WHEN IS LOGGING ROAD EROSION WORTH MONITORING?

David Tomberlin, Teresa Ish

Abstract: Efficient allocation of funds for erosion control on logging roads depends on information about current and potential future erosion on different road segments. Acquiring this information is typically expensive, and may make no immediate contribution to erosion control. Thus, managers face a trade-off between spending funds on information gathering versus on actual erosion control measures. Here, we develop a framework for examining this trade-off when current erosion, future erosion, and the efficacy of erosion control measures are all uncertain. Specifically, casting the manager's problem of allocating funds between erosion control and erosion monitoring as a partially observable Markov decision process (POMDP) allows us to identify the conditions under which costly estimates of erosion levels are worth obtaining as part of an adaptive erosion control program, and, in contrast, under what conditions the better strategy is to skip data acquisition and proceed directly to erosion control treatments. We demonstrate the POMDP approach through an application to a stylized road erosion control problem.

Key words: road erosion, monitoring, partially observable Markov decision process

Introduction

Sediment loading from logging roads impairs water quality and habitat conditions in many Pacific coastal rivers and streams. In this paper, we address the question of whether logging road erosion monitoring is worth the time and expense, given that we could decline to monitor in favor of either applying rehabilitative treatments immediately, without bothering to collect erosion data, or deferring the decision to implement road treatment or monitoring schemes (the more common practice). We examine this question within the framework of a partially observable Markov decision process (POMDP), which is well-suited to this purpose for at least two reasons. First, surface erosion is by its nature difficult to assess, even with special equipment, making the partial observability approach very apt. Second, logging road erosion control can be meaningfully represented in terms of a few states and actions, and the costs of these actions can be reasonably well estimated. Our model assumes the land manager wants to minimize long-run discounted total cost and will engage in monitoring only if it's expected to help with long-term performance. The model's purpose is to help the land manager decide when monitoring is worth the expense. Here, we apply the model to a single road segment, but it could also be applied to an entire watershed.

While there seems to be no consensus on a definition of 'adaptive management,' a necessary condition for management to be adaptive is that it account for the arrival of new information. Within the natural resource management literature, most work has focused on 'passive adaptive management,' in which new information is incorporated into decision making as it becomes available. A more difficult approach is that of 'active adaptive management,' in which new information is sought optimally: the manager considers the short-term cost of information gathering vs. the potential long-term benefits, and decides whether the costly information is worth having[1] .

Markov decision processes (MDPs), when solved with the techniques of stochastic dynamic programming, yield a mapping from the system state into an optimal policy, and may be thought of as a formal representation of adaptive management. However, MDPs assume that state variables are observed perfectly, an assumption that clearly does not hold in many natural resource management problems: animal populations, mineral reserves, and water quality, at least in many situations, cannot be known with certainty, and even developing good estimates is generally expensive and time-consuming.

The theory of partially observable Markov decisions processes (POMDPs) was developed in response to this shortcoming of MDPs, but no numerical algorithms existed for POMDP solution until Sondik (1971). Despite a steady stream of improvements in both exact and heuristic solution techniques, most applied work in dynamic optimization (including control engineering, economics, and behavioral ecology) has continued to rely on MDPs built around

[1] Though these ideas are taken from the control engineering literature, the most thorough treatment in a natural resource management context seems to be Walters (1986).

certainty-equivalent measures, rather than facing the numerical difficulties inherent in POMDPs. These difficulties are two-fold. First, POMDPs inherit from MDPs the well-known 'curse of dimensionality,' by which is meant that solution times explode as the number of admissible states and the length of the time horizon increase. Second, POMDPs are fundamentally Bayesian decision processes, in the sense that an agent's *beliefs* about state variables become the basis for the optimal decision rule. The agent may change these beliefs, via Bayes' Theorem, when new information becomes available. While this is conceptually appealing, the practical result is that we move from a world with finite MDP states to one with an infinite number of possible belief states, since an agent may come to have any set of beliefs, depending on how their prior beliefs and new information combine to yield updated beliefs. Thus, for POMDPs we can no longer use the standard techniques of stochastic dynamic programming as presented in, for example, Bertsekas (2000).

The difficulty in implementing POMDP solutions is a strong incentive to assume certainty-equivalence and stay within the relatively comfortable confines of MDPs. However, the substantial uncertainty inherent in many natural resource management problems really demands a better response. In our own work on salmon habitat management, for example, we have found that the amount of sediment loading from logging roads is highly uncertain, that simulation models do not inspire a great deal of confidence among managers, and that ascertaining empirical estimates of sediment loading requires a commitment of at least several years and tens of thousands of dollars. How can managers reasonably approach sediment control decisions when they don't even know (either at the watershed level or at the operational level of a particular road segment) the magnitude of the problem? More specifically, how much time and money should they sink into developing empirical estimates of sediment loading rates, when that time and money could be spent instead on road upgrades and decommissioning?

Questions of the same form arise anytime we consider resource management under uncertainty with an opportunity to invest in learning, which will generally mean incurring some short-term cost to achieve greater overall long-term net benefit. We believe the POMDP is the best existing tool for addressing such questions. Rather than make that argument directly, however, here we offer an expository example that we hope shows how the POMDP is precisely the tool needed to address the question of when road erosion monitoring programs are worth their costs. We emphasize that this model is for expository purposes only: while we believe the parameter values in our model are quite reasonable, they are not derived from field data. The results reported below relate only to this parameter set and no general lessons can be drawn from these results alone.

While our focus here is on logging road erosion, the question of whether to monitor or not is of considerably broader interest in natural resource management. In the context of Pacific coastal watersheds, various federal, state, and local governments, as well NGOs and community groups, are either monitoring or developing plans to monitor stream conditions, especially as they relate to fish habitat suitability. All this monitoring seems unobjectionable from a conservation point of view, and a lot of it is actually quite fun, so it might seem ungenerous to ask whether it's justified. However, even given the high level of enthusiasm and public funding for fish habitat monitoring, only a small fraction of streams can be monitored, and those only for a few habitat indicators and for a limited period of time. Given these practical constraints on monitoring, our question is an important one if we're serious about seeing that conservation efforts result in as much conservation benefit as possible.

Model

We begin with some general notes on POMDP models and then construct a stylized model that illustrates the use of POMDP for analyzing the desirability of an erosion monitoring program on a particular road segment.

The traditional MDP is a collection of sets $\{S, P, A, W\}$, where S represents state variables, P represents state dynamics as transition probabilities, A represents the actions available to an agent, and W represents the rewards to taking particular actions under particular conditions. A POMDP is an MDP with two additional components, a set of observations, Θ, and an observation model, R. Observations $\theta \in \Theta$ are the only information the agent has on the true state, S, which is unobservable. The observation model R describes the probabilistic relationship between observations θ and the true state S. In other words, the agent uses the observation model R to make inferences about the true state S based on noisy observations θ.

Solution of MDPs and POMDPs proceeds through a recursively defined value function V. In the case of POMDPs, this value function is:

$$V_t(\pi) = \max_a \left[\sum_i \pi_i q_i^a + \beta \sum_{i,j,\theta} \pi_i p_{ij}^a r_{j\theta}^a V_{t+1}[T(\pi \mid a, \theta)] \right] \quad (1)$$

where

π_i = subjective probability of being in state $i \in S$ at time t

q_i^a = immediate reward for taking action $a \in A$ in state $i \in S$ at time t

β = discount factor

p_{ij}^a = probability of moving from state $i \in S$ at time t to state $j \in S$ at time $t+1$ after taking action $a \in A$

$r_{j\theta}^a$ = probability of observing $\theta \in \Theta$ after taking action $a \in A$ and moving to state $j \in S$

T = function updating beliefs based on prior beliefs and observed θ

The value function V is simply the greatest expected net benefit that the agent can achieve over time, taking into account that as conditions change in the future, different actions may be warranted. From the solution of V we can also derive an optimal policy

$$\delta_t^*(\pi_i) = \arg\max_{a \in A} \left[\sum_i \pi_i q_i^a + \beta \sum_{i,j,\theta} \pi_i p_{ij}^a r_{j\theta}^a q_j^a \right] \quad (2)$$

which is a mapping from beliefs about the current state, π, into the optimal action. In other words, for any possible set of beliefs about the true state S at any time in the decision problem, the optimal policy identifies the action that will have the greatest long-term expected net benefit.

While the above formulation may seem abstract, the concept it represents is very intuitive. We live in a world that we understand almost exclusively through limited and imperfect observations. In terms of our road erosion management problem, we do not develop erosion control prescriptions based on known sediment loading rates, rather on the basis of what we *think* those rates are. Even an extensive (and expensive) field study can never tell us exactly how much sediment a particular road segment is producing—in most cases, a field study will only be able to generate an estimate of production on a few segments.

To make the POMDP formulation more concrete and its significance clearer, we now construct a particular POMDP that addresses the question of whether an erosion monitoring scheme should be implemented on a road segment. We consider the problem faced by a manager who has three actions available to address erosion on a road segment suspected of producing an unacceptable level of sidiment: to maintain the road as it is, to monitor the road's erosion level (by installing field instruments), or to treat the road (by adding rock or making design improvements). The first of these is quite inexpensive but does nothing to reduce current erosion rates or generate better estimates of these rates; the second is more expensive and does nothing to reduce erosion, but does provide information for subsequent decision-making; the last is quite expensive but has a good chance of effectively reducing sediment production on the road (though, to reiterate, the manager can't know with certainty either the erosion rate or the effectiveness of treatment).

In terms of the POMDP formulation, the action set A thus consists of {*maintain, monitor, treat*} and the state variable S is erosion. To keep the model computationally tractable and for ease of exposition, we restrict this state variable to only two possible values, *High Erosion* and *Low Erosion*. The observation set consists of the same two possible values, *High Erosion* and *Low Erosion*, but an observation of θ = *High Erosion* does not necessarily mean that the true state S = *High Erosion.* Instead, we define an observation model R as follows:

$$R_{j\theta}^1 = \begin{bmatrix} 0.6 & 0.4 \\ 0.4 & 0.6 \end{bmatrix} \quad R_{j\theta}^2 = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix} \quad R_{j\theta}^3 = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$$

Each matrix, with the state $j \in S$ defined by row and each observation θ defined by column, defines the probabilistic relationship of observation to true state under a different action. Because we have no data on these relationships, we have chosen parameters that provide a reasonable relative information content to observations under different actions. $R_{j\theta}^1$, for example, tells us that after taking action a=1 (*maintain*) and moving to the unobservable state *j=Low Erosion*, we would observe *θ=Low Erosion* with 60% probability and *θ=High Erosion* with 40% probability. That is, maintaining the *status quo* provides some information, presumably through casual observation of the road, but it is weak information. $R_{j\theta}^2$, in contrast, tells us that implementing a monitoring plan (a=2), yields a much stronger basis for inference based on observations: in this case, taking an observation when the true state is *j=Low Erosion* yields *θ=Low Erosion* with 90% probability and *θ=High Erosion* with 10% probability. Finally, $R_{j\theta}^3$ indicates that immediately after treating the road, observations tell us nothing about the true state of erosion. We impose this condition to reflect the fact that treatments themselves often cause transient changes in erosion rates that tell us little about the true state of the road.

The stochastic dynamics of the state S (the sediment production level) are given by transition probability matrices which we define as follows:

$$P_{ij}^1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad P_{ij}^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad P_{ij}^3 = \begin{bmatrix} 0.95 & 0.05 \\ 0.80 & 0.20 \end{bmatrix}$$

The first two matrices indicate that under actions a=1 and a=2 *(maintain* and *monitor*, respectively), a *Low Erosion* road will stay a *Low Erosion* road and a *High Erosion* road will stay a *High Erosion* road. P_{ij}^3 tells us that under a=3 (*treat)*, a *Low Erosion* road stays in that same state with 95% probability, but allows a 5% chance that the treatment will actually backfire and create a *High Erosion* road. Similarly, treating a *High Erosion* road has an 80% chance of successfully creating a *Low Erosion* road and a 20% chance of failure (meaning the *High Erosion* road stays that way). As with the observation model, these values are not derived from field data, but are chosen by us to reflect a plausible scenario for analysis.

Finally, the reward structure (actually, cost structure) in our model is as follows:

$$W_{ij\theta}^1 = \begin{bmatrix} -1 & -20 \\ -1 & -20 \end{bmatrix} \quad W_{ij\theta}^2 = \begin{bmatrix} -3 & -22 \\ -3 & -22 \end{bmatrix} \quad W_{ij\theta}^3 = \begin{bmatrix} -6 & -6 \\ -6 & -6 \end{bmatrix}$$

Here the columns of each matrix represent the possible states j and the rows represent possible observations θ. [We have suppressed the i-dimension of the cost structure since we assume the cost depends only on the state and not how the transition to the state occurred, as in the general POMDP formulation.] In each submatrix of W, the rows are the same because in our case the observation does not directly affect costs, which are in thousands of dollars. The matrix W^1 tells us that maintaining the road in *Low Erosion* state will cost $1000, which is very cheap compared to $20,000, the cost of maintaining the road in *High Erosion* state. W^2, the payoffs to monitoring, are the same as W^1 plus the cost of the monitoring program itself ($2000). That is, monitoring does nothing to change the costs associated with the erosion *per se*, it is a pure additional cost. W^3 tells us that treating the road will cost us the same $6000 regardless of whether the road is in *Low Erosion* or *High Erosion* state. Comparing all these costs, it's obvious that if the decision-maker knew the true state to be *Low Erosion*, the best choice would be to maintain the current situation (a=1), and if the decision-maker knew the true state to be *High Erosion*, the best thing to do would be to treat the road right away (a=3).

However, the premise of our model, and the reality that managers generally face, is that the true state is unknown.

Finally, we assume a discount factor of β=0.95, which completes our model specification.

Results

A POMDP solution consists of the recursively defined value function $V_t(\pi)$ and the associated optimal policy $\delta_t(\pi)$. Because the solution technique is rather complicated, we will not describe it here; interested readers can consult Cassandra (1994, pp. 45-54) for a good discussion of the Monohan/Eagle algorithm we used. POMDP algorithms share with MDP algorithms the basic notion of backward recursion from an arbitrarily defined end of time, T. In our model, time T comes after all decisions have been made and after uncertainty about the true state has been resolved (which allows unambiguous values to be assigned to each of the possible final states).

Figure 1 shows the value function at the final time period in which a decision is to be made, *T-1*. The x-axis is the belief simplex for the two possible states in *S*: p(*Low Erosion*) runs from left to right, and since p(*High Erosion*) must be 1-p(*Low Erosion*), p(*High Erosion*) runs from right to left. The y-axis is the expected value of taking particular actions. The blue line is the value function, giving the expected value at *T-1* of taking whichever action has the lowest expected cost. Here, since there is only one decision period before the end of the problem, these values have very straightforward interpretations. If the manager's current belief is that p(*Low Erosion*) is anything less than 74%, the optimal action is to treat the road, which has a payoff of –6 regardless of the current true state. If, however, the manager's current belief is that p(*Low Erosion*) > 74%, then the optimal action is to maintain the *status quo* road, which has an expected value of [–1*p(*Low Erosion*) + –20*p (*High Erosion*)]. In other words, the more certain the manager is that the true state is in fact *Low Erosion*, the greater the expected payoff to doing simple maintenance work. Thus, Figure 1 not only shows the value function but also partitions the belief space into regions on which each possible action is optimal, i.e., visually lays out the optimal policy.

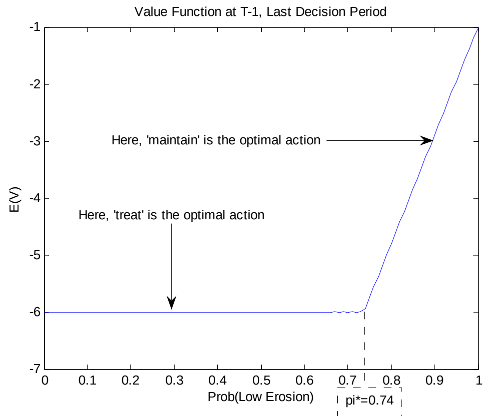

Fig. 1: The value function in the last decision period, showing the partition of the belief space into regions associated with different optimal actions. Here, *treat* is optimal for beliefs p(*Low Erosion*) <= 74%, and *maintain* is optimal for beliefs p(*Low Erosion*) > 74%.

Notice that the action *monitor* does not appear as part of the optimal strategy in Figure 1. The reason is simply that, with no further actions to be taken after *T-1*, there is no justification for paying to gather information that can't provide future benefits in the form of improved decision making.

Of course, we are almost always interested in problems that have at least several decision periods. Figure 2 shows the evolution of the value function over a 10-year time horizon. The most obvious effect of lengthening the time horizon is that the value function moves steadily downward, due to the expectation of increased future costs (a direct result of our model setup). However, the shape of the value function also changes, as do the actions that form the optimal policy. Specifically, for T-3 and all earlier periods, *monitor* becomes part of the optimal strategy. The belief ranges for which *monitor* is optimal are between the two black curves; the beliefs to the left of the left-most black curve are those for which the optimal action is *treat*, and those to the right of the right-most black curve are those for which the optimal action is *maintain*. The belief range over which *treat* is optimal is neither strictly increasing nor strictly decreasing with time, because the optimal policy has not fully converged to a stable mapping, but the general picture is clear. Monitoring enters the optimal policy at T-3, once the time horizon has become long enough that information generated by monitoring can yield sufficient benefits (in expectation) to offset the cost of the monitoring program. As the time horizon deepens, the range of beliefs over which monitoring is part of the optimal strategy increases, from about [0.81 0.90] at T-3 to about [0.81 0.97] at T-10. Immediate treatment is still the optimal action for beliefs up to around p(*Low Erosion*)=0.8, and simply maintaining the status quo is preferred only for beliefs such that p(*Low Erosion*) is well over 0.9.

The above discussion and figures may seem abstract, but in fact they correspond quite nicely to the way most of us get through life. We routinely make decisions that are based not on directly observable facts, but on our beliefs about those underlying facts, which for a variety of reasons we either can't or don't want to know with certainty. Both the facts and our beliefs about them may change over time, but at any point in time we make decisions based on our beliefs at that time. People can't apply Bayes' theorem with the efficiency that a computer can, but the general notion of combining new information with prior beliefs seems to reflect a lot of human decision making. Perhaps more importantly for our present purposes, the partition of the belief space into regions corresponding to different optimal actions is very intuitive and also useful. As Figure 2 shows, one person may believe that p(*Low Erosion*)=0.1, another that p(*Low Erosion*)=0.4, and a third that p(*Low Erosion*)=0.7, but the POMDP makes clear that they should all still be able to agree on treating the erosion problem immediately.

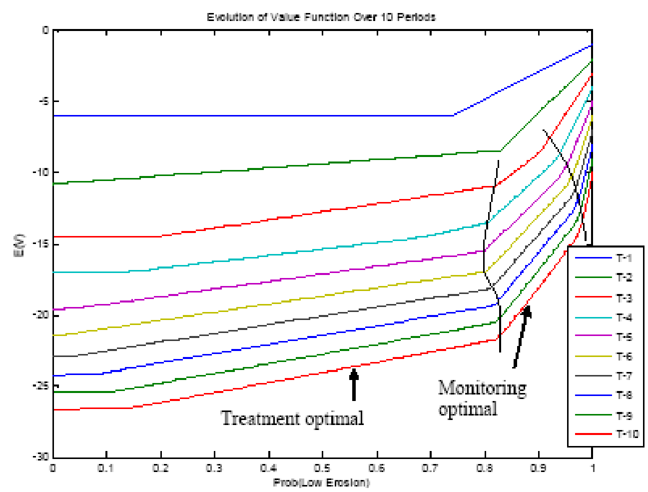

Fig. 2: The evolution of the value function over 10 decision periods. The value function moves monotonically downward as the time horizon increases, which is an artifact of our cost-only model. For T-3 and earlier periods, monitoring becomes part of the optimal strategy for those beliefs between the two black curves.

As we mentioned at the outset, these results are a function of our model's structure and parameters, and are not to be understood as generally applicable.

Conclusions

The partially observable Markov decision process (POMDP) provides a formal framework for exploring when information-gathering is likely to be worth the cost and when not. Given the expense of monitoring programs in natural resource management, budgetary realities ensure that managers have to choose among candidate monitoring programs. The POMDP provides a tool for thinking carefully about such choices.

Here, we have presented an application of POMDP to a stylized problem in logging road management. Because POMDPs are even more subject than traditional MDPs to the 'curse of dimensionality,' research on numerical techniques for POMDP solution is a very active field in engineering and artificial intelligence—more fully developed applications in natural resource management will have to draw on recent advances in heuristic techniques such interior-point methods and witness algorithms. However, even our simple example has shown that, under plausible parameter values, the costs of monitoring can easily exceed the benefits. For the parameter set assumed above, we found that, for problems with a time horizon of at least 3 decision periods, implementing erosion control treatments *without first monitoring* was optimal as long as the subjective probability of the road being highly erosive was more than about 20%. Monitoring was optimal over a narrower range of beliefs, specifically when the belief that existing road conditions were good was between about 80% and about 95%. That is, monitoring in our example was preferred only when the manager had a pretty strong hunch that current conditions were good, in which case the monitoring served essentially to rule out the need for more aggressive and expensive treatment.

In developing our case for the POMDP as a useful decision-making tool, we deliberately touched lightly on the nature of the subjective probabilities π, which are really the heart of the POMDP. While from a technical point of view there's not much to say about π, which is simply a vector of conditional probabilities, we should address a concern that might arise from a philosophical point of view. Some may object that subjective probabilities have no place in management or policymaking, which should strive at all times to be as objective as possible. Without rehashing the centuries-long Bayesian-vs-frequentist struggle, we note that many Bayesians consider subjective probability the only sensible notion of probability, and so would dismiss this criticism as invalid on principle. For our purposes, it doesn't seem necessary to take that rigorous Bayesian position. We are satisfied with the more mundane argument that subjective probabilities are so manifestly the basis for current natural resource management and policymaking that few (if any) decisions would be made without them. In short, we think subjective probabilities in natural resource decision making are perfectly sensible and almost perfectly unavoidable.

The deep uncertainty we face in many aspects of natural resource management requires that we think carefully about when to invest in learning about the systems we manage. The POMDP provides a coherent framework for such thinking.

Literature Cited

Bertsekas, D. 2000. Dynamic Programming and Optimal Control. Belmont, MA: Athena Scientific.

Cassandra, A. 1994. Optimal Policies for Partially Observable Markov Decision Processes. Brown University Department of Computer Science Technical Report CS-94-14.

Sondik, E.J. 1971. The Optimal Control of Partially Observable Markov Decision Processes. Ph.D. dissertation, Stanford University.

Walters, C. 1986. Adaptive Management of Renewable Resources. New York: Macmillan Publishing.

Author Contact Information

David Tomberlin
National Marine Fisheries Service
110 Shaffer Rd.
Santa Cruz, CA, 95060
TEL 831.420.3910
david.tomberlin@noaa.gov

Teresa Ish
University of California, Santa Cruz
Department of Applied Math
1156 High Street
Santa Cruz, CA 95063
Currently Environmental Defense
ishsurf@gmail.com

PROCEEDINGS OF THE

INTERNATIONAL MOUNTAIN LOGGING AND 13TH PACIFIC NORTHWEST SKYLINE SYMPOSIUM

Sponsored by

Department of Forest Engineering
Oregon State University

International Union of Forestry
Research Organizations

Edited by John Sessions and Yvonne Havill

Corvallis, OR
April 1-6, 2007