18th World IMACS / MODSIM Congress, Cairns, Australia 13-17 July 2009
http://mssanz.org.au/modsim09

Knowledge gain or system benefit in environmental decision making?

W. J. M. Probert [1,2,3], E. McDonald-Madden [3], C. E. Hauser [4], P. W. J. Baxter [3,4] and H. P. Possingham [1,3]

[1] *Department of Mathematics, The University of Queensland, St Lucia QLD 4072 Australia*
[2] *Email: wprobert@maths.uq.edu.au*
[3] *Centre for Applied Environmental Decision Analysis, School of Biological Sciences, The University of Queensland, St Lucia QLD 4072 Australia*
[4] *Australian Centre of Excellence for Risk Analysis, School of Botany, University of Melbourne VIC 3010 Australia*

Abstract: The quality of environmental decisions are gauged according to the management objectives of a conservation project. Management objectives are generally about maximising some quantifiable measure of system benefit, for instance population growth rate. They can also be defined in terms of learning about the system in question, in such a case actions would be chosen that maximise knowledge gain, for instance in experimental management sites.

Learning about a system can also take place when managing practically. The adaptive management framework (Walters 1986) formally acknowledges this fact by evaluating learning in terms of how it will improve management of the system and therefore future system benefit. This is taken into account when ranking actions using stochastic dynamic programming (SDP). However, the benefits of any management action lie on a spectrum from pure system benefit, when there is nothing to be learned about the system, to pure knowledge gain. The current adaptive management framework does not permit management objectives to evaluate actions over the full range of this spectrum. By evaluating knowledge gain in units distinct to future system benefit this whole spectrum of management objectives can be unlocked.

This paper outlines six decision making policies that differ across the spectrum of pure system benefit through to pure learning. The extensions to adaptive management presented allow specification of the relative importance of learning compared to system benefit in management objectives. Such an extension means practitioners can be more specific in the construction of conservation project objectives and be able to create policies for experimental management sites in the same framework as practical management sites.

Keywords: *adaptive management; conservation biology; decision theory; optimal monitoring; optimization; stochastic dynamic programming*

1. INTRODUCTION

Environmental decision making in conservation biology is concerned with figuring out how to effectively monitor and manage ecological systems. Our ability to make good decisions is limited by money, time and knowledge. Inadequate conservation funds mean resources are stretched thinly and short funding periods are unsympathetic to the effective study of gradually changing ecological processes (James, Gaston & Balmford 2001, Willis & Birks 2006). Uncertainty may be present in a variety of ways: natural variation of the system, imperfect monitoring of the system, limitations of management on the system (partial controlability) and inadequate mathematical models of the system (structural uncertainty) (Williams 1982). We may reduce uncertainty using monitoring alone or using monitoring in combination with management.

Monitoring informs us of three things: (1) our performance with respect to our management objectives, (2) the state of our system so that we can decide our next management action, and (3) the relative credibility of our models of system function (Nichols & Williams 2006, Field, O'Connor, Tyre & Possingham 2007). Monitoring is therefore an essential part of any decision making process where the efficacy of an action depends on the current state of the system.

The interconnected relationship between management action and monitoring within any conservation project has been highlighted since the introduction of the adaptive management framework in the 1980s (Walters 1986). The adaptive management framework recognises that management action, in combination with a monitoring system, can be a catalyst in reducing uncertainty by acting deliberately to solicit particular information from a system. The framework includes a management objective, a plan for monitoring the system and a regime to implement management action (Parma & the NCEAS Working Group on Population Management 1998, Walters 1986). Management objectives, defined by practitioners of adaptive management, determine what is seen as a 'good' decision and therefore how actions are mathematically evaluated and ranked against other competing actions. In line with management objectives of maximising system benefit, actions in an adaptive management framework are evaluated in terms of system benefit. The value of learning about the system associated with each action is summarised in the same units of system benefit by calculating the expected future system benefit as a result of the improved knowledge of the system (and thus improved future management) (D'Evelyn, Tarui, Burnett & Roumasset 2008).

Although the adaptive management framework acknowledges that *any* management action may simultaneously benefit the system *and* improve knowledge about the system, the capacity of the current framework to explore the full spectrum, from pure knowledge gain through to pure system benefit, is limited. By valuing knowledge gain in units distinct from future improvement of system benefit the algorithms commonly used to create decision policies in adaptive management projects can be expanded to fully explore this spectrum.

We describe six decision making policies and their respective mathematical algorithms defined across the whole range of this spectrum of management objectives. The simplest policy described is that which does not adapt management action to new information about the system and seeks to maximise system benefit over the next time step. Building on this, the second policy looks to maximise population growth rate over the whole length of the conservation project and uses monitoring information to update knowledge about the system. Two popular types of adaptive management that acknowledge multiple models of system function are outlined: passive and active adaptive management. Finally, two algorithms that explicitly value knowledge gain, one that only values knowledge and one that allows specification of the relative importance of both learning and system benefit in the management objective. Using the same mathematical framework, and as each presented algorithm builds upon the previous, the differences between the algorithms are highlighted.

2. METHODS

Uncertainty in how the system functions is described in I different models. Our confidence in each model is summarized in a belief weight, w_i which is defined as the probability we believe model i to be true. Thus, $\sum_{i=1}^{I} w_i = 1$. At each time step, t, of a conservation project of T years we obtain monitoring information about the system, update our belief about our models and implement an action, a, from a pool of m different actions. In this paper population growth rate has been chosen as the measure of system benefit although other quantifiable measures may be chosen. The geometric mean population growth rate over T years is

$$G = \left(\Pi_{t=1}^{T} \lambda_t\right)^{\frac{1}{T}} \quad \text{and} \quad \ln G = \frac{1}{T}\Sigma_{t=1}^{T} \ln\left(\lambda_t\right),$$

where λ_t denotes the population growth rate at time t. We assume the population growth rates are uncorrelated over time and so we can collapse the array of mean geometric growth rates to essential information at each time step:

$$L_t = \Sigma_{l=1}^{t} \ln \lambda_l.$$

The state of the system is a triplet $(\mathbf{w}_t, L_t, t)$: a vector of the current belief weights, $\mathbf{w}_t$, the mean geometric growth rate so far, L_t, and the number of time steps elapsed, t. Actions are chosen from a pool of J possible actions according to the current state of the system and a utility function, $V_j(\mathbf{w}_t, L_t, t)$, $j \in \{1, 2, \cdots, J\}$ defined for each decision making policy. For each state, the optimal action and corresponding optimal utility are denoted with an asterisk, a^* and $V^*(\mathbf{w}_t, L_t, t)$ respectively.

2.1 Non Adaptive Management (NAM)

At the beginning of the management project choose the model that has the highest associated belief weight and at all time steps implement the action that will maximise current system benefit under that model. This policy does not adapt management strategy according to new information that becomes available at each time step. The inclusion of NAM illustrates a policy that has been implemented in the field of conservation in the same Markov decision process (MDP) framework as other proposed policies.

Algorithm. At the beginning of the project choose the most likely model

$$i^M = \arg\max_i \mathbf{w}_0. \tag{1}$$

Choose an action, a^*, that maximises population growth rate according to model i^M and apply this same action every year. Since each λ_t is independent of future growth rates and actions through time, this action will maximise $\ln \lambda_t$ at each time t during the project. The utility function and optimal action to be applied every year are respectively

$$V_j(\mathbf{w}_t, L_t, t) = \int_\lambda \ln \lambda \quad f(\lambda \mid a_j,\ \text{model } i^M) d\lambda, \tag{2}$$

$$a^* = \arg\max_{a_j} V_j(\mathbf{w}_t, L_t, t) \qquad \forall t = 1, \cdots, T, \tag{3}$$

where $f(\lambda \mid a_j,\ \text{model } i^M)$ is the likelihood of observing a particular growth rate, λ, at time t, given model i^M is true and action a_j is taken.

2.2 Myopic Adaptive Management (MAM)

At each time step choose the action that will maximise expected utility in the final time step assuming that optimal actions are taken at all future time steps. To do so, choose the model that has the highest probability of being true and implement the action that will maximise system benefit at the end of the project assuming this model is true. To calculate expected utility, we need to know the optimal action and corresponding maximum utility from the next time step, so this algorithm is initialised by calculating the utility of all system states in the final time step and working backwards iteratively. MAM looks at the utility of actions over the whole scope of the project, rather than just in the next time step. MAM, however, does not anticipate future learning specifically within the utility function. This strategy is identical to choosing a non-adaptive management strategy at each time step instead of only in the first time step. This strategy is *myopic*, or short-sighted, with respect to the models of system function being considered, in the sense that MAM only considers the consequences of an action under one model of system function.

Algorithm. This algorithm uses stochastic dynamic programming (SDP) as follows (Bellman 1957). Terminal optimal utility is calculated as the final mean population growth rate

$$V^*(\mathbf{w}_T, L_T, T) = \exp\left(\frac{L_T}{T}\right) = G. \tag{4}$$

Utility at time step t and corresponding optimal action are respectively

$$V_j(\mathbf{w}_t, L_t, t) = \int_{\lambda_{t+1}} V^*(\mathbf{w}_t, L_{t+1}, t+1) f(\lambda_{t+1} \mid a_{jt},\ \text{model } i_t^M) d\lambda_{t+1} \tag{5}$$

$$a_t^* = \arg\max_{a_j} V_j(\mathbf{w}_t, L_t, t), \tag{6}$$

where model i_t^M is the model with the highest belief weight at time t,

$$i_t^M = \arg\max_i \mathbf{w}_t, \tag{7}$$

and using transitions $L_{t+1} = L_t + \ln \lambda_{t+1}$. In forward simulations, managers update the set of belief weights using Bayes' theorem to passively learn

$$w_{i,t+1} = \frac{\mathrm{P}(\hat{\lambda}_t \mid \text{model i}, a_t^*) w_{it}}{\sum_{k=1}^{I} \mathrm{P}(\hat{\lambda}_t \mid \text{model k}, a_t^*) w_{kt}} \tag{8}$$

where $\hat{\lambda}_t$ is the observed growth rate and a_t^* is the optimal management action taken at time t.

2.3 Passive Adaptive Management (PAM)

Within PAM the expected utility of an action is averaged over the whole range of possible growth rates we might observe in the next time step and over all competing models given the current belief in each being true (the latter point is what differentiates between the PAM and MAM algorithms). As this policy considers the benefits of each action under each model (and weighted according to the belief in each model), the prioritisation is ultimately with respect to actions rather than models, which is in contrast to NAM and MAM (where actions are ranked assuming one model is true, disregarding the others).

Algorithm. Terminal utility is calculated according to equation 4. At time t utility and optimal actions are calculated according to

$$V_j(\mathbf{w}_t, L_t, t) = \sum_{i=1}^{I} w_{it} \int_{\lambda_{t+1}} V^*(\mathbf{w}_t, L_{t+1}, t+1) f(\lambda_{t+1} \mid a_j,\ \text{model } i) d\lambda_{t+1}, \tag{9}$$

$$a_t^* = \arg\max_{a_j} V_j(\mathbf{w}_t, L_t, t), \tag{10}$$

with transitions $L_{t+1} = L_t + \ln \lambda_{t+1}$.

2.4 Active Adaptive Management (AAM)

AAM is the mathematically optimal strategy for solving Markov decision processes with a knowledge state. AAM goes beyond PAM by considering how an action helps to distinguish between competing models of system function and how this ability to discriminate will be beneficial for the remainder of the conservation project. AAM mathematically recognises that increased understanding of the system can mean a more efficiently managed system. AAM assumes that managers will use Bayes' theorem to update belief weights in forward simulations and thus adjusts the future belief weights in the utility calculation accordingly. Learning is *active* as updating of the belief weights is built into the algorithm. AAM looks at the utility of actions over the whole scope of the project, rather than just in the next time step, therefore actions may be chosen that seem to provide sub-optimal system benefit in the next time step but will provide benefit in the long term through ability to distinguish between competing models.

Algorithm. This algorithm also uses SDP. Terminal utility is calculated according to equation 4. At time t utility and optimal actions are respectively calculated according to

$$V_j(\mathbf{w}_t, L_t, t) = \sum_{i=1}^{I} w_{it} \int_{\lambda_{t+1}} V^*(\mathbf{w}_{t+1}, L_{t+1}, t+1) f(\lambda_{t+1} \mid a_j,\ \text{model } i) d\lambda_{t+1} \tag{11}$$

$$a_t^* = \arg\max_{a_j} V_j(\mathbf{w}_t, L_t, t), \tag{12}$$

with transitions $L_{t+1} = L_t + \ln \lambda_{t+1}$ and belief weights updated within the utility function using equation 8. In forward simulations, managers update the set of belief weights also using Bayes' theorem and equation 8.

2.5 Pure Learning

At each time step choose the action that will maximise the expected ability to determine the true model at the conclusion of the project, with no regard for system benefit.

Algorithm. This algorithm uses SDP in the same way as active adaptive management although the utility function in the final time step quantifies learning instead of system benefit. Inability to determine the true model is quantified using negative *information entropy* (Shannon 1948). Negative information entropy is a function of the belief weights and is largest for belief states when one model has a weight of 1 and the other models have a weight of 0 (*i.e.* when there is certainty in only one model being true) and is smallest when all models have equal weight (*i.e.* when there is maximum uncertainty in which model is true). At the final time step the utility function is calculated as the negative information entropy of the belief weights

$$V^*(\mathbf{w}_T, L_T, T) = -\sum_{i=1}^{I} w_{i,T} \ln w_{i,T} \quad > 0 \quad \forall \; L_T. \tag{13}$$

Entropy measures disorder, thus we want to impose order on our knowledge by maximising this expression (*i.e.* large negative entropy). At all previous time steps, $t = 0, 1, \ldots, T-1$ the utility and optimal action, a_t^* are respectively calculated using equations (11) and (12).

Belief weights are updated according to transition equation 8 and the mean geometric growth rate is updated using the transitions $L_{t+1} = L_t + \ln \lambda_{t+1}$.

2.6 Experimental Learning

The extensions that we present allow expressions for a weighted mixture of pure learning and system benefit in action evaluation. A weighting parameter, γ, is introduced to allow manipulation of the relative importance of learning versus system benefit in our algorithm. As with the AAM and pure learning algorithms, utility is initialised at the final time step and the algorithm iterates backward through time. Actions are thus chosen such that utility is maximised at the end of the project. When $\gamma = 0$ this strategy is identical to pure learning, when $\gamma = 1$ this strategy is identical to active adaptive management, between these two limits this strategy is called *experimental learning*.

Algorithm. This algorithm also uses SDP similar to both the active adaptive management and pure learning algorithms, although the utility in the final time step is a weighted combination of these other two utility functions, the mean geometric growth rate and the information entropy. At the final time step the utility function is

$$V^*(\mathbf{w}_T, L_T, T) = \gamma\left(\exp\left(\frac{L_T}{T}\right)\right) + (1-\gamma)\left(-\sum_{i=1}^{I} w_{i,T} \ln w_{i,T}\right) \tag{14}$$

At all previous time steps, $t = 1, 2, \ldots, T-1$, the maximum utility and associated optimal action, a_t^*, are respectively calculated using equations (11) and (12).

Belief weights are updated according to transition equation 8 and the mean geometric growth rate is updated using transitions $L_{t+1} = L_t + \ln \lambda_{t+1}$.

3. DISCUSSION

The differences between the policies can be summarised in a few key criteria (Table 1): whether monitoring information updates knowledge through time; whether the policy assumes one model is true at each time step or weights each model according to their respective belief weight; whether the policy acknowledges that a new observation will change the current knowledge about the system; how improved model discrimination is valued; if the relative value of learning versus system benefit can be specified.

The extensions to the adaptive management framework presented provide a tool for specifying the importance of learning relative to system benefit within a decision making framework common to conservation. The extensions above allow decision policies for experimental sites to be created, for example, where learning is the only objective, using the same adaptive management framework that is used for practical sites, where benefits such as high growth rates are desired, therefore reducing the knowledge load for practioners and unifying two distinct areas of decision making in ecology. Only by quantifying knowledge can actions be chosen to maximise learning.

The policies outlined were only concerned with conservation projects of finite time horizons, as such are sympathetic to the finite funding arrangements most common in conservation projects and since system

Criterion | Policy	NAM	MAM	PAM	AAM	PL	EL
Update knowledge of system function through time?	N	Y	Y	Y	Y	Y
Acknowledge competing models of system function in utility calculation?	N	N	Y	Y	Y	Y
Update knowledge within utility function?	N	N	N	Y	Y	Y
How is knowledge evaluated?	NA	NA	Future system benefit	Future system benefit	Model discrimination	Future system benefit *and* model discrimination, weighted by γ
Allow specification of relative importance of learning versus system benefit?	N	N	N	N	N	Y

Table 1. Key differences between the six decision making policies

function may change over a long period of time (Hauser & Possingham 2008). We presented policies for choosing an optimal action in only one site and ignore costs of actions. The policies presented here can be extended to multiple sites (for example, a metapopulation) by considering a set of J possible actions for all possible action/site combinations and integrating over all possible combinations of observed growth rates in the multiple sites. We have not presented a mechanism in the proposed algorithms to allow varying monitoring effort and have assume that practitioners of the proposed decision making tools are competent with mathematical techniques, such as using Bayes' theorem. No guidelines are given for the choice of proposed models of system function or list of proposed actions.

Future work includes simulating the management of a conservation project using these different policies and comparing the efficacy of each. Such investigations will further highlight the benefits of explicit conservation objectives and the operational considerations required for the outlined decision making policies.

REFERENCES

Bellman, R. (1957), *Dynamic Programming*, Princeton University Press, USA.

D'Evelyn, S. T., N. Tarui, K. Burnett & J. A. Roumasset (2008), 'Learning-by-catching: Uncertain invasive-species populations and the value of information', *Journal of Environmental Management* **89**(4), 284 – 292. Economics of Invasive Species Management.

Field, S.A., P.J. O'Connor, A.J. Tyre & H.P. Possingham (2007), 'Making monitoring meaningful', *Austral Ecology* **32**, 485–491.

Hauser, C.E. & H.P. Possingham (2008), 'Experimental or precautionary? adaptive management over a range of time horizons', *Journal of Applied Ecology* **45**(1), 72–81.

James, A., K.J. Gaston & A. Balmford (2001), 'Can we afford to conserve biodiversity?', *BioScience* **51**(1), 43–52.

Nichols, J.D. & B.K. Williams (2006), 'Monitoring for conservation', *Trends in Ecology and Evolution* **21**(12), 668–673.

Parma, A.M. & the NCEAS Working Group on Population Management (1998), 'What can adaptive management do for our fish, forests, food and biodiversity?', *Integrative Biology, Issues, News and Reviews* **1**, 16–26.

Shannon, C. E. (1948), 'A mathematical theory of communication', *The Bell System Technical Journal* **27**, 379–423, 623–656.

Walters, C.J. (1986), *Adaptive management of renewable resources*, Blackburn Press, Caldwell, New Jersey, USA.

Williams, Byron K. (1982), ‘Optimal stochastic control in natural resource management: framework and examples’, *Ecological Modelling* **16**(2-4), 275 – 297.

Willis, K. J. & H. J. B. Birks (2006), ‘What is natural? the need for a long-term perspective in biodiversity conservation’, *Science* **314**(5803), 1261–1265.