
Accepted 26 December 2010
Available online 3 February 2011

Keywords:
Adaptive management
Expected value of information
Expert elicitation
Whooping crane
Grus americana

ABSTRACT

Natural resource management is plagued with uncertainty of many kinds, but not all uncertainties are equally important to resolve. The promise of adaptive management is that learning in the short-term will improve management in the long-term; that promise is best kept if the focus of learning is on those uncertainties that most impede achievement of management objectives. In this context, an existing tool of decision analysis, the expected value of perfect information (EVPI), is particularly valuable in identifying the most important uncertainties. Expert elicitation can be used to develop preliminary predictions of management response under a series of hypotheses, as well as prior weights for those hypotheses, and the EVPI can be used to determine how much management could improve if uncertainty was resolved. These methods were applied to management of whooping cranes (*Grus americana*), an endangered migratory bird that is being reintroduced in several places in North America. The Eastern Migratory Population of whooping cranes had exhibited almost no successful reproduction through 2009. Several dozen hypotheses can be advanced to explain this failure, and many of them lead to very different management responses. An expert panel articulated the hypotheses, provided prior weights for them, developed potential management strategies, and made predictions about the response of the population to each strategy under each hypothesis. Multi-criteria decision analysis identified a preferred strategy in the face of uncertainty, and analysis of the expected value of information identified how informative each strategy could be. These results provide the foundation for design of an adaptive management program.

1. Introduction

Natural resource management is inevitably an exercise in decision-making under uncertainty. Natural systems are often complex, difficult to observe, highly variable, and understudied, so decisions about them are plagued by uncertainty about how the systems will respond to management actions. No wonder, then, that the concept of adaptive management has seemed like such a beacon, promising to improve long-term management outcomes by taking advantage of learning in the short-term. From its formulation in the field of fisheries in the 1970s (Walters, 1986), adaptive management has become a widely-cited principle of natural resource management, with applications in waterfowl harvest management (Nichols et al., 2007), invasive species management (Shea et al., 2002, 2005), translocation of threatened species (Rout et al., 2009), sustainable forestry (Wintle and Lindenmayer, 2008), revegetation (McCarthy and Possingham, 2007), and many other fields.

Applications of adaptive management emphasize the importance of uncertainty, and the literature includes several taxonomies of uncertainty (Morgan and Henrion, 1990; Regan et al., 2002). Of particular interest is epistemic uncertainty, structural and parametric uncertainty in models that arises due to our incomplete knowledge about how the system works and responds to management actions. Epistemic uncertainty can be an impediment to decision-making, but it is theoretically reducible through appropriate monitoring or research. Herein lies the promise of adaptive management—by addressing epistemic uncertainty in the short-term, we can improve management outcomes in the long-term.

But which uncertainty should we address? Knowing that epistemic uncertainty should be the focus of adaptive management is not enough. There is much we do not know about how ecological systems work, but not all of these uncertainties are important to resolve. How do we know where to invest monitoring and research? Here is where the application of adaptive management has been weak; often, some uncertainty will be identified and an argument will be made that it needs to be reduced, but no careful

* Corresponding author at: US Geological Survey, Patuxent Wildlife Research Center, 12100 Beech Forest Rd, Laurel, MD 20708, United States. Tel.: +1 301 497 5748; fax: +1 301 497 5545.
E-mail addresses: mrunge@usgs.gov (M.C. Runge), sconverse@usgs.gov (S.J. Converse), James_Lyons@fws.gov (J.E. Lyons).

0006-3207/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.biocon.2010.12.020

analysis is undertaken a priori to document the importance of reducing this uncertainty, and no comparative analysis of other sources of uncertainty is made. The risks in being lax in our applications of adaptive management are several: it can lead to unfocused monitoring, where investment is made in reducing many sources of uncertainty, without regard for the expected return on that investment; it can lead to unimportant and possibly costly experimental management, when probing actions are taken in an effort to reduce uncertainty; and it can undermine support for adaptive management, both in specific and general application, by investing scarce resources poorly and failing to noticeably improve management outcomes over time.

There are two key properties of the uncertainty that is important in adaptive management: it affects which management action is preferred, and it can be reduced through monitoring. In the face of important uncertainty, we do not know which action to take, because a different action is preferred under one hypothesis or another, or under one set of parameter values or another, and if we take the wrong action, our expected performance will be lower. Conversely, there are uncertainties that do not affect which management action is preferred. They may affect the expected performance, but we would take no different action if the uncertainty were resolved. In the language of classical decision theory, important uncertainty has a high expected value of information (Raiffa and Schlaifer, 1961). Important uncertainty also can be reduced through monitoring, with or without probing actions designed to accelerate learning. In some cases, uncertainty might be valuable to resolve, but we might not have the monitoring ability or enough resources to substantially reduce it.

The focus of this paper is on using the concept of expected value of information to identify uncertainty that is relevant in an application of adaptive management. After developing this concept more formally, we apply it to a case study involving the Eastern Migratory Population of whooping cranes (*Grus americana*).

1.1. Expected value of information

In classical decision theory, the value of new information is the difference between the expected value of an optimal action after the new information has been collected and the expected value of an optimal action before the new information has been collected (Raiffa and Schlaifer, 1961). The central concept is expressed in the expected value of perfect information (EVPI),

$$\mathrm{EVPI} = E_s[\max_a U(a,s)] - \max_a E_s[U(a,s)] \tag{1}$$

where a is an action taken, s represents a model of the system (a hypothesis about how the system works), and $U(a, s)$ is the utility associated with taking action a under model s (Raiffa and Schlaifer, 1961; Yokota and Thompson, 2004), where utility is the value ascribed by the decision-maker to the outcome, and is thus the measure of management performance. The first term in Eq. (1) represents the expected value once the uncertainty has been resolved, because the optimal action is chosen after knowing which model best describes the system (the expected value over the set of models s is needed in the calculation because prior to resolving the uncertainty, the decision-maker does not know which model will be identified). The second term in Eq. (1) represents the expected value in the face of uncertainty, because the optimal action is chosen without knowing which model best describes the system, thus, the decision-maker takes the expectation over uncertainty before choosing the optimal action. The difference in the terms, EVPI, is the expected improvement in management performance (measured on the scale of U) due to acquiring perfect information.

Calculation of EVPI requires articulation of alternative models of system behavior (these might be discrete models or a continuous set of alternative models described by parameter values), prediction of the outcome (U) of each action under each model, and a priori weights on the alternative models (that is, a probability distribution for uncertainty). While it might seem daunting to provide these elements, they are required already for quantitative application of adaptive management, and represent, in some sense, due diligence in describing the decision-maker's knowledge, and its limitations, about the system under management (Walters, 1986).

The expected value of perfect information is a powerful and useful tool, because it measures how important information is to improved management performance, that is, it measures information from the standpoint of the decision-maker. In economic applications, where the utility can be expressed in dollars or some other currency, the expected value of perfect information can be interpreted as the maximum amount a decision-maker would be willing to pay to acquire the information; this concept can be very valuable in evaluating whether it is worth establishing a monitoring program or funding a research study. In applications where the outcome is not monetary, EVPI is expressed in terms of the improvement in the metric of interest (perhaps survival rate or population size in an ecological setting); it is an additional step for the decision-maker to consider how much to pay for the information that leads to that improvement in performance. In an adaptive management setting, EVPI > 0 is a necessary condition for identifying uncertainty that should be the focus of monitoring and probing (Walters, 1986). It is not sufficient, however, because other conditions need to be considered, for example, whether it is possible to reduce the uncertainty through monitoring, and whether the short-term costs of acquiring the information are off-set by the long-term benefits of the information.

The value of information can also be used to compare the importance of difference sources of uncertainty, through the concept of partial EVPI, or the expected value of perfect x information (EVPXI, Yokota and Thompson, 2004). Let s_i be a subset of the uncertainty (say, a subset of the models), and s_i^c its complement. By resolving the uncertainty associated with just s_i, the performance is expected to increase by

$$\mathrm{EVPXI}(s_i) = E_{s_i}[\max_a E_{s_i^c}[U(a,s_i,s_i^c)]] - \max_a E_{s_i,s_i^c}[U(a,s_i,s_i^c)]. \tag{2}$$

The EVPXI can be compared for different subsets of uncertainty and can provide a measure, on a scale relevant to the decision-maker, of the value of reducing various components of uncertainty. In a decision-making context, this is the most appropriate method for sensitivity analysis (Felli and Hazen 1998, 1999), and it is also a powerful tool for identifying the uncertainty that matters in an adaptive management setting. For example, EVPXI could be used to determine whether it was more important to reduce uncertainty about, say, survival rates compared to reproductive rates in a population model. Note that the EVPXI for subsets of uncertainty are not necessarily additive, thus, the sum of the EVPXI is not EVPI, except in rare circumstances, but this does not undermine the utility of this metric.

It is rare to acquire perfect information, so a more relevant measure is the expected value of imperfect information (EVII), also known as the expected value of sample information (EVSI), which measures the expected improvement in performance from acquiring a sample of information,

$$\mathrm{EVSI} = E_x[\max_a E_{s|x}[U(a,s)]] - \max_a E_s[U(a,s)] \tag{3}$$

where x represents the sample information, and the first expectation is taken over the possible values of x that might arise from the monitoring program (Yokota and Thompson, 2004). Calculating EVSI requires a Bayesian preposterior analysis because it requires calculating posterior distributions (of $s|x$) for all possible values of the sample information, before the sample information is acquired.

(Preposterior analysis is the term that Bayesian statisticians often use in conjunction with experimental design (Berger, 1985). It emphasizes that one is anticipating the posterior distribution prior to collecting the data, by using the prior distribution to generate the likelihood of the observed data.) One very useful extension of EVSI is to evaluate the expected value of sample information for taking any particular action, *a*. This can be used to identify which actions are expected to reduce uncertainty most quickly.

The challenge with these value of information methods is that they require articulation of a set of models (hypotheses), a prior belief about the credibility of those models, and predictions of the outcomes under each model and action combination. That is, in order to evaluate what information is most valuable to collect, we already seem to have to know quite a bit about the system. Is this a cruel catch-22? At first, it seems to be, but all of those pieces are already implicitly required if the decision-maker wants to make a rational decision in the face of uncertainty. What's required, then, is merely an articulation of what is already in the mind of the decision-maker. For this, we turn to expert elicitation.

1.2. Expert elicitation

Of course, any decision-maker would prefer to have empirical evidence to develop the predictions of the outcomes, $U(a, s)$, and the prior probability distributions for the model set *s*, and sometimes such data are available. But for many natural resource management decisions, such data are not available and cannot be made available before an initial decision must be made. In these cases, it is becoming increasingly common to rely on expert judgment to develop estimates for the required quantities (Burgman, 2005; Kuhnert et al., 2010). Expert elicitation is a large and mature field of study in itself (Kahneman et al., 1982; Morgan and Henrion, 1990; O'Hagan et al., 2006), with a growing body of methods for robustly and efficiently eliciting and combining judgments from experts.

In ecological settings, the methods for expert elicitation most commonly used are the Delphi process and its variants, including the Nominal Group Technique (Delbecq et al., 1975; Linstone and Turoff, 1975; Lock, 1987; MacMillan and Marshall, 2006). These are group elicitation processes in which the experts are first asked for independent input on some parameter(s) of interest. These estimates are collated, and then shared with the whole group, along with the rationale and evidence that each expert used. The experts are then allowed to revise their estimates, if they so desire, to reflect the insights that arose from the group. This process can be repeated a number of times, until the experts are comfortable with their responses. The responses can then be aggregated in some manner, often by averaging, with retention of the range of responses to represent uncertainty in the parameter(s). The various methods differ primarily in whether and how they allow the group members to interact, and how they aggregate the individual responses. These methods of behavioral aggregation have been used in so many circumstances, and with so many modifications, that their nomenclature has become vague; the term "Delphi process" is now sometimes used to refer to any method in this general realm, rather than to the strict original procedure. But whatever they are called, these structured methods have the advantages that negative aspects of group dynamics (such as dominance and anchoring) can be avoided and a wide range of independent viewpoints can capture the underlying uncertainty. The insights of many can improve the thinking of individuals, and concerns about language-based misunderstanding can be identified and addressed.

There are detailed methods available to guide parameter elicitation, depending on the type of parameter and the specific concerns about bias that might be present (Kahneman et al., 1982; O'Hagan et al., 2006). These include methods for eliciting intervals or entire distributions as well as point estimates, and asking experts how confident they are in their responses, as a way of gauging uncertainty (Speirs-Bridge et al., 2010).

1.3. The Eastern Migratory Population of whooping cranes

Whooping cranes are listed as endangered under the US Endangered Species Act, the Canadian Species at Risk Act, and by the IUCN. As of late 2008 there were fewer than 400 individuals in the wild, and fewer than 600 individuals in existence (T. Stehn, US Fish and Wildlife Service [USFWS], unpublished data). Wild whooping cranes exist in three populations, including the Aransas-Wood Buffalo Population, which breeds at Wood Buffalo National Park, Northwest Territories, Canada, and winters at Aransas National Wildlife Refuge, Texas, USA; the introduced Florida Non-Migratory Population in central Florida, USA; and the introduced Eastern Migratory Population (EMP).

Beginning in 2001, releases of captive bred birds have been used to establish the EMP. Chicks hatched at the USGS Patuxent Wildlife Research Center's captive breeding facility (PWRC) in Maryland, USA, are trained from hatching for ultralight aircraft-led releases. This training consists of early imprinting of birds on costumed humans (such that they will follow costumed breeding center staff and pilots) and progressive familiarization and exercise behind grounded and, ultimately, airborne ultralight aircraft. Birds are shipped to the USFWS Necedah National Wildlife Refuge (NNWR), located in central Wisconsin, USA, in June or July of their first year (i.e., at 1–2 months of age). Ultralight aircraft lead the birds south during their first autumn to either Chassahowitzka or St. Marks National Wildlife Refuges in Florida, USA. Birds are soft-released there, and survivors return north on their own the following spring (Urbanek et al., 2010). Beginning in 2005, additional chicks have been hatched at the International Crane Foundation, in Wisconsin, USA, for direct autumn release. These birds are also imprinted on costumed humans, and shipped to a pen at NNWR, generally in early July (i.e., at <1–2 months of age). Birds are then released at or near NNWR during their first fall in the vicinity of previously released adult whooping cranes, with the hope that the released chicks will associate with the adult cranes and follow them on their first southward migration. Birds have been released every year since 2001 using the ultralight method and every year since 2005 using the direct autumn release method.

As of October 2009, the size of the EMP was approximately 86 birds (R. Urbanek, USFWS, unpublished data). Survival in the population (83.6% for unpaired birds, 99.1% for paired birds, and 94.1% for nesting birds) (Converse et al., submitted for publication) has been comparable to survival in the Aransas-Wood Buffalo Population (Link et al., 2003) which has been generally increasing for the past 70 years. However, reproduction has been poor. Birds began forming pairs in the spring of 2004, and made the first nesting attempts in the spring of 2005. Through the spring of 2009, of 41 nests containing eggs, only three nests (all renesting attempts) had successfully hatched chicks, and only one of those produced a fledgling (R. Urbanek and R. King, USFWS, unpublished data). Abandonment of nests during incubation appears to be by far the most common proximal cause of nest failure. The cause(s) of abandonment are currently unknown.

2. Methods

A workshop was held March 24–27, 2009 at NNWR and attended by the Refuge Manager (the decision-maker), staff from NNWR, and crane experts from PWRC, USFWS, Florida Fish and Wildlife Conservation Commission, the International Crane

Foundation, Operation Migration, and the USGS Louisiana Fish and Wildlife Cooperative Research Unit. Workshop members were primarily chosen from the membership of the Whooping Crane Eastern Partnership, a consortium of public agencies and non-profit organizations dedicated to establishing a self-sustaining migratory whooping crane population in eastern North America. The workshop agenda used a PrOACT structure (Hammond et al., 1999) to identify the key elements of the decision faced by the Refuge Manager: Problem, Objectives, Alternatives, Consequences, and Tradeoffs. The authors of this paper served as the facilitators and decision analysts for the workshop.

During development of the predictions, a focus was placed on first identifying a set of hypotheses for nest failure. Formal elicitation of the consequences was then organized around the alternative hypotheses. A version of the Delphi process modified for face-to-face interaction, much like the Estimate-Talk-Estimate method of Gustafson et al. (1973) and Lock's (1987) general approach to group judgmental forecasting, was used to elicit predictions from the panel of experts. After definition of the metrics of interest, the experts were asked to independently predict the outcomes of applying each particular management action under each hypothesis. These first responses were collated and discussed in a face-to-face group setting. Experts were then allowed to modify their predictions, if they so desired, based on clarifications and insights that arose during discussion. The results were again collated and discussed, and a number of remaining linguistic uncertainties were identified. A third round of elicitation occurred by email after the workshop. Experts were asked to provide point estimates for each outcome of interest; they were not asked for interval estimates, nor their degree of confidence in their predictions. For this analysis, simple arithmetic averages across experts were used.

A priori weights on the hypotheses were elicited from the experts by giving each expert 40 points to allocate over the eight hypotheses, with the number of points representing the strength of their belief that the hypothesis was operating. The independent weights were collated, and the experts presented the evidence they thought was important in their assignment of weights. After discussion, the experts were allowed to modify the weights they placed on the objectives. The arithmetic average weight across experts was used for further analysis.

Weights on the multiple objectives were elicited by a process of swing weighting (von Winterfeldt and Edwards, 1986). For the measurable attribute associated with each objective, the range of possible values was identified from the elicitation of predictions, with the minimum and maximum taken over the hypotheses and the actions (but not the experts, as their values had already been averaged). From these ranges, a series of hypothetical scenarios was created: the baseline scenario received the worst score on all attributes; the remaining scenarios each scored best on one attribute and worst on the others. An expert panel (the same group as above, but now acting to reflect the values of the decision-makers) was asked to rank the hypothetical scenarios, then assign a score to each between 0 and 100, where 0 was the score attributed to the baseline scenario, and 100 the score attributed to their top-ranked scenario. In this manner, the scores represent how much any one expert wanted to see a particular attribute swing from its worst to its best score. The individual scores were converted to weights by dividing by the sum of the scores, and the weights were averaged over the experts.

The problem was solved using multi-criteria decision analysis (MCDA), specifically a variant of the simple multi-attribute ranking technique (SMART, Goodwin and Wright, 2004). The best action to take in the face of uncertainty was found as

$$a^* = \arg\max_a \sum_o w_o \sum_s w_s U(a, s, o) \tag{4}$$

where $U(a, s, o)$ was the expected response on measurable attribute o, under hypothesis s and action a. The value of resolving all uncertainty expressed by the hypotheses s was found by calculating the EVPI (Eq. (1)).

Several sensitivity analyses were performed, using the concept of value of information. The partial EVPI (EVPXI) for the eight individual hypotheses was calculated using Eq. (2). In this case, the partial information was assumed to be able to perfectly test that hypothesis only, thus the outcome was either confirmation of that hypothesis (thus, its weight became 1) or refutation (its weight became 0); in the latter case, the weights on the remaining hypotheses were rescaled in proportion to their original weights. The expected value of sample information (EVSI, Eq. (3)) was calculated for each action a. This calculation assumed the response (weighted over the multiple objectives) could be observed with a normally distributed sampling error. We investigated sampling errors corresponding to coefficients of variation between 0% and about 30%, rates typical in many natural resource management settings. The posterior distribution of the hypothesis weights was found using Bayes' theorem, and used in the preposterior analysis required for EVSI.

3. Decision framework

The management problem was framed as a multi-criteria decision analysis under uncertainty: what management strategy should be undertaken at NNWR to benefit the EMP of whooping cranes, as expressed through four objectives, in the face of uncertainty about what is causing reproductive failure? This analysis involves managing tradeoffs among several objectives, as well as examining the impact of several specific hypotheses for reproductive failure. To develop this framework, we needed to specify the objectives, articulate the hypotheses for reproductive failure, develop potential management strategies, and predict the outcomes of each strategy on each objective under the alternative hypotheses.

This decision problem is nested within a larger decision problem—management of wetland and associated habitats on NNWR for a variety of purposes, including crane breeding, but also including wetland integrity and function, wildlife habitat, public use, and operational efficiency. While recognizing that this larger context was the real situation he had to make decisions within, the Refuge Manager nevertheless wished to see how he might solve the more focused crane problem, to understand that well before embedding it within a more complicated decision analysis.

3.1. Objectives

The decision-maker, with guidance from the workshop participants, identified four objectives with regard to demographic performance of EMP cranes at NNWR. Each of these objectives contributes to the fundamental objective of the Whooping Crane Eastern Partnership, of which USFWS is a founding member: establishment of a self-sustaining migratory whooping crane population in eastern North America. First, in the long-term, there is a desire to provide suitable nesting sites on the Refuge for whooping cranes. The second objective is to maximize the reproductive success of the pairs breeding on NNWR. Third, the Refuge wants to maximize the survival of the birds during the breeding season while they are at NNWR (April 1–November 1). The fourth objective is to maximize the body condition of birds upon departure for the southward migration.

The expert panel developed measurable attributes for these four objectives. (1) The average number of territorial pairs on NNWR over time. Territorial pairs are bonded, exhibit one or more breeding behaviors, and occupy a core area on the Refuge. (2) The fledging rate, expressed as the average number of young that depart NNWR for southward migration per pair. (3) Average survival of white birds (adults and sub-adults) from April 1 to November 1. (4) Average body condition just prior to fall migration, as measured by an existing constructed scale (1–5) for body condition (Olsen et al., 1996).

The swing weighting was conducted after the elicitation of the predictions (because the range of responses was needed for the swing weighting exercise). The average weights placed on the four objectives were: number of pairs, 13.4%; reproductive success, 59.9%; adult survival, 16.3%; and body condition, 10.4%. The weights did differ somewhat across the panelists; a principal components analysis showed that the first component, which was driven by the weight on reproductive success, explained 75% of the variation in the weights across experts.

3.2. Hypotheses

The expert panel initially identified 34 hypotheses that could be posited to explain the reproductive failure of EMP cranes. Each member of the expert panel was asked to allocate 34 points among these hypotheses. The sum of scores across experts was used to rank the hypotheses, and the top eight were retained for further analyses. After elucidation and discussion of these eight hypotheses, the panel was asked to assign weight to each, based on their perception of the evidence. The eight hypotheses, with the average weights across experts, are briefly described below.

1. Too young (9.4%). Collectively, the EMP birds, which were all released into the wild as juveniles over the past 8 years, are not yet old enough to exhibit mature and successful reproductive behavior. Cranes are long-lived birds with a late age-at-first-reproduction; perhaps these cranes just need longer to exhibit successful reproduction.
2. Black flies (29.1%). Harassment of incubating birds by parasitic black flies causes nest abandonment, which leads to nest failure.
3. Social conditioning (11.9%). All of the birds released to the EMP were costume-reared, that is, they imprinted on humans dressed in costumes designed to conceal human features. The lack of experience with natural pre- and post-hatching conditions leads to poor association between the parents and the nest, then nest abandonment and poor nest defense.
4. Nutrient limitation on NNWR (22.8%). Energetic resources are limited at NNWR, leading to nutritional stress. Adults abandon reproductive attempts to increase their own survival.
5. Nutrient limitation during winter or migration (5.9%). Food resources are inadequate on the winter quarters or in the migratory habitat, thus, the adults arrive at NNWR in poor condition and cannot successfully sustain breeding effort.
6. Nutrient limitation on both NNWR and during winter or migration (6.6%). Arrival condition is poor and cannot be compensated by weight gain at NNWR.
7. Human egg salvage (4.4%). The nests at NNWR are watched carefully by biologists. When nests appear to be abandoned, the biologists will salvage the eggs for captive rearing. Perhaps the nests are not being abandoned, but rather just attended intermittently, and the egg salvage is actually the cause of nest failure.
8. Human disturbance (10.0%). Research activities, refuge operations, and/or public access are disturbing nesting pairs, leading to nest abandonment and failure.

3.3. Alternative management strategies

After the objectives had been identified and the hypotheses articulated, the management team, with guidance from the expert panel, developed seven potential management strategies. As a starting point for development, the group was asked to consider in turn what action would be most fruitful if each hypothesis was in fact the case. The strategies identified were:

1. Status quo, or wait (developed to address the "Too young" hypothesis). Continue to operate as in the past.
2. Kill flies (developed to address the "Black fly" hypothesis). Use insecticides (BTI) and strategic water management to reduce black fly populations during the crane breeding season.
3. Swap eggs (developed to address the "Social conditioning" hypothesis). Swap eggs in nests for eggs further along in incubation, supplied from the captive flock. This would allow adults to spend less time on the nests, yet still have wild-hatched chicks.
4. Restore meadows (developed to address the two hypotheses that involve nutrient limitation on NNWR). Conduct meadow restoration, with supplemental feeding in the interim, to increase food resources at NNWR for cranes.
5. April drawdown and burn (developed to address the two hypotheses that involve nutrient limitation on NNWR). Create a checkerboard of partial drawdowns and burning in impounded wetlands in early April, to increase food resources.
6. No salvage (developed to address the "Human egg salvage" hypothesis). Discontinue salvaging eggs, or at least develop a less invasive set of operational rules.
7. No disturbance (developed to address the "Human disturbance" hypothesis). Minimize the disturbance to nests during the breeding season; keep people outside of the maximum flush distance; eliminate egg swapping and egg salvage.

3.4. Predictions

The expert panel was asked to make predictions for the response on the four measurable attributes, of the seven strategies, under the eight hypotheses. The experts were asked to predict the response averaged over the 10 years after each strategy was fully implemented (assuming that implementation would begin in the year following this analysis). The average response was taken over the 10 experts who participated, after the third round of elicitation (Tables 1–4).

For the objective of maximizing the number of territorial pairs on the refuge, the best action to take was meadow restoration, under all hypotheses (Table 1). For the objective of maximizing reproductive success, the best action to take depended strongly on the hypothesis, with no salvage, no disturbance, April drawdown and burn, and swap eggs all chosen as optimal under one or more hypotheses (Table 2). Interestingly, the best action to take to achieve this objective, given the expected response over all hypotheses was restore meadows, an action not identified as optimal under any of the individual hypotheses. For the objective of maximizing adult survival during the breeding season, the best action depended on the hypothesis, with no salvage, restore meadows, and swap eggs all identified as optimal under individual hypotheses, and restore meadows best under the average response (Table 3). For the objective of maximizing body condition upon departure, the best action was restore meadows under six of the hypotheses, and swap eggs under the other two (Table 4) with restore meadows again best under the average response.

4. Results

The combined score, with each objective first normalized to a 0–1 scale, and then weighted across objectives, reveals that the best action to take depends on the hypothesis. In the face of uncertainty, the best action to take is meadow restoration (not surprising,

Table 1
Predicted number of pairs using NNWR in the long-term, as a function of the management strategy employed and underlying hypotheses about reproductive failure. These predictions are the unweighted average of the responses from 10 experts. The best strategy, conditional on each hypothesis, is shown in bold. The expected value is the weighted average over hypotheses, with weights determined by the same panel of experts.

Hypotheses		Strategies						
Description	Weight (%)	1 Wait	2 Kill flies	3 Swap eggs	4 Restore meadows	5 April DD and Burn	6 No salvage	7 No disturbance
Too young	9.4	14.1	14.0	14.2	**17.2**	14.7	14.4	13.7
Black flies	29.1	12.8	13.5	13.4	**16.3**	13.9	14.2	13.1
Social conditioning	11.9	13.1	13.0	13.2	**16.0**	13.5	13.4	13.8
Nutrient limitation: NNWR	22.8	13.1	13.0	13.2	**18.6**	15.7	13.2	13.4
Nutrient limitation: winter	5.9	13.1	13.0	13.2	**15.5**	13.3	13.2	13.4
Nutrient limitation: both	6.6	13.1	13.0	13.2	**17.2**	14.5	13.2	13.4
Egg salvage	4.4	13.1	14.0	14.0	**15.8**	13.3	13.2	12.8
Disturbance	10.0	13.1	13.8	14.7	**15.8**	13.2	13.2	12.7
Expected Value		13.11	13.36	13.54	**16.79**	14.25	13.60	13.27

Table 2
Predicted fledging rate, as a function of the management strategy employed and underlying hypotheses about reproductive failure. These predictions are the unweighted average of the responses from 10 experts. The best strategy, conditional on each hypothesis, is shown in bold. The expected value is the weighted average over hypotheses, with weights determined by the same panel of experts.

Hypotheses		Strategies						
Description	Weight (%)	1 Wait	2 Kill flies	3 Swap eggs	4 Restore meadows	5 April DD and burn	6 No salvage	7 No disturbance
Too young	9.4	0.246	0.216	0.235	0.244	**0.256**	0.223	0.181
Black flies	29.1	0.069	0.199	0.128	0.124	0.119	**0.204**	0.099
Social conditioning	11.9	0.074	0.097	0.108	0.139	0.141	0.113	**0.191**
Nutrient limitation: NNWR	22.8	0.074	0.094	0.140	0.289	**0.290**	0.103	0.116
Nutrient limitation: winter	5.9	0.074	0.089	0.125	0.144	**0.156**	0.103	0.116
Nutrient limitation: both	6.6	0.074	0.089	0.135	0.244	**0.245**	0.116	0.121
Egg salvage	4.4	0.094	0.226	**0.227**	0.147	0.128	0.166	0.105
Disturbance	10.0	0.091	0.154	**0.252**	0.127	0.108	0.106	0.150
Expected value		0.091	0.147	0.155	**0.185**	0.183	0.148	0.129

Table 3
Predicted survival rate, as a function of the management strategy employed and underlying hypotheses about reproductive failure. These predictions are the unweighted average of the responses from 10 experts. The best strategy, conditional on each hypothesis, is shown in bold. The expected value is the weighted average over hypotheses, with weights determined by the same panel of experts.

Hypotheses		Strategies						
Description	Weight (%)	1 Wait	2 Kill flies	3 Swap eggs	4 Restore meadows	5 April DD and burn	6 No salvage	7 No disturbance
Too young	9.4	0.928	0.927	0.930	**0.933**	0.929	0.929	0.928
Black flies	29.1	0.923	0.925	0.925	0.930	0.927	**0.936**	0.925
Social conditioning	11.9	0.928	0.927	0.930	**0.933**	0.929	0.929	0.932
Nutrient limitation: NNWR	22.8	0.923	0.921	0.925	**0.938**	0.932	0.924	0.924
Nutrient limitation: winter	5.9	0.928	0.926	0.930	**0.938**	0.934	0.929	0.929
Nutrient limitation: both	6.6	0.923	0.922	0.925	**0.935**	0.932	0.924	0.924
Egg salvage	4.4	0.928	0.935	**0.939**	0.934	0.930	0.929	0.927
Disturbance	10.0	0.926	0.933	**0.938**	0.932	0.928	0.927	0.925
Expected value		0.925	0.926	0.928	**0.934**	0.929	0.929	0.926

as this was the best action to take in the face of uncertainty under each objective), with an expected composite score of 0.590 (Table 5). This result is not sensitive to uncertainty in the weights on the objectives; all individual weightings from the advisory panel led to choice of the same action.

If uncertainty can be fully resolved before the action is chosen, the expected performance rises to 0.707, thus, the EVPI is 0.117, an increase in performance of 19.9%. For objective 1 (maximize territorial pairs) alone, the EVPI is 0, because the best strategy (restore meadows) is the same under all hypotheses (Table 1). For objective 2 (maximize reproductive success, Table 2) alone, the expected fledging rate increases from 0.185 to 0.232 (25.7%) through resolution of uncertainty. For objective 3 (maximize adult survival, Table 3) alone, the expected adult survival increases from 0.934 to 0.936 (0.3%), and for objective 4 (maximize fall body condition, Table 4) alone, the expected body condition increases from 3.050 to 3.057 (0.2%) through resolution of uncertainty.

The partial EVPIs due to individual hypotheses reveal that the most value of information comes from discerning hypothesis 2 (black flies) from all the others, followed by hypothesis 8 (human disturbance) (Fig. 1). As it turns out in this case, the partial EVPIs do sum to the total EVPI, and so the partial EVPIs can be shown as a pie chart (Fig. 2). Resolution of hypothesis 2 (black flies) provides 54% of the total EVPI, followed by hypotheses 8 (human disturbance, 32%), 7 (egg salvage, 8%), and 3 (social conditioning, 6%).

Using the expected value of sample information to evaluate how valuable each action is in generating knowledge for future decisions, we find the answer depends on the expected level of

Table 4
Predicted body condition on departure from NNWR, measured on a constructed 5-point scale (Olsen et al., 1996), as a function of the management strategy employed and underlying hypotheses about reproductive failure. These predictions are the unweighted average of the responses from 10 experts. The best strategy, conditional on each hypothesis, is shown in bold. The expected value is the weighted average over hypotheses, with weights determined by the same panel of experts.

Hypotheses		Strategies						
Description	Weight (%)	1 Wait	2 Kill flies	3 Swap eggs	4 Restore meadows	5 April DD and burn	6 No salvage	7 No disturbance
Too young	9.4	2.55	2.55	2.61	**2.92**	2.80	2.57	2.55
Black flies	29.1	2.45	2.65	2.76	**2.98**	2.81	2.97	2.55
Social conditioning	11.9	2.55	2.55	2.67	**2.93**	2.81	2.56	2.75
Nutrient limitation: NNWR	22.8	2.45	2.50	2.56	**3.42**	3.35	2.46	2.45
Nutrient limitation: winter	5.9	2.55	2.55	2.57	**2.88**	2.82	2.56	2.55
Nutrient limitation: both	6.6	2.45	2.50	2.59	**3.25**	3.21	2.51	2.45
Egg salvage	4.4	2.55	2.77	**2.83**	2.77	2.80	2.56	2.53
Disturbance	10.0	2.55	2.67	**2.81**	2.77	2.80	2.51	2.49
Expected value		2.49	2.59	2.67	**3.05**	2.95	2.65	2.54

Table 5
Weighted response across the four objectives, as a function of the management strategy employed and underlying hypotheses about reproductive failure. Each individual objective was normalized to a 0–1 scale before the weighed average was taken, thus this weighted scale is also between 0 and 1, with 1 representing the best possible performance simultaneously on all four objectives, and 0 representing the worst. The optimal performance in the face of uncertainty is 0.590; the expected performance with full resolution of uncertainty is 0.707, thus the EVPI is 0.117.

Hypotheses		Strategies						
Description	Weight (%)	1 Wait	2 Kill Flies	3 Swap eggs	4 Restore meadows	5 April DD and Burn	6 No salvage	7 No disturbance
Too young	9.4	0.586	0.491	0.581	**0.735**	0.663	0.539	0.402
Black flies	29.1	0.021	0.425	0.242	0.373	0.253	**0.589**	0.139
Social conditioning	11.9	0.093	0.145	0.220	0.429	0.321	0.218	**0.485**
Nutrient limitation: NNWR	22.8	0.036	0.081	0.254	**0.992**	0.863	0.128	0.166
Nutrient limitation: winter	5.9	0.093	0.119	0.260	**0.466**	0.405	0.185	0.223
Nutrient limitation: both	6.6	0.036	0.077	0.243	**0.792**	0.703	0.172	0.179
Egg salvage	4.4	0.147	0.622	**0.662**	0.436	0.291	0.354	0.158
Disturbance	10.0	0.120	0.393	**0.740**	0.363	0.216	0.168	0.256
Expected value		0.106	0.284	0.343	**0.590**	0.475	0.331	0.231

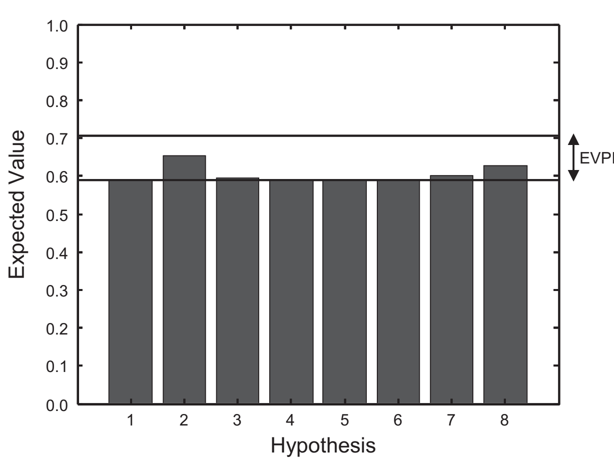

Fig. 1. Partial expected value of perfect information, as a function of the hypothesis resolved. The height of the bars shows the expected performance, on the normalized scale weighted over objectives. In each case, the information acquired allows a full test of the hypothesis, either confirming it or refuting it; in the later case, the remaining hypotheses are assumed to retain their proportional weights. The two reference bars are the expected performance in the face of uncertainty (0.590) and the expected performance under perfect information (0.707); the difference is the EVPI.

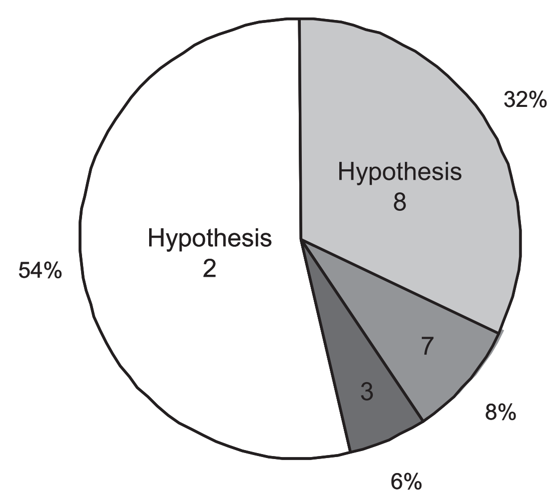

Fig. 2. Partial expected value of perfect information, as a proportion of the full EVPI, for each hypothesis. The partial EVPI for hypotheses 1, 4, 5, and 6 is less than 1% and not shown on the pie graph.

measurement error (Fig. 3). If the weighted response across the four objectives (which is expressed on a 0–1 scale) could be measured quite precisely, with a standard deviation of 0.03 or less, the most informative action would be strategy 2 (kill flies) and it would generate more than 60% of the EVPI. However, if the measurement error is expected to be greater than 0.03, the most informative strategy is 6 (no salvage), followed closely by 3 (swap eggs). The EVSI for strategy 4 (restore meadows) is never high, and drops off very quickly with increases in measurement error.

5. Discussion

In the face of uncertainty and without being able to gather more information, the recommended strategy at NNWR to benefit crane

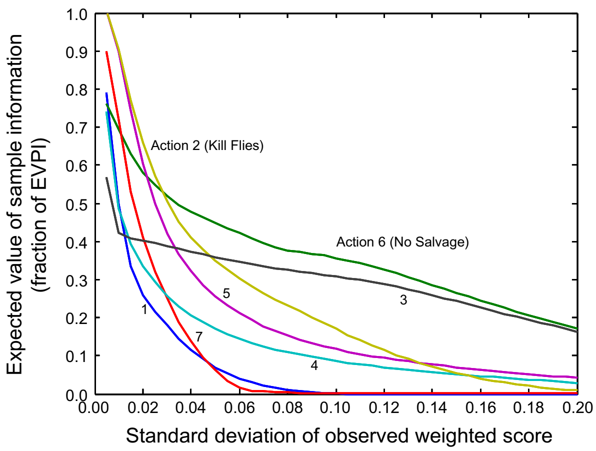

Fig. 3. Expected value of sample information, as a function of measurement error in the observed weighted score. The sample information comes from observing the response to one action. Because the weighted score takes values between 0 and 1, and the expected response for this problem was around 0.6, this range of standard deviations corresponds to coefficients of variation between 0% and about 30%, a range that captures typical errors in measurement of ecological quantities.

demography is meadow restoration. This result is not sensitive to uncertainty in the objective weightings, but it is sensitive to uncertainty about the causal mechanism for reproductive failure. The decision-maker does have the opportunity to gather more information before committing to a final course of action; in fact, the management decision is made and can be adjusted annually, so there is an opportunity to approach this problem adaptively. As adapting to new information is possible, what information is most important to acquire, and what strategies might be most informative in an adaptive setting?

First, the increased performance from reduction of uncertainty (the EVPI) is driven by the response of reproductive success, rather than any of the other three objectives, to management actions. Thus, a monitoring program designed to measure reproductive success will be most valuable; measuring the other response variables is not nearly as important for identifying the best course of action. Second, if an experimental approach is taken, it is most important to test hypotheses 2 (the black fly hypothesis) and 8 (the human disturbance hypotheses); the value of discerning the other hypotheses is minimal. A corollary to this conclusion is that the design of a formal adaptive management framework should focus on uncertainty about the black fly and human disturbance hypotheses, perhaps by including alternative models that differ in these respects; including uncertainty associated with the other hypotheses is not necessary. Third, in an adaptive approach, the most informative strategies are likely to be 6 (no salvage) and 3 (swap eggs)—these strategies can most robustly resolve the important uncertainties, although strategy 2 (kill flies) is most informative if the response can be measured quite accurately. Because the best strategy in the face of uncertainty (action 4) is not the same as the most informative strategy (actions 6 and 3), this is precisely the case when active adaptive management is going to be more beneficial than passive adaptive management (Walters, 1986). An active strategy will favor actions 6 and 3 in the short-term to quickly resolve uncertainty, then switch to the strategy that is optimal once the likely underlying hypothesis has been identified. An active dynamic optimization (McDonald-Madden et al., 2010; Williams, 1996) could be undertaken here to identify the optimal strategy as a function of the information state (the weights on the hypotheses).

In the introduction, we wrote of two key factors in identifying the uncertainty of concern in an adaptive management setting: high expected value of information, and high power of monitoring. We have not fully addressed the second factor for the crane case study, but as an adaptive program for management of cranes on NNWR is developed, the power of the monitoring methods to discern hypotheses needs to be investigated. The key response variable of interest, the fledging rate, is currently monitored with high precision; the more relevant question is whether an experimental or adaptive program can be designed that controls enough external variation to allow the competing hypotheses to be compared and evaluated. This is precisely the issue that is captured in the expected value of sample information. If the sampling variance is high, due to measurement error or uncontrolled stochasticity in the system, the expected value of information decreases. We have investigated the impact of sampling variance over a range that is typical in ecological studies, but we have not yet evaluated what level of sampling variance we could expect in this particular setting.

The crane decision problem we have framed is admittedly a simplification of a larger problem. Some of the details that we omitted may be important. First, we have been vague about the time-frame of the management strategies. Some of the strategies (like the black fly treatment strategy) can be implemented quickly and would be expected to generate responses quickly; others (like meadow restoration) might take years to implement and years more to show differences under the various hypotheses. These time-dependencies could have a strong effect on the value of information. Second, we have not factored in the costs of research and implementation. In part, whooping crane conservation is so important that costs are a minor concern compared to persistence and success of reintroduction; indeed, the American household willingness-to-pay value for whooping cranes is higher than that of bald eagles (Bowker and Stoll, 1988; Loomis and White, 1996). But for NNWR, with a limited annual budget, the costs of some management actions could be a factor. Further refinement of the decision analysis we have presented may be needed to fully capture the decision-makers' concerns.

In classical economic settings, EVPI is calculated on the primary scale of interest, dollars, and thus can be interpreted not only as the increased expected performance (profit) from acquisition of perfect information, but also as the maximum the decision-maker should be willing to pay for that information. If the cost of the information is greater than the EVPI, it is better to make the decision without resolving uncertainty first. The EVPI we have calculated is expressed on a dimensionless weighted multi-criteria scale, and it is not transparent how to translate this into a dollar-equivalent. How much is it worth to increase the weighted multi-criteria response from 0.590 to 0.707 (Table 5), or the expected fledging rate from 0.185 to 0.232 (Table 2)? Answering these questions requires the decision-maker to make a trade-off between biological performance and cost. There are several methods for doing this, one of which resembles the swing weighting methods used to weight the objectives in this example. We did not undertake such an analysis with the decision-maker; as noted above, whooping crane conservation is important enough that it seemed clear the value of information exceeds the cost of acquiring that information, without having to explicitly compare the biological and cost scales. In other settings, however, it may be valuable to make this explicit comparison.

We have formulated this crane problem as a multi-criteria decision analysis and our methods demonstrate the value of EVPI in the multi-criteria setting. But unlike in most multi-criteria settings, the objectives we have used may be viewed as means to a single fundamental objective of maximizing the persistence of the EMP. In this respect, an appropriate weighting among the means objectives could be attained by using a population model that predicted persistence as a function of number of pairs, fledging success, adult survival, and body condition on departure; in this way, the swing-weighting procedures would be unnecessary. On the other hand, separation of these components reflects the particular perspective of management at NNWR, and did allow us to identify the most important proximate life-history variables to monitor, something that could be lost by combining the proximate responses into a single ultimate response (like probability of persistence). A hybrid approach could be pursued in situations like this, one that follows the multiple proximate responses, but uses a population model to determine their weights.

We have focused on the uncertainty as expressed by the alternative hypotheses, using the average responses from the 10 experts to estimate those hypotheses. But there are two additional sources of uncertainty we could have pursued: the uncertainty across experts, and the confidence of individual experts. In the first case, we could look at whether the results of the analysis are sensitive to the individual responses of the experts, using methods similar to what we have shown here. In the second case, we would have needed to elicit the confidence of the experts in their predictions, which we did not do. Analysis at this level of detail is beyond the scope of this paper, but we do believe the uncertainty captured by the hypotheses is more important to the decision than the uncertainty across experts. There are reasons to believe, however, that the aggregate responses of the experts could have a few flaws. For example, the experts predicted that the swap eggs strategy would produce the highest survival rate and body condition under several of the hypotheses (Tables 3 and 4); we cannot explain the logic behind this prediction. As a second example, the best strategy under the black fly hypothesis for maximizing reproductive success is not killing flies, as expected, but preventing egg salvage (Table 2). Perhaps the experts did intend these patterns, but it is also possible that some linguistic uncertainty remained in the elicitation, that the experts had become fatigued and were not able to see all the logical patterns they wished to express in the predictions, or that the averaging across experts somehow produced a spurious pattern. These observations serve to emphasize that expert elicitation does need to be undertaken carefully and slowly, with multiple opportunities for feedback; another round of review with the experts would have been valuable to determine whether all the observed patterns in the predictions were intended.

We evaluated the uncertainty captured by eight hypotheses, but these were only the top ones among a set of 34, and the true reason may lie within that larger set, or indeed, in a hypothesis that has yet to be articulated. Does this undermine the analysis herein? On one hand, if the truth is intermediate between existing hypotheses, that is, a weighted average, then the EVPI calculations may overstate the value of information, because they assume that the truth is one of the stated extremes. On the other hand, if the truth lies far outside of the articulated hypotheses, then the EVPI calculations may understate the value of information. There is not a corresponding EVPI calculus for uncertainty that is not or cannot be articulated; anticipating the value of resolving "unknown unknowns" is a philosophical conundrum, but possibly an important consideration in monitoring design (Wintle et al., 2010).

The central point we wish to emphasize is that the EVPI calculus, in its various forms, is immensely valuable in identifying the important underlying uncertainties in a decision context, and should be common practice for designing adaptive management frameworks. While this calculus is not trivial, it only requires the same amount of information that a formal structured decision analysis would require.

It will seem counter-intuitive to some that the design of an adaptive management program requires the kind of information needed for an EVPI calculation, because one of the points of adaptive management is to acquire information in the process of management. An adaptive management framework does not require EVPI for its design, but an efficient one does. Without this kind of a priori analysis, it is quite possible to invest in the wrong monitoring, to focus on uncertainty among the wrong set of models, and to take the wrong actions to balance management and learning.

The design of an adaptive program, guided by EVPI calculations, does require some a priori predictions and model weights. In the absence of empirical data to build these tools, expert elicitation is a powerful companion tool. The most dedicated empiricists balk at the use of expert elicitation, and there is debate in the judgmental forecasting literature about the accuracy, precision, bias, and confidence of experts (the literature is large, Wright and Ayton, 1987 is one of many entrees). But there are reasons to embrace expert elicitation. First, it is commonly the only recourse a decision-maker has, when the context calls for a decision to be made, or an adaptive program to be initiated, before more information can be collected. Second, in an adaptive program, the predictions made by experts need only to be viewed as initial hypotheses; the monitoring program and the information feedback that arises from it form the empirical tests of the hypotheses.

6. Conclusions

In setting up an experimental or adaptive management program for management of breeding whooping cranes at Necedah National Wildlife Refuge, the most important uncertainty to address concerns management effects on the fledging success from nests of territorial pairs. The most valuable hypotheses to discern are the "black fly" and "human disturbance" hypotheses, and the most informative actions may be the fly insecticide and no salvage strategies, followed by the swap eggs strategy. While there are many uncertainties about the demography of cranes in the EMP, the reproductive behavior of captive-reared (and ultimately wild-hatched) cranes at NNWR, and the effects of various management strategies, not all uncertainty is relevant to decision-making. The components identified above should be central to future learning designed to improve management of this conservation icon.

We come to these conclusions by using the tools of expert elicitation and expected value of information to frame and analyze the particular decision problem at hand. More generally, we argue that these tools should have a central role in design of any adaptive management program.

Acknowledgments

We are grateful to Larry Wargowski, refuge manager at Necedah National Wildlife Refuge, for articulating the need and supporting this work, and to the staff at NNWR, especially Rich King, Jon Olson, and Richard Urbanek, for sharing their insights. We thank the additional members of the expert panel for their contributions, insights, and patience: Marty Folk, John French, Chris Gullikson, Sammy King, Soch Lor, Mike Putnam, and Joel Trick. Jane Austin, Mark Burgman, Brendan Wintle, and an anonymous reviewer provided valuable suggestions that improved this manuscript. This work was supported by the USGS Patuxent Wildlife Research Center, the USGS Biological Resources Discipline Quick Response Research Program, USFWS Region 3, the Australian Centre of Excellence for Risk Analysis, and the Applied Environmental Decision Analysis hub at the University of Melbourne.

References

Berger, J.O., 1985. Statistical Decision Theory and Bayesian Analysis. Springer, New York.
Bowker, J.M., Stoll, J.R., 1988. Use of dichotomous choice nonmarket methods to value the whooping crane resource. American Journal of Agricultural Economics 70, 372–381.

Burgman, M.A., 2005. Risks and Decisions for Conservation and Environmental Management. Cambridge University Press, Cambridge, UK.

Converse, S.J., Royle, J.A., Urbanek, R.P., submitted for publication. Bayesian analysis of multi-state data with individual covariates for estimating genetic effects on demography. Journal of Ornithology.

Delbecq, A.L., Van de Ven, A.H., Gustafson, D.H., 1975. Group Techniques for Program Planning. Scott Foresman, Glenview, Illinois, USA.

Felli, J.C., Hazen, G.B., 1998. Sensitivity analysis and the expected value of perfect information. Medical Decision Making 18, 95–109.

Felli, J.C., Hazen, G.B., 1999. A Bayesian approach to sensitivity analysis. Health Economics 8, 263–268.

Goodwin, P., Wright, G., 2004. Decision Analysis for Management Judgment, third ed. John Wiley & Sons, West Sussex, UK.

Gustafson, D.H., Shukla, R.K., Delbecq, A., Walster, G.W., 1973. A comparative study of differences in subjective estimates made by individuals, interacting groups, Delphi groups, and nominal groups. Organizational Behavior and Human Preference 9, 280–291.

Hammond, J.S., Keeney, R.L., Raiffa, H., 1999. Smart Choices: A Practical Guide to Making Better Life Decisions. Broadway Books, New York, NY.

Kahneman, D., Slovic, P., Tversky, A., 1982. Judgment Under Uncertainty: Heuristics and Biases. Cambridge University Press, Cambridge, UK.

Kuhnert, P.M., Martin, T.G., Griffiths, S.P., 2010. A guide to eliciting and using expert knowledge in Bayesian ecological models. Ecology Letters 13, 900–914.

Link, W.A., Royle, J.A., Hatfield, J.S., 2003. Demographic analysis from summaries of an age-structured population. Biometrics 59, 778–785.

Linstone, H.A., Turoff, M. (Eds.), 1975. The Delphi Method: Techniques and Applications. Addison-Wesley, Reading, MA, USA.

Lock, A., 1987. Integrating group judgments in subjective forecasts. In: Wright, G., Ayton, P. (Eds.), Judgmental Forecasting. John Wiley & Sons, Chichester, UK, pp. 109–127.

Loomis, J.B., White, D.S., 1996. Economic benefits of rare and endangered species: summary and meta-analysis. Ecological Economics 18, 197–206.

MacMillan, D.C., Marshall, K., 2006. The Delphi process – an expert-based approach to ecological modelling in data-poor environments. Animal Conservation 9, 11–19.

McCarthy, M.A., Possingham, H.P., 2007. Active adaptive management for conservation. Conservation Biology 21, 956–963.

McDonald-Madden, E., Probert, W.J.M., Hauser, C.E., Runge, M.C., Possingham, H.P., Jones, M.E., Moore, J.L., Rout, T.M., Vesk, P.A., Wintle, B.A., 2010. Active adaptive conservation of threatened species in the face of uncertainty. Ecological Applications 20, 1476–1489.

Morgan, M.G., Henrion, M., 1990. Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. Cambridge University Press, Cambridge, UK.

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. Journal of Ornithology 148, S343–S349.

O'Hagan, A., Buck, C.E., Daneshkhah, A., Eiser, J.R., Garthwaite, P.H., Jenkinson, D.J., Oakley, J.E., Rakow, T., 2006. Uncertain Judgements: Eliciting Experts' Probabilities. John Wiley & Sons, West Sussex, UK.

Olsen, G.H., Carpenter, J.W., Langenberg, J.A., 1996. Medicine and surgery. In: D.H. Ellis, G.F. Gee, C.M. Mirande (Eds.), Cranes: Their Biology, Husbandry, and Conservation. US Department of Interior, National Biological Service, Washington, DC, USA and International Crane Foundation, Baraboo, Wisconsin, USA, pp. 137–174.

Raiffa, H., Schlaifer, R.O., 1961. Applied Statistical Decision Theory. Graduate School of Business Administration. Harvard University, Cambridge, MA, USA.

Regan, H.M., Colyvan, M., Burgman, M.A., 2002. A taxonomy and treatment of uncertainty for ecology and conservation biology. Ecological Applications 12, 618–628.

Rout, T.M., Hauser, C.E., Possingham, H.P., 2009. Optimal adaptive management for the translocation of a threatened species. Ecological Applications 19, 515–526.

Shea, K., Kelly, D., Sheppard, A.W., Woodburn, T.L., 2005. Context-dependent biological control of an invasive thistle. Ecology 86, 3174–3181.

Shea, K., Possingham, H.P., Murdoch, W.W., Roush, R., 2002. Active adaptive management in insect pest and weed control: intervention with a plan for learning. Ecological Applications 12, 927–936.

Speirs-Bridge, A., Fidler, F., McBride, M.F., Flander, L., Cumming, G., Burgman, M.A., 2010. Reducing overconfidence in the interval judgments of experts. Risk Analysis 30, 512–523.

Urbanek, R.P., Fondow, L.E.A., Zimorski, S.E., Wellington, M.A., Nipper, M.A., 2010. Winter release and management of reintroduced migratory Whooping Cranes *Grus americana*. Bird Conservation International 20, 43–54.

von Winterfeldt, D., Edwards, W., 1986. Decision Analysis and Behavioral Research. Cambridge University Press, Cambridge, UK.

Walters, C.J., 1986. Adaptive Management of Renewable Resources. Macmillan, New York, USA.

Williams, B.K., 1996. Adaptive optimization and the harvest of biological populations. Mathematical Biosciences 136, 1–20.

Wintle, B.A., Lindenmayer, D.B., 2008. Adaptive risk management for certifiably sustainable forestry. Forest Ecology and Management 256, 1311–1319.

Wintle, B.A., Runge, M.C., Bekessy, S.A., 2010. Allocating monitoring effort in the face of unknown unknowns. Ecology Letters 13, 1325–1337.

Wright, G., Ayton, P. (Eds.), 1987. Judgmental Forecasting. John Wiley & Sons, Chichester, UK.

Yokota, F., Thompson, K.M., 2004. Value of information literature analysis: a review of applications in health risk management. Medical Decision Making 24, 287–298.