



ELSEVIER

Contents lists available at ScienceDirect

Resource and Energy Economics

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



Monitoring as a partially observable decision problem



Paul L. Fackler^{a,*}, Robert G. Haight^b

^a Department of Agricultural and Resource Economics, North Carolina State University, Box 8109, Raleigh, NC 27695, USA

^b U.S. Forest Service Northern Research Station, 1992 Folwell Avenue, St. Paul, MN 55108, USA

ARTICLE INFO

Article history:

Received 5 April 2012

Received in revised form 20 December 2013

Accepted 27 December 2013

Available online 9 January 2014

JEL classification:

C61

C63

Q20

Q50

Keywords:

Environmental monitoring

Dynamic programming

Partial observability

Value of information

POMDP

ABSTRACT

Monitoring is an important and costly activity in resource management problems such as containing invasive species, protecting endangered species, preventing soil erosion, and regulating contracts for environmental services. Recent studies have viewed optimal monitoring as a Partially Observable Markov Decision Process (POMDP), which provides a framework for sequential decision making under stochastic resource dynamics and uncertainty about the resource state. We present an overview of the POMDP framework and its applications to resource monitoring. We discuss the concept of the information content provided by monitoring systems and illustrate how information content affects optimal monitoring strategies. Finally, we demonstrate how the timing of monitoring in relation to resource treatment and transition can have substantial effects on optimal monitoring strategies.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Resource managers undertake monitoring activities to estimate the state of the resource and learn about its dynamics. Because monitoring activities are costly, it is appropriate to integrate them into a larger resource management plan (Nichols and Williams, 2006; McDonald-Madden et al.,

* Corresponding author. Tel.: +1 1919 515 4535; fax: +1 1919 515 1824.

E-mail addresses: paul.fackler@ncsu.edu (P.L. Fackler), rhaight@fs.fed.us (R.G. Haight).

2010). For example, resource monitoring can be coupled with treatment in an adaptive management framework to gather information and reduce uncertainty about resource dynamics (Walters and Hilborn, 1976; Probert et al., 2011; Williams, 2011). Monitoring can also be used to reduce uncertainty about the state of the resource and improve the quality of treatments. This paper focuses on the problem of determining optimal monitoring strategies when the state of the resource is uncertain.

Researchers have used the Partially Observable Markov Decision Process (POMDP) framework to address resource management problems in which the resource state is uncertain. We review the POMDP framework and its applications to resource monitoring. For example, Haight and Polasky (2010) use the POMDP framework to determine a management strategy for controlling an invasive species with imperfect information about level of infestation. Other applications include monitoring strategies for erosion control (Tomberlin and Ish, 2007), environmental compliance (White, 2005), the presence of an invasive species (Regan et al., 2006), and the presence of an endangered species (Chadès et al., 2008).

A distinguishing feature of the POMDP framework is its recognition that, although managers undertake monitoring activities to discover the true state of the resource, those monitoring activities may not provide correct information. For example, a monitoring system may not detect the presence of an invasive species when in fact there is a moderate infestation (Haight and Polasky, 2010). The POMDP framework includes an observation model that predicts the probabilities of observing different resource states as a function of the actual resource state and the type of monitoring activity that is undertaken. The optimal resource treatment and monitoring activities over time depend on their associated observation models and costs.

The applications of the POMDP framework to resource monitoring have paid scant attention to the information content and timing of monitoring activities. We discuss the concept and measure of information content and show how the relative information content provided by alternative monitoring activities affects the optimal monitoring strategy. In addition, we address the issue of timing the monitoring activity in relation to the treatment decision and resource transition. We present a novel formulation in which monitoring is performed prior to treatment decisions, which in turn are conditional on the monitoring results.

2. Applications of POMDPs to environmental issues

The literature applying POMDPs to environmental issues is relatively small but growing. Lane (1989) appears to be the first paper to use the POMDP framework in a resource management problem which involves fishing decisions when the fish stock is not directly observed. The problem is to determine where to fish during the season given current beliefs about the level of the aggregate fish stock. Rather than viewing monitoring as a separate and costly activity, monitoring is assumed to occur in conjunction with the fishing activity. The model includes an observation function for each fishing location that specifies the probabilities of catch levels given the aggregate level of fish stock. The outcome of the fishing activity is used to update beliefs about the aggregate fish stock and decide where next to fish. Partial observability in fisheries management has also been addressed by making the decision directly dependent on the monitoring outcome rather than on belief states (Clark and Kirkwood, 1986; Moxnes, 2003; Sethi et al., 2005).

Studies that focus on monitoring as a separate and costly activity fall into three main application areas that address various land management issues and the management of endangered and invasive species. In land use applications, a site is classified into two or more categories but the current state of the site is not known. Monitoring can be undertaken to reduce uncertainty about the state of the site and treatment activities can also be initiated to alter the state of the site. White (2005) addresses the problem of choosing a monitoring system to support decisions concerning conservation activities on land sites when the current state of a site is not known with certainty. The site is classified according to which of two alternative vegetative covers dominates. Four different monitoring systems are considered, differing by their information content and cost. Tomberlin and Ish (2007) consider the problem of when to monitor or repair a logging road to reduce erosion when the degree of erosion is not known with certainty. Crowe and White (2007) consider the optimal use of potentially degraded

land that can either be in conservation or agricultural uses but the land use and degree of degradation are not known with certainty.

Studies that address endangered species management include [Chadès et al. \(2008\)](#), [Tomberlin \(2010\)](#), and [McDonald-Madden et al. \(2011\)](#). These studies model the state variable as the presence or absence of a target species. [Chadès et al. \(2008\)](#) consider when to switch a program to protect an endangered species from active management (e.g. anti-poaching activity) to a monitoring program aimed at determining if the species is extinct. The uncertain state is whether or not the species is still extant and the monitoring program seeks to determine if it is worth while to continue active management or if conservation resources should be shifted to other activities. [McDonald-Madden et al. \(2011\)](#) build on [Chadès et al. \(2008\)](#) by considering the management and monitoring of two sub-populations. [Tomberlin \(2010\)](#) considers the management of potential disturbance to an endangered species. The unknown state is whether the species is present or absent in a given period. If present it is beneficial to restrict human activities that might cause disturbance, such as nest or den abandonment. In all of these studies, active management and monitoring are treated as mutually exclusive activities.

Studies that focus on invasive species management include [Regan et al. \(2006, 2011\)](#) and [Haight and Polasky \(2010\)](#). These studies assume the state is one of a small number of categories describing the population of the species. [Regan et al. \(2006\)](#) examines the question of when monitoring should be stopped if an invasive species has not been sighted recently. Although they do not structure the problem as a POMDP, they do introduce a belief state and, as such, it could be posed as a POMDP. In their model, an invasive species is either present on a site or not and the site can be either surveyed or not. [Haight and Polasky \(2010\)](#) model the problem of monitoring and treating a site for an invasive species using a state variable that takes on one of three values: no, moderate or severe infestation. In their model, the manager can do nothing or can treat, monitor, or both. [Regan et al. \(2011\)](#) address the problem of selecting a management activity to eradicate an invasive plant given uncertainty about its presence. The plant population can be in one of three states: not present, dormant (seeds are present) or active (the site contains adult plants), and the management actions include two treatment options and doing nothing. This study differs from those above in that monitoring always occurs rather than being a choice variable.

An additional study that addresses invasive species is [Moore \(2008\)](#), which involves an invasive predator with the state variable being the predator population size. This study assumes that the unobserved state is a continuous variable rather than one of a small set of categories; the methodology required to address such problems is discussed later.

An important feature of the POMDP framework is the existence of an observation variable that provides information about the unobserved state. When both the observation and the state variables take on discrete sets of values, the relationship between the state and the observation variable can be described by a matrix that specifies the probability of obtaining each value of the observation variable given each value of the state variable. For the most part the studies described above interpret the observation variable as an indicator of the most likely state. This means that there are the same number of values of the state and observation variables and the diagonal elements of the matrix are the probabilities of a correct assessment.

Three types of observation matrices are found in the above studies. The first, which will be termed the simple form, assigns the same probability to the diagonal elements and allocates the remaining probability equally to the other observations. For example, suppose that there are three state values and that the probability of a correct assessment is c . The observation matrix is

$$\begin{bmatrix} c & (1-c)/2 & (1-c)/2 \\ (1-c)/2 & c & (1-c)/2 \\ (1-c)/2 & (1-c)/2 & c \end{bmatrix}$$

This type of observation matrix is used in [White \(2005\)](#) and [Tomberlin and Ish \(2007\)](#). Another special type of observation matrix can be called the detection type and applies to monitoring situations in which the goal is to determine presence or absence. A common assumption in such applications is

that there are no false positives (i.e. if detected then presence is certain). In this case the observation matrix is

$$\begin{bmatrix} 1 & 0 \\ 1-d & d \end{bmatrix}$$

where d is the detection probability (note that rows are associated with the true state, absent or present, and columns with the observation variable, absent or present). Detection type observation matrices are used in [Regan et al. \(2006\)](#), [Chadès et al. \(2008\)](#), [Tomberlin \(2010\)](#), and [McDonald-Madden et al. \(2011\)](#) (in the last case there are implicitly two observation matrices, one for each subpopulation). [Regan et al. \(2011\)](#) also use a detection type matrix but one for which detection can only occur in the adult plant stage and not in the not-present or dormant stages. This is the one case in which the observation variable had a different number of categories than the state.

General square observation matrices are used in [Haight and Polasky \(2010\)](#) and in [Crowe and White \(2007\)](#). With the previous types of observation matrices it is intuitively reasonable to conjecture that increases in c or d can be interpreted as increasing the informational content of the observation variable (this is verified below). [Haight and Polasky \(2010\)](#) conjectured that increasing the diagonal elements of a general square observation matrix increased its informativeness. It is shown below that this is not correct.

Another important consideration for monitoring systems is the timing of the monitoring program in relationship to when decisions concerning it and other control actions are made and when these activities are actually carried out. With the exception of [Moore \(2008\)](#), little attention has been given to this topic. The most common timing alternative can be represented as: *state transition, monitor, update beliefs, treatment/monitoring decision, treat* (time periods are assumed to always begin with the state transition). This sequence corresponds to the way most POMDP software defines the decision problem. Note that monitoring and treatment decisions are made at the same time but monitoring is conducted after the effects of the treatment and next state transition have occurred. We will examine this sequence of events and some alternatives later in the paper. In many of the studies, monitoring and other control actions are mutually exclusive ([Tomberlin and Ish, 2007](#); [Chadès et al., 2008](#); [McDonald-Madden et al., 2011](#); [Tomberlin, 2010](#)). Such an assumption might be made because of implicit budget limitations but a number of the studies mentioned do not explicitly consider monitoring and treatment costs. [White \(2005\)](#), [Crowe and White \(2007\)](#) and [Haight and Polasky \(2010\)](#), on the other hand, allow the decision maker to both treat and monitor at the same time although this option may not be used. In [Regan et al. \(2006\)](#) and [Regan et al. \(2011\)](#), the issue does not arise; in the former there is no treatment action and in the latter monitoring always occurs.

To date, no studies have considered the possibility that treatment and monitoring decisions could be made at different times, with the results of monitoring used to inform treatment decisions. Given that the standard framework for specifying dynamic optimization problems involves a sequence of identically structured decisions, it is not surprising that no studies have examined this possibility. The software used for the current analysis allows for this possibility, which is explored more fully later in the paper.

3. The POMDP framework

As generally used, POMDP models are extensions of finite, discrete state and discrete time Markov Decision Processes (MDPs) in which the state variables are not known with certainty. Recognizing that state variables are uncertain makes the POMDP more difficult to specify and solve than a MDP. Unlike MDPs, we can no longer base our action on the resource state because we do not know exactly what it is. Further, a monitoring system can be used to gather information about the resource state and reduce uncertainty. It generally is not optimal to base treatment activities on observations from the monitoring system alone because those observations may provide an imprecise signal. A rigorous way to proceed is to characterize the problem in terms of the degree of belief that each resource state is, in fact, the current state. In other words, we need to assign probabilities to each resource state and base our management action on these probabilities. We update our beliefs about the resource

state based on our understanding of the state transition probabilities, the observation probabilities associated with the monitoring system, and the observations themselves.

The POMDP framework is well-developed in the artificial intelligence literature, where the main application has been the control of robots. We use the notation and language of Kaelbling et al. (1998), who provide a thorough review of work on POMDPs from the operations research literature and describe its connection to closely related work in artificial intelligence. Let $S = \{s_1, s_2, \dots\}$ represent a finite set of resource states and $A = \{a_1, a_2, \dots\}$ represent a finite set of possible actions. The decision-maker may not know for sure the current resource state. Instead the decision-maker has a set of beliefs represented by a probability distribution b over the set of states S where $b(s)$ represents the probability of being in state s . The axioms of probability require that $0 \leq b(s) \leq 1$ for all $s \in S$ and $\sum_{s \in S} b(s) = 1$. Each period, the decision-maker chooses an action, a , from the set A based on the current beliefs. Following the action, the state of the resource may change. We define a state-transition function, T , giving for each state and action a probability distribution of the possible states in the next period, where $T(s, a, s') = Pr\{s'|s, a\}$ is the conditional probability of moving from state s to state s' after taking action a .

After the action and state transition, the manager makes an observation to gather information about the resource state. We define a finite set of observations, $\Omega = \{o_1, o_2, \dots\}$, that represents the possible observations (signals) the decision maker may receive. Further, we define an observation function, O , which gives, for each action and subsequent state, a probability distribution over possible observations. We write $O(s', a, o) = Pr\{o|s', a\}$ as the conditional probability of making observation o given that the decision-maker took action a and landed in state s' .¹

Note that the observation function O gives conditional probabilities of observations as a function of the action taken. Each action includes an assumption about the degree of monitoring and its associated observation matrix. In our examples below, we show how the information content of the observation function associated with different degrees of monitoring may affect optimal management strategies. Note also that the observation function assumes that observations depend on the future state of the system after action and state transition. Although this definition is commonly used in applications of POMDP in robotics and other areas of engineering, one can define an alternative observation function in which observations depend on the current resource state prior to the action and state transition. In our examples below, we show how the conditioning of the observation affects results.

Finally, note that the definition and incorporation of an observation function in the POMDP framework distinguishes it from adaptive management frameworks in which the resource state is assumed to be known with certainty while monitoring and treatments are used to reduce uncertainty about resource dynamics (see Williams, 2011; Probert et al., 2011).

Bayes rule is used to compute a new set of beliefs, b' , given an old set of beliefs b , an action, a , and an observation, o . The new belief concerning resource state s' , $b'(s')$, can be computed with basic probability theory:

$$b'(s')|b, a, o = \frac{O(s', a, o) \sum_{s \in S} T(s, a, s') b(s)}{\sum_{s' \in S} O(s', a, o) \sum_{s \in S} T(s, a, s') b(s)}. \quad (1)$$

The numerator is the product of two probabilities: the conditional probability of receiving observation o given that the decision maker took action a and landed in state s' and the probability of landing in state s' given the old set of beliefs b and action a . The denominator is the probability of making observation o overall possible resource states and serves as a normalizing factor that causes $\sum_{s' \in S} b'(s') = 1$.

The solution to a POMDP is an optimal management strategy that maximizes expected total reward. Let R be the per-period reward function, where $R(s, a)$ gives the decision-maker's immediate reward for taking action a in resource state s . The objective is to choose a sequence of actions that maximizes the expected total reward. For infinite-horizon POMDPs, the sequence of actions has infinite length and we specify a discount factor $\gamma \in [0, 1)$ so that the total reward is finite and the problem is well

¹ We have followed the standard notation used in the POMDP literature but it should be noted that this notation is potentially misleading because the observation variable is conditional on and generally assumed to occur after the state transition. It would be clearer, therefore, to indicate this using o' rather than o .

defined. In this case, the expected total reward is given by $E \left[\sum_{t=0}^{\infty} \gamma^t R(s_t, a_t) \right]$ where s_t and a_t denote the resource state and decision-maker's action in period t .

The decision-maker's sequence of actions is characterized by a strategy (policy in the language of POMDP), π , which maps the decision-maker's beliefs, represented by the probability distribution b , to a prescribed action a . Note that the information about the system from the observation variable is incorporated into the beliefs and hence the strategy does not directly depend on the observations. When the objective is to maximize the expected discounted value of rewards over an infinite horizon, the optimal policy is stationary (Howard, 1960): the choice of action depends only on the decision-maker's beliefs, b , and is independent of the time period.

POMDP policies are often computed using a value function over the belief space. The value function $V_{\pi}(b)$ for a given policy, π , is defined as the total expected reward the decision-maker receives starting with belief state b and executing the policy π over the infinite horizon. The optimal POMDP policy maximizes the value function, which satisfies the Bellman equation:

$$V^*(b) = \max_{a \in A} \left[\sum_{s \in S} b(s)R(s, a) + \gamma \sum_{s \in S} \sum_{s' \in S} \sum_{o \in \Omega} b(s)T(s, a, s')O(s', a, o)V^*([b'(s'_1), b'(s'_2), \dots]) \right] \quad (2)$$

The first term inside the brackets of Eq. (2) is the expected reward associated with taking action a in the current period given the current belief probabilities b . The second term inside the brackets of Eq. (2) is the expected total discounted reward over subsequent periods after taking action a in the current period. Expected total discounted reward is the weighted sum of the optimal values associated with updated beliefs where each weight, $b(s)T(s, a, s')O(s', a, o)$, represents the probability of beginning in state s , moving to state s' , and making observation o , after taking action a . Note that the updated beliefs $[b'(s'_1), b'(s'_2), \dots]$ are computed using the updating rule given in Eq. (1).

Thus far the discussion of POMDPs has focused on models with a discrete set of states. When one or more state variables are defined on a continuum, the belief distribution must be expressed as a density function. For such problems to be computationally feasible, a parameterized family of density functions must be used, with the parameters of that family acting as state variables. This is the approach taken by Moore (2008) in a predator control application in which the unobserved state was the size of the predator population. Often with such models the updated belief distribution for next period's state does not have the same functional form as the belief distribution for the current state. Moore (2008) addressed this by finding the parameters of the assumed distribution that most closely approximates the actual updated belief distribution.

4. Solution algorithms

Exact solutions for finite time problems can be obtained using the fact that the value function can be expressed in terms of a finite number of linear functions (Sondik (1971)). Although numerical procedures based on linear programming have been developed to find exact solutions to POMDP problems (see Monahan, 1982; Cassandra, 1994 for reviews), these procedures have severe curse-of-dimensionality issues in that the number of linear functions can grow very quickly as the number of time periods and the number of states grows. This approach is therefore not a realistic option except for very small problems.

In the last decade considerable progress has been made in obtaining approximate solutions to POMDPs. Among the approaches used are factoring the problem so it can be divided into smaller subproblems and using various search techniques to limit the scope of the problem to areas of the belief space that are likely to be visited. Examples of the literature in this area include Guestrin et al. (2003), Smith and Simmons (2005), Spaan and Vlassis (2005), Pineau et al. (2006), Shani et al. (2007), Kurniawati et al. (2008) and Ong et al. (2009).

Another alternative, used here, is to discretize the belief space. Lovejoy (1991) discusses how this can be done using a regular grid on an $(n - 1)$ -dimensional simplex, where n is the number of alternative values of the state variable. He suggests using Freudenthal interpolation to assign probability values to grid points when Eq. (1) results in a non-grid value. It is also possible to discretize using a scattered set of points in the belief space and even to pick these points adaptively so one obtains a

greater concentration in an area of that space in which the system spends more time. All of the resource monitoring studies discussed above have been replicated using Lovejoy's discretization approach. The solution algorithm is incorporated in the MATLAB based MDPsolve package for solving dynamic optimization problems which is available for download at <https://sites.google.com/site/mdpsolve/>. The package also has a procedure implementing the exact solution based on Monahan's (1982) method.

5. The information content of observations

An important consideration in monitoring problems is the choice of monitoring systems. It is often the case that some information about the state of the system is obtained even without an explicit (and often costly) monitoring program. Information can be obtained, for example, by reported sightings of a rare species or from informal reports of land owners, recreational users, or others concerning the extent of an infestation. Although these reports may not be as informative as a systematic monitoring program, they may nonetheless be useful for updating beliefs about the state of the system and making treatment decisions without incurring the extra cost of a formal monitoring system. In addition, there may be a menu of explicit monitoring systems available with different degrees of informativeness, as in White (2005).

The possibility of receiving more than one signal raises the question of when one monitoring system is more informative than another. This was first addressed in a rigorous way by Blackwell (1951) and subsequently by, among others, Marschak and Miyasawa (1968) and Hirshleifer and Riley (1992). In this literature, two signals are compared, both of which contain information about the underlying state S . Signal X is defined by its conditional probability matrix $O_x = \Pr(X|S)$; the pair Y and O_y is defined similarly. Blackwell demonstrated that a signal X is at least as informative as another signal Y if there is a probability matrix A (a non-negative matrix with rows summing to one) that satisfies $O_x A = O_y$. For example, a perfect signal is at least as informative as an arbitrary signal. To see this, note that $O_x = I$ implies that X is a perfect signal. Setting $A = O_y$ ensures that A is a probability matrix and implies that X is at least as informative as Y . Similarly an arbitrary signal is at least as informative as a perfectly non-informative signal. To see this, note that $O_y = 1w^T$ for some probability vector w (i.e. a matrix with identical rows) implies that Y is a completely non-informative signal because conditioning on S does not affect the signal probability (so $\Pr(Y|S) = \Pr(Y)$). For any O_x set $A = O_y$, thereby demonstrating that X is at least as informative as Y .

The essence of Blackwell's proof is to show that $O_x A = O_y$ implies that $\Pr(S|X, Y) = \Pr(S|X)$, i.e. that Y adds nothing to our knowledge of S that is not already contained in X . This is equivalent to showing that $\Pr(Y|X, S) = \Pr(Y|X)$, i.e. that knowledge of Y adds nothing to our knowledge of S that is not already contained in X . The equivalence relies on the fact that

$$\frac{\Pr(Y|X, S)}{\Pr(Y|X)} \equiv \frac{\Pr(S|X, Y)}{\Pr(S|X)} \equiv \frac{\Pr(S, Y|X)\Pr(Y|X)}{\Pr(S|X)}$$

This expression equals 1 when X is as informative as Y (some technical issues arise when O_x is not full rank and so X is not informative about one or more sets of states; see Marschak and Miyasawa, 1968 for details).

In their POMDP model of invasive species management, Haight and Polasky (2010) suggest that a signal is more informative if it moves probability mass horizontally from the off-diagonal elements to the diagonal ones. It is easy to show, even in the two state/two signal case, that this is neither necessary nor sufficient for a signal to be considered more informative. Consider

$$O_x = \begin{bmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{bmatrix}$$

$$O_y = \begin{bmatrix} 3/8 & 5/8 \\ 2/3 & 1/3 \end{bmatrix}$$

In this case

$$A = \begin{bmatrix} -1/2 & 3/2 \\ 5/4 & -1/4 \end{bmatrix}$$

contains negative elements and hence it cannot be concluded that signal X is more informative than signal Y even though the diagonal elements of O_x are larger than those of O_y . Alternatively, consider the same O_x and

$$O_y = \begin{bmatrix} 7/12 & 5/12 \\ 1/2 & 1/2 \end{bmatrix}$$

In this case

$$A = \begin{bmatrix} 5/6 & 1/6 \\ 1/3 & 2/3 \end{bmatrix}$$

contains all positive elements. Thus signal X is more informative than signal Y even though O_y contains a diagonal element that is larger than the associated element in O_x .

Blackwell's approach can be used to illustrate a number of intuitively reasonable propositions. For example, for detection type systems with $d_1 < d_2$, we have

$$\begin{bmatrix} 1 & 0 \\ 1 - d_2 & d_2 \end{bmatrix} A = \begin{bmatrix} 1 & 0 \\ 1 - d_1 & d_1 \end{bmatrix}$$

implying that

$$A = \begin{bmatrix} 1 & 0 \\ 1 - d_1/d_2 & d_1/d_2 \end{bmatrix}$$

All of the elements of A are non-negative, implying that higher values of d are more informative. Extreme cases are $d = 1$ (perfect information) and $d = 0$ (no information).

For simple type monitoring systems that have equal diagonal elements and equal non-diagonal elements, there is a single probability, c , of a correct assessment and the remaining probability is evenly distributed over the other states. With n possible states, the observation matrix can be written as the sum of a diagonal matrix and a constant matrix:

$$O = \frac{nc - 1}{n - 1} I_n + \frac{1 - c}{n - 1} 1_n 1_n^T$$

A perfectly informative system is one with $c = 1$ and a non-informative system is one with $c = 1/n$. Suppose that $1/n \leq c_1 < c_2 \leq 1$. It is easy to verify that

$$A = O_2^{-1} O_1 = \frac{nc_1 - 1}{nc_2 - 1} I_n + \frac{c_2 - c_1}{nc_2 - 1} 1_n 1_n^T$$

and that A has all positive elements with rows summing to 1, thereby demonstrating that monitoring system 2 (with higher value c) is at least as informative as system 1 (with lower value c). To illustrate this suppose that $n = 3$, $c_1 = 4/6$ and $c_2 = 5/6$, then

$$A = \begin{bmatrix} 7/9 & 1/9 & 1/9 \\ 1/9 & 7/9 & 1/9 \\ 1/9 & 1/9 & 7/9 \end{bmatrix}$$

Using Blackwell's approach it is also possible to show that the order of the signal is irrelevant to the informativeness of the monitoring system and hence the diagonal elements have no particular significance. To see this consider the comparison of the observation matrix O with OP , where P is a permutation matrix that reorders the columns of O . P consists of all zeros except a single 1 in each row and column and has the property that $P^T P = I$. Using Blackwell's theorem $OPA = O$ implies that $A = P^T$

and $OA = OP$ implies that $A = P$. Since the elements of P are all non-negative this implies that O and OP are equally informative. For example

$$O = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is exactly as informative as $O = I$.

As has already been noted most of the studies discussed interpret the signal as an assessment of the system state and hence the number of values the signal may take is equal to the number of values that the state may take (i.e. the observation function O is a square matrix). This need not be the case. For example a signal defined by

$$O = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

provides perfect information that the state is either 1 (if the signal equals 1) or that the state is either 2 or 3 (if the signal equals 2).

6. Control of invasive species

To examine alternative monitoring systems and timing issues we use the invasive species management study of [Haight and Polasky \(2010\)](#). In this study it is assumed that an ecosystem is in one of three possible states: no infestation, moderate infestation, or high infestation. There are four mutually exclusive actions: no action (NN), monitoring only (NM), treatment only (TN), and both monitoring and treatment (TM). Based on beliefs about the level of infestation, the decision-maker chooses actions to minimize the sum of discounted costs of management and damage caused by the infestation.

The following parameter values are used in their base model. The state transition matrix depends on whether or not treatment occurs. The transition matrices with no treatment and with treatment are, respectively:

$$T_n = \begin{bmatrix} 0.8 & 0.2 & 0 \\ 0 & 0.8 & 0.2 \\ 0 & 0 & 1 \end{bmatrix}, \quad T_t = \begin{bmatrix} 0.9 & 0.1 & 0 \\ 0.8 & 0.2 & 0 \\ 0.6 & 0.4 & 0 \end{bmatrix}$$

where rows represent current states and columns represent future states. For example, once a high infestation occurs it cannot be reduced without treatment ($T_n(3,3) = 1$). Furthermore, if treated, a high infestation is guaranteed to be reduced ($T_t(3,3) = 0$). It is assumed that treatment has no effect on the information used in updating beliefs (an assumption that is relaxed later) and hence there are two observation matrices for this model. The observation matrix for the actions without explicit monitoring (NN or TN) returns an imperfect signal:

$$O_a = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.3 & 0.4 & 0.3 \\ 0.1 & 0.4 & 0.5 \end{bmatrix}$$

Notice that this is an informative signal and, as already discussed, it is more correct to think of this an implicit monitoring system. The observation matrix for the (costly and explicit) monitoring options (NM or TM) returns a perfect signal (i.e. $O_m = I_3$). The objective is to minimize the combined damage, monitoring and treatment costs over an infinite time horizon with a discount factor $\gamma = 0.95$. The per period damage costs are 0, 10 and 20 for no infestation, moderate infestation, and high infestation, respectively. The monitoring cost is 4 units and the treatment cost is 20 units.

The optimal solution for the base case is shown in [Fig. 1](#) (post-transition monitoring), which displays the optimal action as a function of the current belief vector b . The figures show ternary plots in

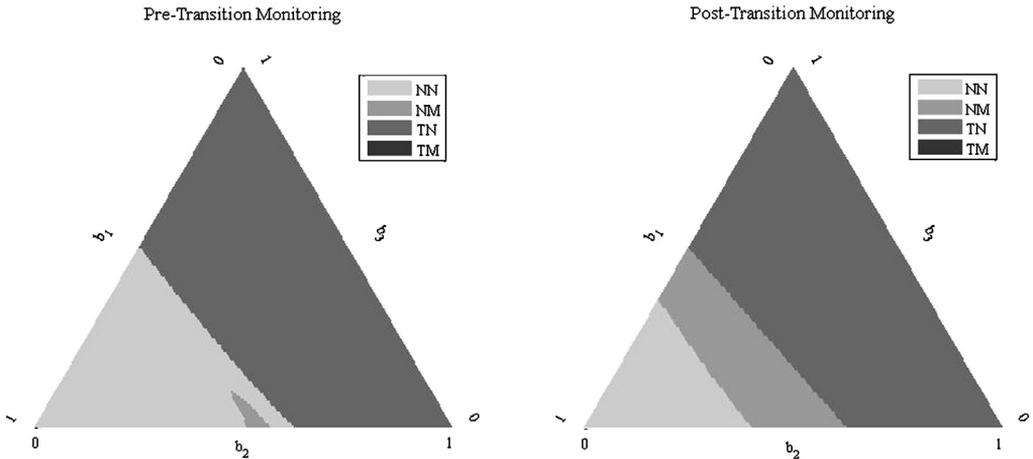


Fig. 1. Optimal treatment and monitoring decisions as a function of the belief state for a POMDP where the treatment and monitoring decision is made simultaneously and observations are made either before the state transition or after the state transition. NN: no-treatment/no-monitoring; NM: no-treatment/monitoring; TN: treatment/no-monitoring; TM: treatment/monitoring. Note: the treatment/monitoring option (TM) is never optimal in either case.

which each corner represents certainty in one of the three alternative states with the lower left corner representing certainty in state 1, the lower right in state 2 and the upper corner in state 3. Certainty in a given alternative diminishes as one moves toward the boundary opposite the certainty corner (diagonally to the right for state 1, diagonally left for state 2 and down for state 3).

With the base parameter values, it is never optimal to both treat and monitor. Since treating is guaranteed to reduce a high infestation, even without monitoring beliefs about the state in the next period are sufficiently accurate that monitoring is not cost effective. Treatment is conducted when it is fairly certain that the infestation is either moderate or high (northeast area of figure) and the do-nothing action is optimal when it is fairly certain that the infestation is low (lower left corner). Explicit monitoring does take place in the middle region where there is a relatively even split between low infestation on the one hand and moderate or high infestation on the other.

As previously mentioned, POMDP applications typically assume that the monitoring signal is conditioned on the post transition state, even though the decision to monitor is made prior to the treatment and state transition. Suppose instead that monitoring occurs immediately after a decision to monitor is made and prior to the treatment: *state transition, update beliefs, treatment and monitoring decision, monitor, treat*. In this case, the signal obtained from monitoring provides information about the current state and the observation function, O , gives, for each current state and action, a probability distribution over possible observations. We write $O(s, a, o) = pr(o|s, a)$ as the conditional probability of making observation o given that the current state is s and the decision-maker took action a . In this case, the belief updating rule given in Eq. (1) must be modified to:

$$b'(s')|b, a, o = \frac{\sum_{s \in S} O(s, a, o)T(s, a, s')b(s)}{\sum_{s' \in S} \sum_{s \in S} O(s, a, o)T(s, a, s')b(s)}$$

The Bellman equation given in Eq. (2) would also be modified by replacing $O(s', a, o)$ with $O(s, a, o)$.

The optimal decision rule for Haight and Polasky’s base case model using this alternative is also displayed in Fig. 1 (pre-transition monitoring). Making the pre-transition monitoring assumption results in far less explicit monitoring than under the post-transition monitoring assumption. Given that monitoring does not influence the current treatment, information about the post-transition state is more valuable than information about the current state. With a perfect monitoring system, post-transition monitoring assures that the state is known when next period’s decision is made. With

Table 1

Cost function comparisons for alternative models at selected belief values.

Model comparisons	Known low state [1 0 0]	Known medium state [0 1 0]	Known high state [0 0 1]	Equally likely states [1/3 1/3 1/3]
(1) Pre- versus post-transition monitoring				
Pre-transition monitoring	162.40	192.40	210.37	193.75
Post-transition monitoring ^a	147.89	177.89	195.16	179.12
(2) Costless monitoring ^b				
Implicit and imperfect monitoring	141.38	171.38	188.17	172.54
Imperfect monitoring only	142.17	172.17	188.79	173.28
(3) Implicit monitoring ^c				
No information: (O_n, O_n)	177.66	207.66	225.22	208.98
Treatment information: (O_n, O_a)	174.10	203.23	220.34	204.50
Public/treatment information: (O_p, O_a)	169.33	198.57	215.66	199.81
(4) Implicit and explicit monitoring ^d				
No information: (O_n, O_m, O_n, O_m)	162.61	192.61	209.67	193.94
Treatment information: (O_n, O_m, O_a, O_m)	160.39	189.69	206.43	190.89
Public/treatment information: (O_p, O_m, O_a, O_m)	160.00	189.40	206.12	190.59
(5) Two-stage models ^e				
Public information in stage 1: (O_p, O_m), (O_n, O_a)	158.59	187.98	204.67	187.72
Public information in stage 2: (O_n, O_m), (O_p, O_a)	157.56	186.97	203.74	186.73

^a This case uses the base case model parameters from Haight and Polasky (2010).

^b Results in these two cases are not directly comparable to the other cases in this table because these cases involve costless monitoring.

^c Observation matrices associated with management options: no-treatment, treatment.

^d Observation matrices associated with management options: no-treatment/no-monitoring, no-treatment/monitoring, treatment/no-monitoring, treatment/monitoring.

^e Observation matrices associated with management options in stage 1 (no-monitoring, monitoring) and stage 2 (no-treatment, treatment).

pre-transition monitoring, the belief distribution next period will reflect the noisiness of the transition probabilities. Therefore, when monitoring is informative about the current state it is less likely to be undertaken. Indeed, the area in which explicit monitoring is optimal is quite small in the pre-treatment/transition monitoring case. It is only when there is about a 50/50 split between no and moderate infestation that explicit monitoring is undertaken. In this situation, explicit monitoring at least resolves the uncertainty about whether there will be a transition to severe infestation in the next period because the transition probabilities do not allow a transition from no infestation to severe infestation. However, monitoring does not resolve the uncertainty about a transition to low or moderate infestations.

Table 1 Part 1 displays value function comparisons for four alternative current belief states. These include the three corners, which represent current certainty about the state and the equally weighted (perfectly uninformed) state. In each case the use of pre-transition monitoring results in anywhere from a 7% to a 10% increase (depending on the current belief state) in the sum of the discounted expected costs. Costs increase because pre-transition monitoring does not resolve the uncertainty of the state transition prior to next period's decision. Haight and Polasky (2010) also explore the use of a costless but imperfect monitoring system described by the observation matrix: $\hat{O}_m = \begin{bmatrix} 0.75 & 0.25 & 0 \\ 0.15 & 0.7 & 0.15 \\ 0.05 & 0.2 & 0.75 \end{bmatrix}$ Thus, there are two alternative costless monitoring systems:

“implicit monitoring” with observation matrix O_a (defined above) and “imperfect monitoring” with observation matrix \hat{O}_m .

Observation matrix O_a is used in two management options: no treatment/implicit monitoring (NN) and treatment/implicit monitoring (TN), and observation matrix \hat{O}_m is used in two management options: no treatment/imperfect monitoring (NM) and treatment/imperfect monitoring (TM).

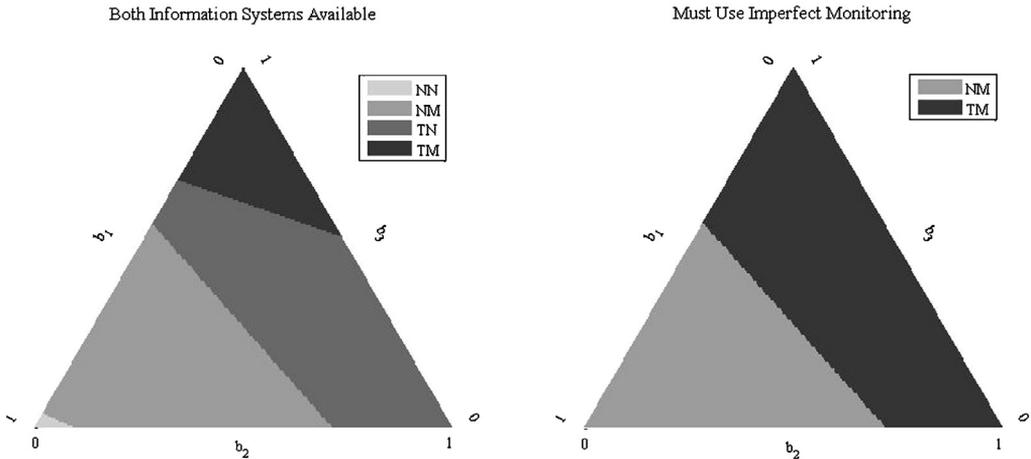


Fig. 2. Optimal monitoring and treatment decisions as a function of the belief state for a POMDP with two costless information systems (no explicit monitoring and imperfect monitoring), each providing an imperfect signal, and a POMDP in which the imperfect monitoring option must be used. NN: no treatment, no explicit monitoring; NM: no treatment, imperfect monitoring; TN: treatment, no explicit monitoring; TM: treatment, imperfect monitoring.

Haight and Polasky’s (2010) results (reproduced in Fig. 2, imperfect monitoring only) are based on the assumption that \hat{O}_m is more informative than O_a because its diagonal elements are larger. If this were true, management options with the imperfect monitoring system (NM and TM) would always be used and the decision would be simply to treat or not treat. Applying the Blackwell criteria by computing

$$A = \hat{O}_m^{-1}O_a = \begin{bmatrix} 0.565 & 0.543 & -0.109 \\ 0.304 & 0.307 & 0.326 \\ 0.014 & 0.399 & 0.587 \end{bmatrix}$$

however, demonstrates that \hat{O}_m is not more informative than O_a .

Haight and Polasky (2010) incorrectly assumed that if both monitoring systems are costless, the \hat{O}_m signal would always be used to the exclusion of the O_a signal. This assumption led them to incorrect results in their Figs. 7 and 8. In their Fig. 7 (reproduced in Fig. 2, imperfect monitoring only), the no treatment/imperfect monitoring (NM) and the treatment/imperfect monitoring (TM) options are the only actions that appear as optimal. We correct this result in our Fig. 2 (implicit and imperfect monitoring), displaying areas of the belief space where actions with the implicit monitoring signal O_a (NN and TN) are optimal. Table 1, Part 2 displays value function comparisons for four alternative belief states. In each case, the ability to choose implicit or imperfect monitoring results in about a 1% reduction in the discounted expected cost of management and damage caused by the infestation relative to the case where management options only involve imperfect monitoring.

An important consideration in determining whether explicit monitoring should be used concerns the information that would be available without such a system. With an invasive species, implicit monitoring might involve obtaining information about the state of the infestation from public reporting of sightings or from informal observations by management staff. In addition, there might be additional information obtained from treatment activities, especially where these activities involve on-site applications of treatments or where the treatment itself requires some level of monitoring.

To explore the value of implicit monitoring, we apply Haight and Polasky’s (2010) model with three different implicit monitoring systems. Implicit monitoring with observation matrix O_a is interpreted as the information obtained from treating the resource. Public monitoring with observation matrix O_p provides information without treatment and is strictly less informative than the implicit

monitoring with treatment (O_a). The observation matrix for the public monitoring system is obtained by multiplying O_a by a matrix with constant diagonals and constant off-diagonals

$$O_p = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.3 & 0.4 & 0.3 \\ 0.1 & 0.4 & 0.5 \end{bmatrix} \begin{bmatrix} 0.65 & 0.175 & 0.175 \\ 0.175 & 0.65 & 0.175 \\ 0.175 & 0.175 & 0.65 \end{bmatrix} = \begin{bmatrix} 0.4125 & 0.4125 & 0.1750 \\ 0.3175 & 0.3650 & 0.3175 \\ 0.2225 & 0.3650 & 0.4125 \end{bmatrix}$$

Thus, the public monitoring system agrees with the information obtained from treatment 65% of the time. The “no information” monitoring system with observation matrix O_n represents a situation in which no information is obtained (the observation matrix has identical rows).

To estimate the value of alternative implicit monitoring systems, we first compute value functions using [Haight and Polasky's \(2010\)](#) model for three different sets of two management options: not treating or treating ([Table 1](#), Part 3). Each set employs different implicit monitoring systems without any explicit monitoring. In the first set, the no-information observation matrix O_n used for both the no-treat and treatment options. In the second set, the no-information observation matrix O_n used for the no-treat option and observation matrix O_a is used with the treatment option. In the third set, the observation matrix for public monitoring O_p is used with the no-treat option and observation matrix O_a is used with the treatment option. Comparing the situation in which no information is obtained (set 1) to the situation where signal O_a is obtained if treatment occurs (set 2), there is a reduction in cost of approximately 2%. If both public monitoring in the absence of treatment and monitoring with treatment are available (set 3), then an additional 2% reduction in cost is obtained.

When explicit monitoring using observation matrix O_m is included as an option in [Haight and Polasky's \(2010\)](#) model, the costs are substantially reduced. To show this, we examine three different sets of four management options: no-treatment/no-monitoring, no-treatment/monitoring, treatment/no-monitoring, and treatment/monitoring ([Table 1](#), Part 4). In the first set, the no-information observation matrix O_n is used in management options with no monitoring and observation matrix O_m is used in options with monitoring. In the second set, the no-information observation matrix O_n used for the no-treatment/no-monitoring option, observation matrix O_a is used with the treatment/no-monitoring option, and observation matrix O_m is used in options with monitoring. In the third set, the observation matrix for public monitoring O_p is used with no-treatment/no-monitoring option. When explicit monitoring provides the only information (first line in [Table 1](#), Part 4) there is a 7–9% cost reduction compared to the situation where no information is obtained (first line of [Table 1](#), Part 3). The addition of monitoring (O_m) information to treatment information (O_a) provides a 7–9% reduction in cost (lines 2 of [Table 1](#), Parts 3 and 4). Even when public (O_p) and treatment (O_a) information is available, the addition of monitoring (O_m) still reduces costs by 4–5% (lines 3 of [Table 1](#), Parts 3 and 4). These experiments demonstrate the sensitivity of costs to the precise definitions of the available information systems.

In the standard POMDP model, with the signal conditional on the post-transition state, the manager nonetheless makes a simultaneous choice about both treatment and monitoring based only on the belief state at the beginning of the current period. An alternative approach is to separate the timing of the treatment and monitoring decisions during the period, with treatment decisions made after updating the belief state mid period with the monitoring information received earlier in the period. Although it cannot be accommodated in a standard POMDP framework, the MDPSolve software used here allows for problems with multiple sub-periods.

In the first stage, the decision is to not monitor or to monitor. It is assumed that the state does not change from the first stage to the second so the state transition for either action is an identity matrix. The signal received is either from the fully informative system $O_m = I_3$ (if monitoring is chosen) or is from either the no information system (O_n) or the public information system (O_p). The cost in this stage is either 0 if implicit monitoring (O_n or O_p) is chosen or c_m if explicit monitoring (O_m) is conducted. The discount factor is set to 1. In the second stage the decision is to not treat or to treat. The state transition equation is either T_n or T_t (defined above for [Haight and Polasky's \(2010\)](#) model) and the signal received after treatment is from the observation matrix O_a (if treatment is chosen) or is from either the public information (O_p) or the no information (O_n) observation matrices

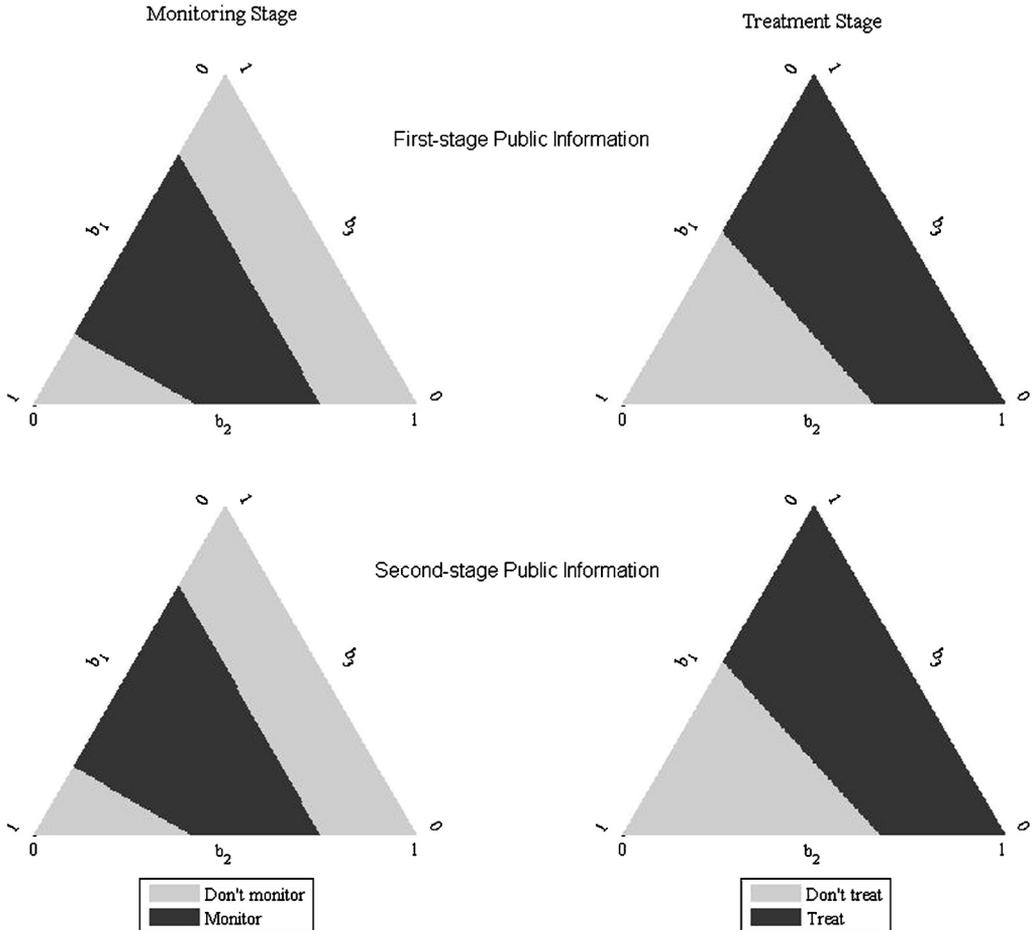


Fig. 3. Optimal monitoring and treatment decisions as a function of the belief state for a POMDP where the monitoring decision is followed by the treatment decision within each period. *Note:* If explicit monitoring has occurred in the first stage the belief state in the second stage will be at one of the three corners as the monitoring system used in this example provides perfect information. If there is no monitoring in the first stage the belief weights in top figure (implicit monitoring) will nonetheless change whereas they will not in the bottom figure (no information without explicit monitoring).

(these are distinguished by reference to the public information in stage 1 or stage 2). The cost is the damage cost plus the treatment cost (if the treatment action is chosen) and the discount factor is γ .

The optimal strategy for this problem is displayed in Fig. 3. No monitoring is undertaken when certainty concerning the current state of the infestation is relatively high, roughly when $Pr(High) + Pr(Moderate)/2 < 0.25$ (in which case no treatment is used) or when $Pr(No\ infestation) < 0.25$ (in which case treatment is used). If uncertainty is relatively high, on the other hand, monitoring is conducted. The decision rules shown in Fig. 3 are potentially misleading in this situation because the monitoring will move the belief state to one of the corners, with no treatment undertaken if the beliefs move to corner [100] and with treatment undertaken if the beliefs move to one of the other corners. Table 1, Part 5 displays value function comparisons for the four alternative current belief states. In each case the ability to monitor before making a treatment decision results 1–2% decrease in the cost function compared with the model results shown in line 3 of Part 4.

7. Discussion

This paper has attempted to summarize and clarify the growing literature that examines questions relating to the conditions under which environmental monitoring should be undertaken. Monitoring is undertaken to reduce uncertainty about variables that measure environmental conditions and influence the results of control activities. When there is little uncertainty about environmental conditions it is not optimal to undertake costly monitoring programs. The POMDP framework provides a way to explicitly incorporate the value of a monitoring program and to determine when such monitoring activity should be undertaken. The current paper clarifies issues relating to the timing and informativeness of alternative monitoring systems.

Most of the literature using POMDPs has been fairly stylized, using very small state spaces that can be easily solved. Further progress in this area will depend on the use of algorithms that can handle larger state spaces, such as the approximation methods being developed in the robotics literature or approaches such as the one developed by Moore (2008).

An important issue not addressed here is that of structural uncertainty, which refers to incomplete knowledge of the values of the state transition equations and/or the mapping from states and actions into utility outcomes. Like observational uncertainty, structural uncertainty can be addressed using a belief state representation. Recently, there have been extensions to the POMDP framework that enable both structural uncertainty and partial observability to be handled in a common framework (Chadès et al., 2012; Fackler and Pacifici, 2014).

Another way the POMDP approach could be made more useful would be to add more institutional realism to the monitoring problem. Monitoring generally takes place over multiple sites and monitoring programs have limited budgets. It would be useful to extend the POMDP approach to situations in which decisions about which of a variety of monitoring systems are used in each of a collection of sites.

Acknowledgements

Thanks to Jinhau Zhao for pointing to the literature on informativeness of signals and to Steve Polasky for useful conversations on observational uncertainty. We thank James Nichols, Krishna Pacifici, Stephanie Snyder, and David Tomberlin for comments on the manuscript. Research was supported in part by Hatch Act funds (Fackler) and the U.S. Forest Service Northern Research Station (Haight).

References

- Blackwell, D., 1951. *The comparison of experiments*. In: Neyman, J. (Ed.), *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. University of California Press, Berkeley, CA, pp. 93–102.
- Cassandra, A., 1994. *Optimal Policies for Partially Observable Markov Decision Processes*. Technical Report CS-94-14. Brown University Department of Computer Science.
- Chadès, I., McDonald-Madden, E., McCarthy, M.A., Wintle, B., Linkie, M., Possingham, H.P., 2008. *When to stop managing or surveying cryptic threatened species*. *Proceedings of the National Academy of Sciences of the United States of America* 105 (37), 13936–13940.
- Chadès, I., Sabbadin, R., Carwardine, J., Martin, T.G., Nicol, S., Buffet, O., 2012. *MOMDPs: a solution for modelling adaptive management problems*. In: *Twenty-Sixth AAAI Conference of the Association for the Advancement of Artificial Intelligence (AAAI-12)*, Toronto, Canada, July 22–26, pp. 267–273.
- Clark, C.W., Kirkwood, G.P., 1986. *On uncertain renewable resource stocks: optimal harvest policies and the value of stock surveys*. *Journal of Environmental Economics and Management* 13, 235–244.
- Crowe, B., White, B., 2007. *A bioeconomic analysis of the duration of conservation contracts*. In: Presented at Australian Agricultural and Resource Economists Society 51st Annual Conference, Queenstown, New Zealand, February 13–16, 2007 <http://ageconsearch.umn.edu/handle/9457>
- Fackler, P.L., Pacifici, K., 2014. *Addressing structural and observational uncertainty in resource management*. *Journal of Environmental Management* 133, 27–36.
- Guestrin, C., Koller, D., Parr, R., Venkataraman, S., 2003. *Efficient solution algorithms for factored MDPs*. *Journal of Artificial Intelligence Research* 19, 399–468.
- Haight, R.G., Polasky, S., 2010. *Optimal control of an invasive species with imperfect information about the level of infestation*. *Resource and Energy Economics* 32, 519–533.
- Hirshleifer, J., Riley, J.G., 1992. *The Analytics of Uncertainty and Information*. Cambridge University Press, Cambridge.
- Howard, R.A., 1960. *Dynamic Programming and Markov Processes*. MIT Press, Cambridge, MA.

- Kaelbling, L.P., Littman, M.L., Cassandra, A.R., 1998. Planning and acting in partially observable stochastic domains. *Artificial Intelligence* 101, 99–134.
- Kurniawati, H., Hsu, D., Lee, W.S., 2008. SARSOP: efficient point-based POMDP planning by approximating optimally reachable belief spaces. In: *Proceedings of Robotics: Science and Systems*.
- Lane, D.E., 1989. A partially observable model of decision making by fishermen. *Operations Research* 37 (2), 240–254.
- Lovejoy, W.S., 1991. Computationally feasible bounds for partially observed Markov decision processes. *Operations Research* 39, 162–175.
- Marschak, J., Miyasawa, K., 1968. Economic comparability of information systems. *International Economic Review* 9 (2), 137–174.
- McDonald-Madden, E., Baxter, P.W.J., Fuller, R.A., Martin, T.G., Game, E.T., Montambault, J., Possingham, H.P., 2010. Monitoring does not always count. *Trends in Ecology and Evolution* 25 (10), 547–550.
- McDonald-Madden, E., Chades, I., McCarthy, M.A., Linkie, M., Possingham, H.P., 2011. Allocating conservation resources between areas where persistence of a species is uncertain. *Ecological Applications* 21, 844–858.
- Monahan, G.E., 1982. A survey of partially observable Markov decision processes: Theory, models, and algorithms. *Management Science* 28, 1–16.
- Moore, A.L., (Ph.D. Dissertation) 2008. Managing populations in the face of uncertainty: adaptive management, partial observability and the dynamic value of information. University of Melbourne.
- Moxnes, E., 2003. Uncertain measurements of renewable resources: approximations, harvesting policies and value of accuracy. *Journal of Environmental Economics and Management* 45, 85–108.
- Nichols, J.D., Williams, B.K., 2006. Monitoring for conservation. *Trends in Ecology and Evolution* 21, 668–673.
- Ong, S.C., Png, S.W., Hsu, D., Lee, W.S., 2009. POMDPs for robotic tasks with mixed observability. In: *Proceedings of Robotics: Science and Systems V*.
- Pineau, J., Gordon, G.J., Thrun, S., 2006. Anytime point-based approximations for large POMDPs. *Journal of Artificial Intelligence Research* 27, 335–380.
- Probert, W.J.M., Hauser, C.E., McDonald-Madden, E., Runge, M.C., Baxter, P.W., Possingham, H.P., 2011. Managing and learning with multiple models: objectives and optimization algorithms. *Biological Conservation* 144, 1237–1245.
- Regan, T.J., Chades, I., Possingham, H.P., 2011. Optimally managing under imperfect detection: a method for plant invasions. *Journal of Applied Ecology* 48, 76–85.
- Regan, T.J., McCarthy, M.A., Baxter, P.W.J., Panetta, F.D., Possingham, H.P., 2006. Optimal eradication: when to stop looking for an invasive plant. *Ecology Letters* 9 (7), 759–766.
- Sethi, G., Costello, C., Fisher, A., Hanemann, M., Karp, L., 2005. Fishery management under multiple uncertainty. *Journal of Environmental Economics and Management* 50, 300–318.
- Shani, G., Brafman, R., Shimony, S., 2007. Forward search value iteration for POMDPs. In: *Proceedings of International Joint Conference on Artificial Intelligence*, pp. 2619–2624.
- Smith, T., Simmons, R., 2005. Point-based POMDP algorithms: improved analysis and implementation. In: *Proceedings of Uncertainty in Artificial Intelligence*, pp. 542–547.
- Sondik, E.J., (Ph.D. Dissertation) 1971. The optimal control of partially observable Markov processes. Stanford University.
- Spaan, M.T.J., Vlassis, N., 2005. Perseus: randomized point-based value iteration for POMDPs. *Journal of Artificial Intelligence Research* 24, 195–220.
- Tomberlin, D., 2010. Endangered seabird habitat management as a Partially Observable Markov Decision Process. *Marine Resource Economics* 25 (1), 93–104.
- Tomberlin, D., Ish, T., 2007. When is logging road erosion worth monitoring? In: Sessions, J., Havill, Y. (Eds.), *Proceedings of the International Mountain Logging and 13th Pacific Northwest Skyline Symposium*. Oregon State University, College of Forestry Corvallis, Oregon, pp. 239–244.
- Walters, C., Hilborn, R., 1976. Adaptive control of fishing systems. *Journal of the Fisheries Research Board of Canada* 33, 145–159.
- White, B., 2005. An economic analysis of ecological monitoring. *Ecological Modelling* 189, 241–250.
- Williams, B.K., 2011. Adaptive management of natural resources – framework and issues. *Journal of Environmental Management* 92, 1346–1353.