On the Potential Use of Adaptive Control Methods for Improving Adaptive Natural Resource Management

Craig A. Bond
Assistant Professor
Department of Agricultural and Resource Economics
Colorado State University
Ft. Collins, CO 80526
Ph: (970) 491-6951
Fax: (970) 491-2067

DARE CSU Working Paper 002-08
December 9, 2008

Abstract: The paradigm of adaptive natural resource management (AM), in which experiments are used to learn about uncertain aspects of natural systems, is gaining prominence as the preferred technique for administration of large-scale environmental projects. To date, however, tools consistent with economic theory have yet to be used to either evaluate AM strategies or improve decision-making in this framework. Adaptive control (AC) techniques provide such an opportunity. This paper demonstrates the conceptual link between AC methods, the alternative treatment of realized information during a planning horizon, and AM practices; shows how the different assumptions about the treatment of observational information can be represented through alternative dynamic programming model structures; and provides a means of valuing alternative treatments of information and augmenting traditional benefit-cost analysis through a decomposition of the value function. The AC approach has considerable potential to help managers prioritize experiments, plan AM programs, simulate potential AM paths, and justify decisions based on an objective valuation framework.

KEYWORDS: adaptive control, adaptive management, dynamic programming, value of experimentation, value of information

On the Potential Use of Adaptive Control Methods for Improving Adaptive Natural Resource Management

1. Introduction

In recent decades, ecologists and other natural scientists have increasingly encouraged the use of "adaptive management" (AM) to address relatively large-scale environmental projects (e.g., Everglades and Grand Canyon restoration) characterized by considerable uncertainty about the effects of controllable actions [1, 2]. A key principle of AM is that management projects or policies are considered means of testing hypothesis about the response of the ecological system, and the results of these "experiments" are monitored and evaluated before taking additional action (such as continuing the project/policy or changing course). The experiments thus provide information about the system and reduce uncertainty [3, 4]. Because the results of actions are not known ex ante and ecosystems are constantly evolving (in an often unknown manner), traditional benefit-cost techniques are not well-suited to inform decision-makers about the ramifications of alternative management choices [5].

In a mostly separate literature, macroeconomists have focused on similar problems of uncertainty, experimentation, and learning in a formal dynamic framework termed "dual control" or "adaptive control" (AC) [6, 7]. With an interesting intellectual history starting in the engineering literature and briefly sidelined due to the embrace of rational expectations (see [8] for the entertaining full story), the dual control literature has documented many of the challenges related to parameter uncertainty and learning in dynamic models (e.g., non-continuities and non-convexities, an exponentially-expanding state space, etc…), yet these methods have yet to penetrate mainstream environmental and resource economics [9]. Of course, there are examples that involve either Bayesian updating in a management context (e.g., [10] related to nonpoint source pollution control) and/or passive AM in which the optimization and updating stages are decoupled (e.g., [11] related to climate change and [12] related to the problem of polluting a shallow lake), but little formal analysis of active AM strategies has been completed. Exceptions include papers related to stock pollution/climate change [13], ship examination and invasive species [14], and general environmental policy [15].

Although it has evolved from the macroeconomic modeling literature, the AC methodology is well-suited to the analysis of many environmental and natural resource management problems for which basic scientific stimulus/response relationships are uncertain or unknown. In these circumstances, managers must take into account not only their physical objectives (e.g., population restoration of an endangered species), but also the fact that actions will ultimately provide information about the response of the ecosystem. This information can be subsequently used to improve future decisions as the degree of uncertainty about the system has been reduced. In addition, managers may have to choose between competing actions or experiments subject to a budget constraint, and thus must take the expected value of future information into account. In the next few sections, we illustrate how AC can be used to these ends.

This paper contributes to the literature in several ways. First, we demonstrate the conceptual link between AC methods, the alternative treatment of realized information during a planning horizon, and adaptive natural resource management. Thus, we bring together strands of research and management that have heretofore evolved separately, yet exhibit a number of complementarities that can be exploited by natural resource managers to make more efficient decisions. Second, we show how the different assumptions about the treatment of observational information can be represented through alternative dynamic programming model structures, including the differences between “passive” and “active” learning on the necessary first-order conditions that ultimately define the feedback decision rules that govern optimal management behavior. These conditions thus define the optimality of experimentation (defined as deviation from a similar non-learning feedback rule), though generalizations regarding optimal experimentation appear to be problem-specific. Finally, we provide a means of valuing alternative treatments of information through a decomposition of the value function, and illustrate how this decomposition can be used to augment traditional benefit-cost analysis through the incorporation of the expected value of information. This and other similar uses of AC techniques have considerable potential to help adaptive natural resource managers make more objective and informed decisions, distinguish between management plans, and choose between potential experiments based on a theoretically-consistent model of dynamic uncertainty that incorporates both natural and informational states.

This paper proceeds as follows. We first provide some background on AM in practice, then present the structure of a general set of models that can be used to generate decision rules and ex post and expected ex ante values of following those rules under different treatments of information. We then discuss how these results can be applied for decision-making in an adaptive natural resource management setting, and discuss the limitations of the AC technique. A final section concludes.

2. Adaptive Management (AM) in Practice

The optimal management of large, complex environmental systems is characterized with uncertainties and incomplete information along multiple dimensions and scales. AM techniques feature the resolution of uncertainty through closed-loop experimental learning on a relatively smaller scale than the grand project. Experiments are used to test hypotheses on the response of ecosystems to intervention, then monitored, and continued/expanded if the hypothesis regarding the response is not rejected, but subject to revision if rejected [16]. The resultant management path is thus path-dependent and dependent on both the value of information and the costs of acquiring it [3].

Billions of dollars are being spent directly on AM in the Everglades, while hundreds of millions in foregone hydropower is being incurred due to AM implementation in Grand Canyon [1, 17]. AM is also being implemented for projects related to the Missouri river [18], and is being considered by public land agencies such as the Bureau of Land Management and the Council on Environmental Quality (for use in economic impact statements). As noted in [1, 2], environmental and natural resource economists have been largely absent and silent on the benefits and costs of AM. However, there is interest among decision makers and those that fund AM for evaluative tools that can help agencies properly allocate resources in a context of dynamic, closed-loop decisions, rather than the traditional open-loop (and sometimes static) benefit-cost framework.

One technique that offers promise in this regard, but has not been fully explored in the environmental and natural resource economics literature, is AC. In the next section, we discuss a general representation of an AM model in an AC framework, with particular attention to the treatment of information and the expected values of treating information in different manners.

3. A General Adaptive Management (AM) Model using Adaptive Control (AC)[1]

The canonical model of an autonomous environmental AM problem can be constructed as follows. Define the instantaneous payoff (utility) function as $u(c_t, x_t)$, where c_t is an M-dimensional vector of control variables (e.g., management options, production decisions, etc…)

[1] Much of the notation in this section was inspired by [18]. In the discussion that follows, we will assume that the relevant functions (objective, state, and value) exist and are differentiable in all of their arguments, and that the Bellman equation is a contraction mapping (and thus can be approximated through, say, numerical methods).

and x_t is a K-dimensional vector of environmental states (populations of species, stocks of pollution, etc…). States evolve according to the system of difference equations

$$x_{t+1} = f\left(c_t, x_t; \gamma, \bar{\alpha}\right) + \varepsilon_{t+1}, \tag{1.1}$$

where γ is a vector of parameters known to the manager, $\bar{\alpha}$ are the true, but unknown (to the controller) parameters of the system, and ε_t are *i.i.d.* white noise shocks to the system with known variance.

The manager, however, does have prior beliefs over the uncertain parameters represented by the prior probability distribution $p(\alpha \mid \theta_t)$, where θ_t are the N sufficient statistics describing the joint probability distribution at time *t*. Assuming updating of prior beliefs using Bayes rule prior to decision in the next period, the posterior distribution can be described as $p(\alpha \mid \theta_{t+1}) = \dfrac{p(x_{t+1} \mid c_t, x_t, \alpha, \theta_t)\, p(\alpha \mid \theta_t)}{p(x_{t+1} \mid c_t, x_t, \theta_t)}$.[2] Assuming existence, this implicitly defines the evolution of the sufficient statistics (and thus the evolution of the beliefs over the uncertain parameters) as a differential equation dependent on the choice of controls and resultant states, or

$$\theta_{t+1} = G\left(c_t, x_{t+1}, x_t, \theta_t\right). \tag{1.2}$$

Assuming an infinite horizon and discounting at factor $0 < \beta < 1$, the dynamic programming formulation of this problem becomes[3]

$$\begin{aligned} V\left(x_t, \theta_t\right) &= \max_{c_t} \left\{ u(c_t, x_t) + \beta E\left[V\left(x_{t+1}, \theta_{t+1}\right)\right]\right\} \\ &= \max_{c} \left\{ u(c_t, x_t) + \beta E\left[V\left(f\left(c_t, x_t; \gamma, \alpha\right) + \varepsilon_{t+1}, G\left(c_t, f\left(c_t, x_t; \gamma, \alpha\right) + \varepsilon_{t+1}, x_t, \theta_t\right)\right)\right]\right\}, \end{aligned} \tag{1.3}$$

[2] Of course, this is but one of a number of updating rules possible.
[3] The assumption of an infinite time horizon suggests that calendar time is inconsequential in the Bellman equation; however, for notational convenience, we retain the time subscripts to indicate adjacent periods and the information set in time *t*.

where the expectation is over all random parameters in the model conditional on the information available at time *t*. Let $q(\varepsilon)$ denote the distribution of the white noise terms on the state equations, and substitute the relevant distributions into (1.3) to obtain a (somewhat) operational Bellman equation:

$$\begin{aligned} V\left(x_t,\theta_t\right) &= \max_c \left\{u(c_t,x_t)\right. \\ &+\beta\int_{\mathbb{R}^K}\int_{\mathbb{R}^N} V\left(f\left(c_t,x_t;\gamma,\alpha\right)+\varepsilon_{t+1},G\left(c_t,f\left(c_t,x_t;\gamma,\alpha\right)+\varepsilon_{t+1},x_t,\theta_t\right)\right)p(\alpha\mid\theta_t)q(\varepsilon)d\alpha d\varepsilon\left.\right\}. \end{aligned} \tag{1.4}$$

This specification is fairly general, potentially nonlinear in the objective and state transition equation, and allows for multiple uncertain parameters. Note the value function depends on physical as well as informational states, as do the updating equations for information.

Alternative Information Treatments

As noted in [13, 8] and implied by [19], the manager has a number of ways of dealing with the expected impact of information and the uncertainty surrounding the parameters. Options include: a) treating the uncertain parameters as true values at their expected levels with zero variance and no updating of sufficient statistics; b) treating the uncertain parameters as uncertain (positive variance) and no updating of sufficient statistics; c) treating the uncertain parameters as uncertain with passive updating of sufficient statistics (i.e., decoupling the optimization and updating procedures); and d) treating the uncertain parameters as uncertain with active updating (or active learning) of sufficient statistics (i.e., endogenous updating of the distribution of the uncertain parameters). In an environmental AM framework, all of these solutions might be of interest, depending on the circumstance; however, it seems likely that the expected value of active experimentation, as given by the difference in value between c) and d), is of prominent importance, given the question of when (and how) to experiment.

Generalizing [8] and concentrating on the dynamic programming framework, the four information treatments translate directly into alternative specifications of the dynamic programming problem, and thus the value function. If the decision maker does not take into account parameter uncertainty (but does take into account the stochastic nature of the state variables) and does not learn, then the appropriate Bellman equation is

$$V^{ce}\left(x_t;\theta_0\right)=\max_{c_t}\left\{u(c_t,x_t)+\beta\int_{\mathbb{R}^K}V^{ce}\left(f\left(c_t,x_t;\gamma,\hat{\alpha}\right)+\varepsilon_{t+1};\theta_0\right)q(\varepsilon)d\varepsilon\right\}, \tag{1.5}$$

where $\hat{\alpha}$ is the expected value of the unknown parameters based on the prior distribution characterized by the constant set of sufficient statistics θ_0. Note that the value function $V^{ce}\left(\cdot\right)$, in this case, is only a function of the physical (as opposed to informational) state variables, while the prior distribution is assumed fixed over the planning horizon.

Case b), in which the decision maker takes the uncertainty of the parameters into account but does not learn, is characterized by

$$\begin{aligned}V^{s}\left(x_t;\theta_0\right)&=\max_{c_t}\left\{u(c_t,x_t)\right.\\&\left.+\beta\int_{\mathbb{R}^K}\int_{\mathbb{R}^N}V^{s}\left(f\left(c_t,x_t;\gamma,\alpha\right)+\varepsilon_{t+1};\theta_0\right)p(\alpha\mid\theta_0)q(\varepsilon)d\alpha d\varepsilon\right\},\end{aligned} \tag{1.6}$$

where once again, the optimal value function is only a function of the physical states with the prior distribution taken as given and fixed. The fundamental difference between the two problems is the fact that (1.6) represents an expectation over the (fixed) parameter distribution, while (1.5) just substitutes the certainty-equivalence parameter values.

Following [13], we now investigate the difference in incentives with respect to the control variables between a passive and active management strategy. In each case, the value function is of the form $V\left(x_t,\theta_t\right)=\max_{c_t}\left\{u(c_t,x_t)+\beta E\left[V\left(x_{t+1},\theta_{t+1}\right)\right]\right\}$, but by assumption, the passive learner assumes (in the maximization only) $\theta_{t+1}=\theta_t$. Prior to the next decision, however, the passive learner updates based on (1.2), resulting in a (potentially) new posterior distribution with updated sufficient statistics. However, this updating is *not* anticipated in the formation of the feedback decision rule.

In the most general form, then, the first-order conditions with respect to the controls, assuming differentiability, are

$$
\begin{aligned}
&u_c(c_t, x_t) + \beta \frac{\partial E\left[V\left(x_{t+1}, \theta_{t+1}\right)\right]}{\partial c} \\
&\quad = u_c(c_t, x_t) + \beta \int_{\mathbb{R}^{K+N}} \left(\frac{\partial V}{\partial x}\frac{\partial f}{\partial c} + \frac{\partial V}{\partial \theta}\left(\frac{\partial G}{\partial c} + \frac{\partial G}{\partial x_{t+1}}\frac{\partial f}{\partial c} \right) \right) p(\alpha \mid \theta_t) q(\varepsilon) d\alpha d\varepsilon \overset{\text{set}}{=} 0,
\end{aligned}
\tag{1.7}
$$

where the partial derivatives are evaluated at the state values at time *t*. Thus, the optimal policy follows an augmented form of the standard first-order condition, which states that at the margin, instantaneous effects of the chosen controls must be offset by the discounted expected net present value of all future effects, now including effects on both the state path through $\frac{\partial V}{\partial x}\frac{\partial f}{\partial c}$ and the effects on the expected value of information $\frac{\partial V}{\partial \theta}\left(\frac{\partial G}{\partial c} + \frac{\partial G}{\partial x_{t+1}}\frac{\partial f}{\partial c} \right)$.

For the passive management problem, since $G(\cdot) = \theta_t$, the term $\left(\frac{\partial G}{\partial c} + \frac{\partial G}{\partial x_{t+1}}\frac{\partial f}{\partial c} \right) = 0.$ This term can be interpreted as the sum of the marginal effect of the control variables on the updating of sufficient statistics and the marginal effect of the resultant states on the updating of the sufficient statistics, working through the controls. In other words, it is the aggregate marginal effect of a change in control on the beliefs of the controller. So long as this aggregate effect is different than zero, and the shadow value on at least one of the sufficient statistics is non-zero (i.e., $\frac{\partial V}{\partial \theta} \neq 0$), some degree of experimentation (defined here as a deviation from the passive management strategy), will be optimal. If, however, a change in control does not change the prior, then clearly there is no benefit to deviation (and, in fact, the control rule is independent of θ). This result is a generalization of Theorem 1 in [13].

While this may seem a trivial result, consider a single-dimension, linear state equation with either an uncertain constant term and/or state term with uncertain constant marginal effects on the future realization of that state. In our notation, this implies a state equation of the form $x_{t+1} = \alpha_0 + \alpha_1 x_t + \gamma c_t + \varepsilon_{t+1}$, where α_1 and α_2 are uncertain. In addition, in order to simplify either analytical or numerical analysis, discrete distributions are often used to describe the uncertainty; e.g., $P\left[\alpha = \alpha_i\right] = \pi_i$, $\sum_{\mathrm{i}=1}^{\mathrm{Q}} \pi_{\mathrm{i}} = 1$ for the Q support values α_i of the uncertain parameter. As shown in appendix A, if this is the only form of parameter uncertainty, then it will never be optimal to deviate from the passive learning solution, as deviation would simply shift the entire distribution regardless of the true value of the unknown parameter [13].

Second, this result has significance for selection of experiments in environmental AM; namely, ceteris paribus, higher valued experiments are likely to be those in which the direct marginal effects of a management action are relatively unknown (in which the updating is fairly large), rather than ones in which the marginal effect is more certain. Note, however, that this is *not* to say that undertaking such experiments is always optimal, since the benefits of the information gathering must be balanced against the expected effects on the value of the future state variables and instantaneous utility (in other words, the endogenous costs).

Turning to the first-order conditions with respect to the state variables, application of the envelope theorem to (1.3) yields:

$$
\begin{aligned}
V_x\left(x_t,\theta_t\right) &= u_x(c_t,x_t)+\beta\,\frac{\partial E\left[V\left(x_{t+1},\theta_{t+1}\right)\right]}{\partial x} \\
&= u_x(c_t,x_t)+\beta\int\left(\frac{\partial V}{\partial x}\frac{\partial f}{\partial x}+\frac{\partial V}{\partial \theta}\left(\frac{\partial G}{\partial x_{t+1}}\frac{\partial f}{\partial x}+\frac{\partial G}{\partial x}\right)\right)p(\alpha\,|\,\theta_t)q(\varepsilon)d\alpha d\varepsilon,
\end{aligned}
\tag{1.8}
$$

$$
V_\theta\left(x_t,\theta_t\right)=\beta\,\frac{\partial E\left[V\left(x_{t+1},\theta_{t+1}\right)\right]}{\partial \theta}=\beta\int\left(\frac{\partial V}{\partial \theta}\frac{\partial G}{\partial \theta}\right)p(\alpha\,|\,\theta_t)q(\varepsilon)d\alpha d\varepsilon. \tag{1.9}
$$

For the passive management problem, the term $\left(\frac{\partial G}{\partial x_{t+1}}\frac{\partial f}{\partial x}+\frac{\partial G}{\partial x}\right)=0$ in (1.8) and $\left(\frac{\partial V}{\partial \theta}\frac{\partial G}{\partial \theta}\right)=\frac{\partial V}{\partial \theta}$ in (1.9). Note that from (1.8), the marginal value of the physical states for the active manager now includes not only the marginal utility now and in the future resulting from changing that physical state, but also the impact of that state on the future value of information through the information updating procedure (in the active management case). Similarly, the evolution of the marginal value of information itself depends on the expectation of *all* future information gains. For the passive manager, however, there is no expectation that the prior and posterior distributions will differ from one period to the next. Thus, there is no incentive to deviate from the non-learning case in which the uncertain parameters are treated as such.

An Alternative Decomposition of the Active Learning Value Function

In seminal works in the engineering literature, Bar-Sahlom and Tse [6] and Kendrick [7] develop what is known as the “dual control” method of solving stochastic dynamic programming problems in the presence of uncertain parameters and learning (as well as potential errors in measurement, not discussed in the present paper). This method uses a second-order Taylor-series expansion around the future certainty-equivalent, or “nominal” path, and the resultant quadratic problem has a (complicated) analytical solution. The approximate value function from this technique, termed "cost-to-go" as the problem is posed in a dynamic cost-minimization framework, can be divided into what are termed deterministic, cautionary, and probing terms, where the first is self-explanatory, the second is a function of the covariance of the projected one-period-ahead state variable covariance matrix and the stochastic shock covariance matrix, and the third (probing) term is a function of the expected covariance matrix of the state equations in each remaining period after the state variables are measured.

This decomposition has been studied extensively, with one of the main results showing that the optimal value function can exhibit nonconvexities caused by the cautionary or probing terms [20, 21, 22, 23, 24]. Of course, these nonconvexities can lead to local, rather than global, solutions using numerical techniques, especially those that use gradient methods.

Although not completely analogous to the "cost-to-go" decomposition, we can use the associated value functions of our cases discussed above to value the different treatment of information in the system, which could be of considerable interest to managers engaged in adaptive environmental management. For the sake of notation, define the optimal feedback rules as

$$c^i\left(x;\theta_0,\gamma,\beta\right),\ i=\{ce,s\}, \tag{1.10}$$

for the certainty-equivalent problem (1.5) and the stochastic parameter problem (1.6), respectively. The resultant state at time period t from following this rule from the beginning of the planning horizon in each case is thus recursively defined by

$$x_t^i = f\left(c^i\left(x_{t-1}^i;\theta_0,\gamma,\beta\right),x_{t-1}^i;\gamma,\overline{\alpha}\right)+\varepsilon_t,\ i=\{ce,s\},\ t=(1,\ldots,t), \tag{1.11}$$

Ex post, the cumulative *realized* value of treating the parameter vector as random up until time t is thus the difference in the net present value of the objective following the decision rules of the certainty-equivalent and stochastic parameter problems, or

$$\sum_{t=0}^{t-1}\beta^t\left[u\left(c_t^s\left(x_t^s\right),x_t^s;\theta_0\right)-u\left(c^{ce}\left(x_t^{ce}\right),x_t^{ce};\theta_0\right)\right], \tag{1.12}$$

taking the prior as fixed. On the other hand, the ex ante *expected* value of treating the parameter vector as random, assuming certain starting stock levels of x_t at time *t*, is

$$V^s\left(x_t;\theta_0\right)-V^{ce}\left(x_t;\theta_0\right). \tag{1.13}$$

A similar expression can be derived for the cumulative realized values of optimal experimentation (letting the superscripts *al* and *pl* indicate passive and active learning, respectively):

$$\sum_{t=0}^{t-1}\beta^t\left[u\left(c^{al}\left(x_t^{al},\theta_t^{al}\right),x_t^{al}\right)-u\left(c^{pl}\left(x_t^{pl}\right),x_t^{pl};\theta_t\right)\right], \tag{1.14}$$

noting that the decision rule for the passive learning problem is independent of the state of the prior (save for its use in calculating the expectation over uncertain parameter values), and with the notation θ_t indicating that the prior is updated, but decoupled from the optimization problem. This fact, however, complicates the expression for the ex ante expected value of experimentation. Specifically, $V^{pl}\left(x_t,\theta_t\right)$, the net present value of expected net benefits *assuming* that the prior distribution will not change (recall that $\theta_{t+1}=\theta_t$ for the passive learner), is not the correct term to use when valuing the expected gains from active experimentation, and, in fact, is equal to $V^s\left(x_t;\theta_0\right)$.

Instead, define the evolution of the sufficient statistics vector from following the passive learning feedback rule $c_t^{pl}(x_t^{pl})$ as $\theta_{t+1}^{pl}=G\left(c_t^{pl}(x_t^{pl}),x_{t+1}^{pl},x_t^{pl},\theta_t^{pl}\right)$, and the associated value function

$$\begin{aligned}V^{pl,\theta}\left(x_t,\theta_t\right)&=u\left(c^{pl}\left(x_t\right),x_t\right)+\\ &\beta\int_{\mathbb{R}^K}\int_{\mathbb{R}^N}V\left(f\left(c^{pl}\left(x_t\right),x_t;\gamma,\alpha\right)+\varepsilon_{t+1},G\left(c_t^{pl}(x_t^{pl}),x_{t+1},x_t,\theta_t\right)\right)p(\alpha\mid\theta_t)q(\varepsilon)d\alpha d\varepsilon.\end{aligned}$$

Note that these expressions are not "optimal" in the sense that they do not satisfy the restricted first-order conditions (1.7)-(1.9), but rather are based on the passive learning feedback rule. As such, they reflect the expected value of following that rule *and* the decoupled updating that occurs. The appropriate expression for the value of (optimal) experimentation is thus

$$V^{al}\left(x_t,\theta_t\right)-V^{pl,\theta}\left(x_t,\theta_t\right). \tag{1.15}$$

Finally, the ex post and ex ante values of passive learning over the stochastic parameter treatment can be defined as

$$\sum_{t=0}^{t-1}\beta^t\left[u\left(c_t^{pl}\left(x_t^{pl}\right),x_t^{pl};\theta_t\right)-u\left(c^s\left(x_t^s\right),x_t^s;\theta_0\right)\right], \tag{1.16}$$

and

$$V^{pl,\theta}\left(x_t,\theta_t\right)-V^s\left(x_t;\theta_0\right). \tag{1.17}$$

Putting the ex ante expectation terms together, we can use these different value functions to break the active learning value function into certainty equivalent, uncertainty, passive learning, and active learning terms; namely,

$$\begin{aligned}V^{al}\left(x_t,\theta_t\right)&=V^{ce}\left(x_t;\theta_0\right)+\left[V^s\left(x_t;\theta_0\right)-V^{ce}\left(x_t;\theta_0\right)\right]\\&\quad+\left[V^{pl,\theta}\left(x_t,\theta_t\right)-V^s\left(x_t;\theta_0\right)\right]+\left[V^{al}\left(x_t,\theta_t\right)-V^{pl,\theta}\left(x_t,\theta_t\right)\right].\end{aligned} \tag{1.18}$$

In contrast to the decomposition of the quadratic-quadratic approximation in [6, 7], (1.18) gives an exact decomposition of the expected values of different informational treatments, rather than separating the approximate "cost-to-go" into one deterministic and two stochastic terms, with the latter decomposed into two terms depending on the effect of the controls on each. In practice, however, a closed form of the decomposition is extraordinarily unlikely to be found, though numerical methods could provide estimations.

In the next section, we explore how the terms in the decomposition of the value function can be used to augment traditional benefit-cost analysis in the presence of deep uncertainty and potentially improve the efficiency of adaptive natural resource management.

4. Usefulness of Adaptive Control in Adaptive Natural Resource Management Applications

Many natural resource and environmental management issues are characterized by deep uncertainty stemming from a pure lack of information, disagreements about previously collected data (leading to disparate priors), and the fact that stochastic and complex processes are the norm, rather than the exception. Natural resource managers are faced with the challenge of making decisions in the face of this deep uncertainty, yet the standard benefit-cost paradigm offers little guidance when a) the response parameters of the system are uncertain and the system is stochastic; and perhaps more importantly, b) the allocation of scarce resources must be divided between learning about the system and managing it for a specific result. This latter tradeoff, which has, in practice, resulted in the adoption of AM practices in which experimentation is actively used to test hypotheses about natural system response, introduces a new frontier over which decisions must be made, in addition to the (more standard) complex calculus of trading off benefits and costs between the short- and long- runs.

AC techniques, and the results presented above, may help to shed light on this additional dimension. Clearly, the general model incorporates many of key aspects of AM, including stochastic dynamic processes, uncertain marginal effects of actions on the states of the system, and the need to trade off the benefits of information gathering with the (endogenous) costs of deviation from the stated objective. In particular, the potential contributions of using AC in environmental AM include:

(1) the ability to value alternative treatments of information for use in AM decision-making;

(2) the ability to cast the results of the modeling exercise in a familiar benefit-cost framework, including a decomposition of the relevant expected informational values of the system; and

(3) the ability to represent and/or simulate optimal alternative AM paths, including the "switching" points where a previously employed experiment is no longer optimal relative to another in the feasible set.

We discuss each in turn.

Valuation of alternative treatments of information

A key deficiency of traditional benefit-cost analysis in environmental decision-making is that the relationship between an action and the resultant states is either known with certainty, or that the uncertain states can be characterized by a reasonably informative distribution. As such, ex ante valuation of information collection is completely ignored using this technique. In AM, however, experimental learning is a key component of the overall management strategy, and thus the use of information becomes paramount [16].

AC offers a way to decompose the expected value of alternative types of system information into its component parts, thus providing a significant amount of insight (information) to a decision-maker. Equation (1.18) provides this decomposition of value for the optimal active learning path, which incorporates all expected forward-looking information, and provides the expected value at each physical and informational state from treating parameters as uncertain, the value of updating priors based on passive strategies, and the value of experimentation. In the presence of, for example, disparate priors amongst groups of scientists, this decomposition provides a means of illustrating the expected tradeoffs between passive and active learning (the value of optimal experimentation) under alternative assumptions, potentially leading to more efficient management decisions, or at least documenting the opportunity costs of acting under differing assumptions.

Note that the value of information can be calculated for sub-optimal paths as well, which can be especially useful if only a subset of management options are being considered. For any general decision rule $c(x,\theta)$ corresponding to a pre-defined experiment path (possibly including no experimentation whatsoever), the expected value of alternative information treatments can be calculated. For example, the expected value of anticipatory information updating can be estimated by finding an approximation to the value function

$$\begin{aligned} V\ & (x_t,\theta_t) = u\left(c\left(x_t,\theta_t\right),x_t\right) + \\ & \beta\int_{\mathbb{R}^K}\int_{\mathbb{R}^N} V\left(f\left(c\left(x_t,\theta_t\right),x_t;\gamma,\alpha\right)+\varepsilon_{t+1}, G\left(c(x_t,\theta_t),x_{t+1},x_t,\theta_t\right)\right)p(\alpha\,|\,\theta_t)q(\varepsilon)d\alpha d\varepsilon. \end{aligned} \tag{1.19}$$

under the two conditions $G\left(c(x),x_{t+1},x_t,\theta_t\right)=\theta_t$ and (1.2), then taking the difference. Similar calculations could be performed for the other terms if needed. In a political context where objectives may differ across individuals or groups, such calculations can provide a means of objective valuation of differing management approaches (e.g., ex ante plans that maximize the perceived information collection, minimize disruptions to the natural ecosystem, etc…), and in so doing potentially improve management decisions.

Benefit-cost Analysis and Adaptive Control

Closely related to the value of alternative informational treatments is the ability of researchers to use AC methods to consistently augment traditional benefit-cost analysis to include informational aspects, especially in the presence of deep uncertainties. Although AC methods are complex, the resultant information can be placed into the more familiar B/C framework to help inform policy-makers about the magnitudes and tradeoffs related to alternative information treatments.

To illustrate, assume that instantaneous utility in each period can be decomposed into (presumably additively separable) benefit and cost components, or $u(c_t, x_t) = B(c_t, x_t) - C(c_t, x_t)$, where the first term represents the benefits and the second the costs of an action and state vector. Using the now familiar notation, a number of alternative benefit-cost streams can be defined. For example, the certainty-equivalent streams from the current period through *T* are defined as $\sum_{t=0}^{T} \beta^t \left[B\left(c^{ce}(x_t^{ce}), x_t^{ce}\right) \right]$ and $\sum_{t=0}^{T} \beta^t \left[C\left(c^{ce}(x_t^{ce}), x_t^{ce}\right) \right]$, and subtracting the cost term from the benefit term yields the second term in (1.12), or the realized discounted value of following the certainty-equivalent path. This is akin to calculation of the traditional benefit-cost analysis when the resultant state vectors and responses are assumed known (i.e., no parameter uncertainty).

Similar expressions can be developed for the stochastic parameter, passive learning, and active learning cases, and the differences between each can be the expected gain or loss in benefits or costs from alternative informational treatments. For example, experimentation may be valuable ex ante, but due to reduced future costs rather than impacts on future benefits. Formally, this situation could be described as:

$$\Delta B = \sum_{t=0}^{T} \beta^t \left[B\left(c^{al}(x_t^{al}, \theta_t^{al}), x_t^{al}, \theta_t^{al}\right) - B\left(c^{pl,\theta}(x_t^{pl,\theta}), x_t^{pl,\theta}, \theta_t^{pl,\theta}\right) \right] < 0$$

and $\Delta B - \sum_{t=0}^{T} \beta^t \left[C\left(c^{al}(x_t^{al}, \theta_t^{al}), x_t^{al}, \theta_t^{al}\right) - C\left(c^{pl,\theta}(x_t^{pl,\theta}), x_t^{pl,\theta}, \theta_t^{pl,\theta}\right) \right] > 0$, with this last term similar to (1.15). Thus, the researcher or manager can document the effect of alternative information treatments on both the benefit and cost side, allowing for a natural extension of B/C analysis into the realm of deeply uncertain information. Of course, sub-optimal management plans can be decomposed this way as well.

Simulation of optimal adaptive management paths

Finally, by providing a framework in which optimal decision rules that incorporate the trade offs of management objectives with learning can be derived, the AC framework is well-suited for simulation analysis related to adaptive environmental management. As such, given alternative future simulated stochastic realities, entire AM plans can be developed and/or evaluated for planning and other purposes, as opposed to the planning of just one experiment in the immediate future.

Similar to the "information exploitation" of [25] and the agent-based modeling approach described in [16], researchers and managers have the ability to use the AC model as a virtual lab by exploring the implications of certain future outcomes on the optimal adaptive path, and planning accordingly. In particular, this could conceivably include the circumstances under which it is optimal to stop a given experiment and start another, even if the full resolution of the uncertainty has not yet occurred. In other words, it may not be optimal to completely resolve the uncertainty in all parameters if the underlying endogenous opportunity cost of experimentation is too great. Unfortunately, as discussed in the next section, there may be occasions where the "optimal" solution leads to convergence on incorrect parameters.

4. Limitations of Adaptive Control in Environmental Adaptive Management Applications

As with all modeling techniques, the use of AC to help inform adaptive environmental management has several major limitations, several of which are well-documented in the dual-control and Bayesian learning literature. First, it is possible that limit beliefs (those that predominate asymptotically) do not converge to the true parameter values in certain cases (see [26, 27, and 19, plus references therein]. However, Wieland [19] finds that in a linear regression framework (in which the transition function is essentially a regression equation with unknown parameters), there are discontinuities in the optimal policy function (and a value function which is potentially non-differentiable) at the incorrect limit beliefs, suggesting a large degree of optimal experimentation at those points in the state space. Further research is necessary to determine if this is a general result, or one specific to the linear regression context or parameter values chosen.

Second, the curse of dimensionality may be a very real problem in the AM context. Assuming N uncertain parameters characterized by a mean, variance, and covariance matrix, this results in $2N + \frac{N(N-1)}{2}$ unique information parameters in θ, independent of the number of state equations. In many cases, it is possible that this number can be reduced depending on the assumptions made by the modeler; however, in many environmental applications, even this reduced number may render numerical approximation infeasible given the current state of technology.

As such, it may be that AC methods are most useful providing information for a separable subset of the large-scale environmental problem under consideration. For example, while the restoration projects of the Everglades and Grand Canyon would involve hundreds if not thousands of uncertain parameters when taken as a whole, an AC application that helps determine the optimal strategy to increase the humpback chub population in a certain section of river may be more tractable (say, one or two uncertain parameters).

Third, environmental systems can be characterized by complex, non-linear dynamics, thresholds, and irreversibilities that may not be known to the researcher or manager a priori.

Indeed, this fact is a likely driver of the adoption of AM practices in practice. However, to the author's knowledge, there has been no research that compares the performance of active learning control strategies with data generating processes that do not match those assumed by the model. In some respect, linear or quadratic state equations can be considered first- or second-order approximations to this true process, but still may not capture all of the ecological processes necessary to make a truly "optimal" decision. One promising option might be to use orthogonal polynomial forms, such as Chebychev polynomials, to model state dynamics. In any case, more research is needed to confirm the usefulness of AC under circumstances where the true and assumed data generating processes differ.

Finally, as no analytical solutions are likely for even simple AM/AC problems, approximate solutions must be found numerically. Dual control algorithms use Monte-Carlo simulation analysis using a Taylor-series expansion of the future value function, while functional approximation and grid-space methods are typically used for smaller numeric dynamic programming applications [19, 28, 29]. As noted by Wieland [19], these approaches are likely complementary, with each having its own advantages and disadvantages.

Nevertheless, there are some specific challenges of the AC methodology. The possible existence of non-convexities, non-continuities, or non-differentiable value functions, mentioned previously, may result in numerical difficulties with gradient search methods, and the exponentially-increasing state space creates problems relating to solution time. As computing power advances and more research is completed regarding computational economics and methods, however, these problems may become less costly. As with all numerical analyses, specific functional forms and known parameter values must be assumed, as well as the distributional form of the unknown parameters, limiting the ability to generalize results across applications.

5. Conclusions

Experiment-based adaptive natural resource management is being increasingly adopted by a number of public agencies and applied to a number of high-profile, large-scale environmental management projects, including Everglades restoration and endangered species management in the Grand Canyon. However, the natural resource and environmental economics profession has not been active in providing appropriate tools for ex ante and/or ex post evaluation of AM strategies, which in part has resulted in exclusion from these important management issues.

This paper argues that AC techniques have considerable potential to provide theoretically-consistent evaluation of AM strategies. Although originally developed for use in models of the macro economy, the incorporation of the notion of learning about structural uncertainties of a complex system renders AC well-suited for application to many natural resource management problems. More specifically, we illustrate that alternative treatments of information can be valued using a dynamic programming framework, and these values can be used to augment traditional benefit-cost analysis to include informational aspects of the management problem. Thus, a key limitation of standard analysis is overcome. While this approach is not without its limitations, the use of AC could help managers prioritize experiments, plan AM programs, simulate potential AM paths, and justify decisions based on an objective valuation framework.

References

1. Milon, W., Kiker, C., and Lee, D. Adaptive ecosystem management and the Florida Everglades: More than trail-and-error? *Water Resources Update* 1998; **113**: 37-46.

2. Loomis, J., Douglas, A., and Harpman, D. Recreation use values and nonuse values of Glen and Grand Canyons. In S. Gloss, J. Lovich and T. Melis, *The State of the Colorado River Ecosystem in Grand Canyon: A Report of the Grand Canyon Monitoring and Research Center 1991-2004*, USGS Circular 1282, Reston, VA., 2005.

3. Walters, C. *Adaptive Management of Renewable Resources*. McGraw Hill, New York, 1986.

4. Thrower, A and Martinez, J.M. Reconciling Anthropocentrisim and Biocentrism through Adaptive Management. *Journal of Environment and Development* 2000; **9**(10): 68-97.

5. Holling, C. and Meffe, G. Command and control and the pathology of natural resource management. *Conservation Biology* 1996; **10**(20): 328-337.

6. Bar-Shalom, Y. and Tse, E. Caution, probing, and the value of information in the control of uncertain systems. *Annals of Economic and Social Measurement* 1976; **5**(3): 323-337.

7. Kendrick, D.A. *Stochastic Control for Economic Models.* McGraw-Hill, Inc., New York, 1981.

8. Kendrick, D.A. Stochastic control for economic models: Past, present, and the paths ahead. J*ournal of Economic Dynamics and Control* 2005; **29**: 3-30.

9. Mercado, P.R. and Kendrick, D.A. Parameter uncertainty and policy intensity: Some extensions and suggestions for further work. *Computational Economics* 2006; **27**: 483-496.

10. Kaplan, J.D., Howitt, R.E., and Farzin, Y.H. An information-theoretical analysis of budget-constrained nonpoint source pollution control. *Journal of Environmental Economics and Management* 2003; **46**: 106-130.

11. Kelly, D.L. and Kolstad, C.D. Bayesian learning, growth, and pollution. *Journal of Economic Dynamics and Control* 1999; **23**: 491-518.

12. Peterson, G.D., Carpenter, S.R. and Brock, W.A. Uncertainty and the management of multistate ecosystems: An apparently rational route to collapse. *Ecology* 2003; **84**(6): 1403-1411.

13. Cunha-e-Sá, M.A. and Santos, V. Experimentation with accumulation. *Journal of Economic Dynamics and Control* 2008; **32**: 470-496.

14. Springborn, M. 2008. Bayesian adaptive management with optimal learning. Working paper presented at Colorado State University, January 28, 2008.

15. Brock, W.A. and Carpenter, S.R. Panaceas and diversification of environmental policy. *Proceedings of the National Academy of Sciences* 2007; **104**(39): 15206-15211.

16. Loomis, J., Bond, C.A., and Harpman, D. The potential of agent based modeling for evaluating the economic efficiency of adaptive natural resource management. Working Paper, Colorado State University, 2008.

17. Bureau of Reclamation. *Operation of Glen Canyon Dam Final Environmental Impact Statement*. Bureau of Reclamation, Salt Lake City, Utah, 1995.

18. Prato, A. Adaptive Management of Large Rivers with Special Reference to the Missouri River. *Journal of the American Water Resources Association* 2003; **39**: 935-946.

19. Wieland, V. Learning by doing and the value of optimal experimentation. *Journal of Economic Dynamics and Control* 2000; **24**: 501-534.

20. Kendrick, D.A. Non-convexities from probing in an adaptive control problem. *Economic Letters* 1978; **1**: 347-351.

21. Norman, A., Norman, M., and Palash, C. Multiple relative maxima in optimal macroeconomic policy: An illustration. *Southern Economic Journal* 1979; 46: 274-279.

22. Mizrach, B. Non-convexities in a stochastic control problem with learning. *Journal of Economic Dynamics and Control* 1991; **15**(3): 515-538.

23. Amman, H.M. and Kendrick, D.A. Nonconvexities in stochastic control models: An analysis. In W.W. Cooper and A.B. Winston (eds.), *New Directions in Computational Economics*. Kluwer, Dordrecht, 1994.

24. Tucci, M.P. The nonconvexities problem in adaptive control models: A simple computational solution. *Computational Economics* 1998; **12**: 203-222.

25. Tesfatsion, L. Agent based computational economics: A constructive approach to economic theory. In L. Tesfatsion and K. Judd, editors, *Handbook of Computational Economics, volume 2, Agent Based Computational Economics.* Elsevier, Amsterdam, 2006.

26. Easley, D. and Kiefer, N. Controlling a stochastic process with unknown parameters. *Econometrica* 1988; **56**: 1045-1064.

27. Kiefer, N. and Nyarko, Y. Optimal control of an unknown linear process with learning. *International Economic Review* 1989; **30**: 571-586.

28. Judd, K.L. *Numerical Methods in Economics*. MIT Press: Cambridge, MA; 1998.

29. Miranda, M.J. and Fackler, P.L. *Applied Computational Economics and Finance*. MIT Press : Cambridge, MA; 2002.

Appendix A

Assume a state equation of the form $x_{t+1} = \alpha_0 + \alpha_1 x_t + \gamma c_t + \varepsilon_{t+1}$, where α_1 and α_2 are uncertain. Assume a discrete probability distribution of the form $P\left[\alpha_j = \alpha_{ij}\right] = \pi_{ij},\ \sum_{\mathrm{i}=1}^{\mathrm{Q}} \pi_{\mathrm{ij}} = 1,\ j \in (0,1)$, for the Q support values α_{ij} of the uncertain parameter indexed by j. Using Bayesian updating, the posterior is obtained by

$$G_{ij}(\cdot) = \frac{\pi_{ij} f\left(x_{t+1} - \alpha_{i0} - \alpha_{i1} x_t - \gamma c_t\right)}{\sum_{k=1}^{Q} \pi_{kj} f\left(x_{t+1} - \alpha_{k0} - \alpha_{k1} x_t - \gamma c_t\right)},\ i = 1,2,\ldots,Q,\ j = 0,1.$$

Calculating the elements of the term $\left(\frac{\partial G}{\partial c} + \frac{\partial G}{\partial x_{t+1}} \frac{\partial f}{\partial c}\right)$ from (1.7), obtain

$$\frac{\partial G_{ij}(\cdot)}{\partial c} = \frac{-\gamma \pi_{ij} f_\varepsilon(\cdot) \sum_{k=1}^{Q} \pi_{kj} f(\cdot) + \gamma \pi_{ij} f(\cdot) \sum_{k=1}^{Q} \pi_{kj} f_\varepsilon(\cdot)}{\left(\sum_{k=1}^{Q} \pi_{kj} f(\cdot)\right)^2}, i = 1,2,\ldots,Q,\ j = 0,1,$$

where $f(\cdot)$ and $f_\varepsilon(\cdot)$ are the probability distribution function and marginal of the white noise stochastic shock evaluated at $x_{t+1} - \alpha_{k0} - \alpha_{k1} x_t - \gamma c_t$, and

$$\frac{\partial G_{ij}(\cdot)}{\partial x_{t+1}} = \frac{\pi_{ij} f_\varepsilon(\cdot) \sum_{k=1}^{Q} \pi_{kj} f(\cdot) - \pi_{ij} f(\cdot) \sum_{k=1}^{Q} \pi_{kj} f_\varepsilon(\cdot)}{\left(\sum_{k=1}^{Q} \pi_{kj} f(\cdot)\right)^2}, i = 1,2,\ldots,Q,\ j = 0,1.$$

As such, $\frac{\partial G_{ij}(\cdot)}{\partial x_{t+1}} = -\frac{\partial G_{ij}(\cdot)}{\partial c} \gamma^{-1}\ \forall i, j,$ and since $\frac{\partial f(\cdot)}{\partial c} = \gamma$, $\left(\frac{\partial G}{\partial c} + \frac{\partial G}{\partial x_{t+1}} \frac{\partial f}{\partial c}\right) = 0.$ Note this result would not hold if γ was uncertain.