Journal of Environmental Economics and Management 66 (2013) 609–624

Contents lists available at ScienceDirect

Journal of Environmental Economics and Management

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

A density projection approach for non-trivial information dynamics: Adaptive management of stochastic natural resources

Michael Springborn [a,*], James N. Sanchirico [b]

[a] Department of Environmental Science and Policy, University of California, Davis, United States
[b] Department of Environmental Science and Policy, University of California, Davis and University Fellow, Resources for the Future, Washington DC, United States

ARTICLE INFO

Article history:
Received 4 December 2012
Available online 17 July 2013

Keywords:
Dynamic programming
Bayesian learning
Adaptive management
Bioeconomic
Decision-making under uncertainty
Partially-observed Markov decision process

ABSTRACT

We demonstrate a density projection approximation method for solving resource management problems with imperfect state information. The method expands the set of partially-observed Markov decision process (POMDP) problems that can be solved with standard dynamic programming tools by addressing dimensionality problems in the decision maker's belief state. Density projection is suitable for uncertainty over both physical states (e.g. resource stock) and process structure (e.g. biophysical parameters). We apply the method to an adaptive management problem under structural uncertainty in which a fishery manager's harvest policy affects both the stock of fish and the belief state about the process governing reproduction. We solve for the optimal endogenous learning policy—the active adaptive management approach—and compare it to passive learning and non-learning strategies. We demonstrate how learning improves efficiency but typically follows a period of costly short-run investment.

1. Introduction

The traditional method of attempting to resolve structural uncertainty on ecosystem dynamics is to invest in natural science research (Moore and McCarthy, 2010). Because management decisions are made on a repeated basis, however, there is also the possibility of using these actions to learn about the ecosystem, i.e. to practice adaptive management. In one of the foundational works on adaptive management, biologist Carl Walters argued that the primary means of reducing uncertainty in environmental models is "through experience with management itself rather than through basic research or the development of (theory)" (Walters, 1986).

Optimal investment in learning depends on the expected value of information generated (Hartmann et al., 2007; Springborn et al., 2010). All else equal, the faster a resource manager resolves structural uncertainty, the sooner the manager is able to reduce management error resulting from the difference between the true structure of a system and the manager's beliefs about that structure. However, accelerating the process of learning, relative to the optimal non-learning policy, comes with an opportunity cost, in the form of short run losses due to learning-driven deviations. An optimal endogenous

* Corresponding author. Fax. +1 530 752 3350.
E-mail addresses: mspringborn@ucdavis.edu (M. Springborn), jsanchirico@ucdavis.edu (J.N. Sanchirico).

0095-0696/$ - see front matter
http://dx.doi.org/10.1016/j.jeem.2013.07.003

learning approach—sometimes called active adaptive management (AAM)—involves balancing these tradeoffs by pursuing informative actions, but only when the investment is justified by the value of reducing uncertainty.

We employ an adaptive control framework to model a fishery manager choosing the optimal harvest in each period subject to both irreducible uncertainty (stochasticity) and reducible (structural) uncertainty in the population dynamics as a partially observed Markov decision process (POMDP) (see also Zhou et al., 2010; McDonald-Madden et al., 2010, and Bond and Loomis, 2009).[1] Learning to reduce structural uncertainty is accomplished by observing the population, which a manager can influence by altering harvest policy.[2] A POMDP is a sequential decision problem in which at least one underlying state is not known with certainty. Using the concept of a *belief state*, i.e. encoding the decision-maker's probabilistic beliefs about the uncertain state, the POMDP can be converted to a continuous-state Markov decision process (MDP), or *belief* MDP (Bertsekas, 1995).

Our paper makes three important contributions to the growing literature on adaptive control in natural resource economics. First, we use a hierarchical uncertainty model to flexibly represent uncertainty. Second, we relax constraints imposed in the existing AAM literature and much of the economics literature that limit the action space or belief space to a few discrete alternatives. Finally, we employ for the first time in the resource economics and adaptive management literature (to our knowledge) a *density projection* approximation method that enables us to solve an otherwise intractable high-dimensional control problem. The approach is general: the density projection method results in concise analytical conditions for determining information dynamics which hold across a broad set of applications (exponential family distributions).

Hierarchical modeling—also called multilevel or state-space modeling—is a framework in which complex processes, such as uncertain fish population dynamics, are broken down into a set of conditional sub-models that are coherently tied together using simple probability structures (Wikle, 2003).[3] We use a hierarchical model of uncertainty to allow for a more realistic and flexible representation of uncertainty in ecosystem dynamics and therefore the beliefs of the managers. Practical usage of hierarchical models in dynamic resource management models, however, has been limited by state space dimensionality concerns.

To address the challenge of the growing dimensionality of the belief state, researchers often discretize and constrain the belief state space. For example, extending a shallow lake pollution model to incorporate learning, Bond and Loomis (2009) limit uncertainty over a critical nutrient threshold to two possible levels. Another recent example is Johnson (2011), who describes adaptive management of waterfowl harvest by the U.S. Fish and Wildlife Service, who aims to discern between four competing models of population dynamics.

Rather than constrain the belief state space that could limit the ability to map model outputs to management actions, we use a density projection method to solve for the optimal AAM policy in the presence of a hierarchical model of uncertainty. A limited set of hierarchical models are, in Bayesian terminology, conjugate, meaning that an explicit closed form solution exists for the belief state dynamics. However, hierarchical models are typically non-conjugate, as is the case here. We gain tractability for our non-conjugate hierarchical model by approximating the posterior distribution of the manager's beliefs using a function that minimizes the Kullback–Leibler (KL) divergence between true updated beliefs and an approximation (Zhou et al., 2010). The approximation of the belief state through this density projection[4] method converts the *belief* MDP to a *projected-belief* MDP with lower dimensionality. Relatively simple analytical expressions can be derived to identify the parameters of the projected or approximate beliefs for a wide range of applications. A further advantage of this conversion is the ability to solve for the AAM solution using standard dynamic programming methods, such as value or policy iteration (Zhou et al., 2010).

Finally, a number of studies consider applications similar to ours (though not with a hierarchical uncertainty model) where a manager can learn about a system by observing population dynamics (e.g., Hauser and Possingham (2008), Rout et al. (2009) and McDonald-Madden et al. (2010)). However, in these studies, the researchers constrain the set of management actions to a set of 2–3 discrete choices for tractability (e.g., high and low harvest levels). While the limited set of control actions can be informative for how learning might change management responses and it reduces the dimensionality of the problem, we allow for a continuous choice control variable that is more consistent with how fish populations (and many other natural resources) are managed. A notable exception to the use of limited action and/or belief spaces is Wieland (2000) who allows for continuous beliefs and actions in a generic, linear regression-style learning model (Kalman filter). However, Wieland assumes a convenient, conjugate structure using normal distributions for tractability, in contrast to the framework here which does not require selecting from a limited set of conjugate models.

[1] These two types of uncertainty are sometimes referred to as aleatory uncertainty, which refers to randomness in samples, and epistemic uncertainty stemming from basic lack of knowledge about particular processes (Paté-Cornell and Elisabeth, 1996).

[2] Learning by observing population dynamics without the requirement of perturbing the system is similar to the information dynamics in Rout et al. (2009). An alternative approach often presented in the engineering literature on adaptive dual control (see Filatov and Unbehauen (2000) for an overview) and in the foundational text of Walters (1986) is to model information as emerging from "exciting" or perturbing the system from the status quo (e.g., perturbing the system by changing the stock of fish). Our core methodological contributions are adaptable to either setting, though specific numerical results will obviously differ depending on the nature of the learning process.

[3] A common hierarchical model used in econometrics is the random effects model. More than simply a technical approach, Royle, Dorazio (2008) argue that hierarchical modeling is a conceptual framework for doing science that is flexible, fosters the fundamental activity of model construction and elucidates the nature of inference. Several recent reviews remark on the increasing popularity of the framework (Gelfand, 2012; Schaub and Kéry, 2012). Halstead et al. (2012) state that "we have likely only seen the tip of the iceberg with regard to the utility of these models in ecology" (p. 134).

[4] This general approach is known by other terms, e.g. *assumed density filtering* (Maybeck, 1982) and *belief compression* (e.g. Roy et al., 2005).

We illustrate the contribution of density projection methods by solving an otherwise intractable stochastic bioeconomic model. We consider the investment decisions of a manager employing AAM in a noisy environment to maximize the expected net present value from fishing given both continuous action and belief state spaces. We find that differences in the manager's belief state can drive strong differences in resource use policy. Not surprisingly, we also find that AAM can result in significant efficiency gains relative to a non-learning approach. However, the AAM gains relative to a passive learning approach are less dramatic and we show how they depend strongly on the nature of the learning process. We also show that long-run gains associated with AAM are achieved by *investing* in learning, which can take the form of costly deviations from the optimal path without learning. Thus, when setting expectations, advocates of AAM must be aware that implementation of such optimal endogenous learning may bring short-run sacrifice.[5]

The paper proceeds as follows. After outlining the basic ecological and economic components of the bioeconomic model we develop the hierarchical model of survivorship and Bayesian learning model in detail. Next we describe how belief state approximation, or density projection, is accomplished via minimization of the KL divergence. We then discuss the solution of the stochastic dynamic programming problem and contrast the performance of various management approaches to learning.

2. Model

We present the stochastic bioeconomic model in three parts: (1) the fish population dynamics, (2) the economics and management component, and (3) the model of uncertainty and beliefs of the manager.

2.1. Fish population model

The population model describes the dynamics of a biological (fish) species whose life history is summarized in two stages: juvenile and adult.[6] The juvenile and adult stages are subject to different types of competition and mortality and do not necessarily occur in the same habitat. For example, our model is a special case of Sanchirico and Springborn (2011) who consider a species with ontogenetic migrations where adults reside on a coral reef and the juveniles are in sea grass and mangrove habitats.[7]

In each period, the adult population given by N_t produces larvae that become juveniles according to the function $Z(N_t) \equiv \theta N_t^{\gamma}$, where γ and θ are both nonnegative. Theta is a scaling factor that maps adults to juvenile biomass ($\gamma > 1$ ($\gamma < 1$) implies increasing (decreasing) returns to scale).

While there are a number of ways to introduce uncertainty into the population dynamics (state uncertainty, parameter uncertainty), we assume uncertainty enters in the survivorship of juveniles.[8] In a deterministic and stable environment, the new recruits to the adult population, K, would be given by Z^*S, where S is a constant survivorship rate. Here, we assume that survivorship is a binomial random variable with a survivorship ("success") probability of S_t:

$$K_t \sim \text{Binomial}\,(Z_t, S_t), \tag{1}$$

where S_t can vary over time.

We assume that recruits enter the adult population according to a Beverton–Holt recruitment function, $b_1 K_t/(1+ b_2 K_t)$, where b_1 describes the survival rate at low densities, and b_1/b_2 is the saturation limit with respect to the recruitment. This formulation is consistent with a density-dependent process stemming from recruits competing with other recruits for space and resources during settlement. Note that this is different than the survivorship in the juvenile stage that is driven by predation, environmental conditions, etc.

Because of tractability concerns, we make a simplifying assumption that the juvenile stage occurs on a time scale shorter than the adult population dynamics. This assumption permits us to substitute the juvenile stage directly into the adult population dynamics rather than including a separate state equation for them (see e.g., Sanchirico, 2005; Grimsrud and Huffaker, 2006, and Springborn and Sanchirico, Mumby, 2009). Combining recruitment, fishing and natural mortality, the change in the adult fish stock is

$$N_{t+1} = F(N_t, H_t, K_t) = N_t(1-\psi) + \frac{b_1 K_t}{1+b_2 K_t} - H_t. \tag{2}$$

where ψ is a natural mortality of the adult stage, and H_t specifies harvest.

[5] Another important issue discussed by Doremus (2007) is that agency discretion might be constrained by statute. The implication is that the ability of managers to experiment and learn could be more limited than what is currently being modeled in the adaptive control literature.

[6] Other examples of bioeconomic models with complex age-structured population dynamics include Smith et al. (2008) and Tahvonen (2009). Smith et al., (2008) emphasize the importance of capturing life history features and details of production in population models in their empirical analysis of a complex age-structured model. Tahvonen (2009) focuses on the management implications of age-structured information in a theoretical model.

[7] The model considered here, however, also captures a richer set of possibilities, such as when the stages occupy different depths or the adults and juveniles are present in the same habitat but are subject to different rates of predation and competition (density-dependent processes).

[8] Massey et al. (2006) also utilize a population model which highlights the importance of juvenile survivorship, albeit in a context of valuing improvements in water quality for enhancing fishery productivity.

With no harvest, the deterministic equilibrium population size (assuming $K = Z^{*}S$, and $\gamma = 1$) is

$$N^{eq} = \frac{1}{\psi}\left[\frac{b_1}{b_2} - \frac{\psi}{S\theta b_2}\right]$$

where the higher the survivorship rate of juveniles, the greater the equilibrium population and the faster rate of growth (note that as $S- > 0$, the equilibrium population and growth rate go to zero). That is, unlike the logistic growth function where two separate parameters are used for the growth and carrying capacity (long-run equilibrium population), in this formulation an increase in S increases the rate of growth and long-run equilibrium, all else being equal. The uncertainty the manager faces, therefore, spans both the rate of growth and the long-run equilibrium population.

Eq. (2) differs from some previous normative bioeconomic models in fisheries with uncertainty (see, e.g., Reed, 1979 and Sethi et al., 2005) in that we assume that reproduction and recruitment occur before harvest. In other words, we are not recasting the problem in terms of escapement (N_t–H_t). We do, however, follow the previous literature in assuming that at the time the optimal harvest is chosen in each period the fish stock is known. The fish stock in the future is subject to stochasticity and as we describe below, we model the changing beliefs of the manager with respect to the process governing stock dynamics over time.

2.2. Economic and management model

We model a benevolent manager who chooses the level of fish catch in each period to maximize the present discounted value of expected profits. The catch yields instant economic returns that accompanies the extraction and selling of the fish. The catch also affects the signal the manager uses for updating beliefs, as it impacts the size of the fish population (discussed further below).

Let $\pi(H_t|N_t)$ represent the immediate profits of harvest conditional on the fish stock. Fishing profit is assumed to be increasing at a decreasing rate in harvest and fish population ($\pi_h > 0$, $\pi_{hh} \le 0$, $\pi_N > 0$, $\pi_{NN} \le 0$). The discount factor is specified by δ. The decision maker's belief state is I_t and the belief state dynamic equation is $I_{t+1} = G(I_t, Z(N_t), K_t)$. The Bellman equation specifying the manager's stochastic optimal control problem is given by

$$\begin{aligned} &V(N_t, I_t) = \max_{H_t}\{\pi(H_t|N_t) + \delta E_{\{N,I\}}V(N_{t+1}, I_{t+1})\} \\ &N_{t+1} = F(N_t, H_t, K_t) \\ &I_{t+1} = G(I_t, Z(N_t), K_t) \\ &N_{t=0}, I_{t=0}, 0 \le N_t, 0 \le H_t \end{aligned} \tag{3}$$

where $E_{\{N,I\}}$ is the expectation operator with expectations taken with respect to the population of the fish stock and the manager's belief, $V(N_t, I_t)$ is the value function in period t, and $N_{t=0}$, $I_{t=0}$ are the initial conditions for the fish stock and beliefs.

Following our assumptions on the curvature of fishing profits, we assume fishing profits are:

$$\pi(H_t, N_t) = (a_0 - a_1 H_t)H_t - \frac{c}{N_t}H_t, \tag{4}$$

where a_0 is the choke price, a_1 is the slope of the demand curve, and c is a cost parameter. The per unit cost of harvest in Eq. (4) depends inversely on the fish population. By selecting fishery profit to serve as the objective function we are implicitly assuming risk neutrality. The manager's control problem described in (3) is conditioned on the stock dynamics $F(\cdot)$ specified in Eq. (2). In addition, the problem is constrained by belief state dynamics $G(\cdot)$, which we now discuss in detail.

2.3. Survivorship and belief models

We employ a flexible hierarchical approach consisting of two nested levels to model the uncertain juvenile survivorship. The first level is Eq. (1), where the number of surviving juveniles K_t is a binomial random variable given by an initial number of juveniles Z_t ("trials") and the success rate S_t. Eq. (1) implies that the greater the number of juveniles (due to higher N or greater average larval production from higher θ or γ), the more trials that are observed and thus the more the decision maker learns about survivorship in a given period. In the second level, S_t is a random variable that can vary from one period to the next.

Since the survivorship rate is on the unit interval, we assume that S_t follows a beta distribution. The beta distribution is typically depicted as a function of two shape parameters, α and β. We re-parameterize the beta in terms of its mean, $\overline{S} = \alpha/(\alpha+\beta)$, and *concentration*, $\rho = \alpha+\beta$ (see e.g. Kruschke, 2011) where ρ conveys a sense of the inverse of spread.

Fig. 1 depicts the two-stage hierarchical survivorship and belief models. The true parameters $\overline{S}$ and ρ are fixed scalars. The belief state model presents two additional distributions, which capture the decision maker's information on the true values of $\overline{S}$ and ρ. The structural uncertainty in our model creates a setting of ambiguity in which the probabilities of

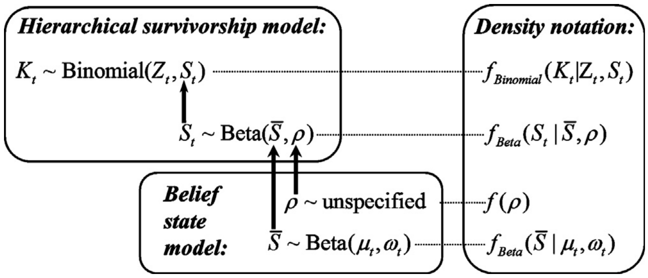

Fig. 1. Schematic of combined survivorship and belief state models.

potential outcomes are not known with certainty. The belief state framework is an example of modeling ambiguity by introducing second-order probabilities, or beliefs over beliefs (Siniscalchi, 2008).[9]

In what follows, we first discuss the more traditional non-hierarchical approach as a means to highlight the value, challenges, and differences when using a hierarchical approach. In each case, the form of the belief model will depend on whether or not the population model is hierarchical.

2.4. Traditional non-hierarchical model

A special case of our approach is the more traditional non-hierarchical (NH) approach. Under this approach, one level of uncertainty such as S_t is assumed fixed: $S_t = \overline{S}, \forall\, t = 1, \ldots, \infty$. This assumption implies that the belief distribution $f(\rho)$ represents strong beliefs, i.e. no variation in S_t ($\rho \to \infty$) since its true density is concentrated on a single point.

A computational advantage of the NH approach is that Bayesian updating of the belief state given observations on the system (Z_t, K_t) can result in a simple, closed form for the belief state dynamics. For example, assume that the decision maker's prior beliefs about the true value of $\overline{S}$ are given by a beta distribution with mean μ_t and concentration ω_t, which is a common approach for modeling beliefs about the success parameter of a binomial distribution (see Gelman et al., 2004, p. 34). Given observations (Z_t,K_t), the beta distributional assumption, and the application of Bayes' theorem, the posterior distribution for $\overline{S}$ that describes the updated belief state also follows a beta distribution—a convenient property known as conjugacy. In this NH case, the belief state dynamics, $I_{t+1} = G(I_t, Z_t, K_t)$, are given by

$$\begin{bmatrix} \mu_{t+1} \\ \omega_{t+1} \end{bmatrix} = \begin{bmatrix} \frac{\mu_t \omega_t + K_t}{\omega_t + Z_t} \\ \omega_t + Z_t \end{bmatrix}. \tag{5}$$

We follow the Bayesian convention of referring to the arguments of the prior, μ_t and ω_t, as *hyperparameters* to distinguish them from the parameters of the underlying system (survivorship model).

In Eq. (5), the prior belief state (μ_t, ω_t) maps to the posterior belief state $(\mu_{t+1}, \omega_{t+1})$ in simple fashion. This framework is computationally convenient and arguably appropriate for a number of applications, such as the fishery recovery rate model of Hauser and Possingham (2008). Given a large number of juveniles observed in each period (high Z_t), however, the NH approach leads to a trivial learning process. In fact, beliefs can become highly concentrated in a single period, even with a relatively diffuse (non-concentrated) prior. The NH model is also unrealistic in applications such as ours in which the process governing juvenile survivorship is unlikely to remain fixed from year to year.[10]

2.5. Hierarchical model

The hierarchical model is more consistent with the way managers learn over time about complex systems. That is, observations are made in a *noisy environment* and learning about the behavior of the system takes time. In the hierarchical approach, both levels of the survivorship model are stochastic: S_t varies from period-to-period, which is equivalent to assuming that $0 < \rho < \infty$. A fully-flexible learning model incorporates belief states on the true level of $\overline{S}$ and ρ, which results in four hyperparameters as state variables (two for $\overline{S}$ and two for ρ).

Because of the computational issues associated with the dimensionality of the fully-flexible model, we follow the convention in the literature (see, e.g., Gelman et al., 2004, pg. 46) and assume that the scale parameter (ρ) is fixed and

[9] Second order probabilities—beliefs in which a probability measure is used to model ambiguity with respect to another probability measure—were first formalized mathematically in the 1950s (see Skyrms (1980) for an overview). This terminology is most commonly used in the ambiguity aversion literature, which focuses on individual choice. For example an individual may be uncertain about the probability of drawing a ball of a particular from and urn (Seo, 2009) or the probability of human survivorship (Hudomiet and Willis, 2012). Uncertainty over these probabilities may be modeled using second order beliefs, or a distribution capturing the likelihood of different likelihoods, in fashion similar to the belief models in this article for uncertain survivorship. For further discussion of a range of approaches for modeling ambiguity, including the merits of both Bayesian versus non-Bayesian approaches see Gilboa and Marinacci (2011).

[10] See chapters 4 and 5 of Walters (1986) for a discussion of multiple instances in which it is "practically impossible" to develop natural resource models in which parameters may confidently be assumed constant.

known.[11] Thus, S_t is drawn each period from a distribution with known concentration, ρ, but an unknown mean, $\overline{S}$. Even in this setting tractability is still a challenge.

Unlike the non-hierarchical model, the posterior in the hierarchical model is non-conjugate and does not result in an explicit closed form solution for belief dynamics. Specifically, given the prior belief state (μ_t, ω_t), observations (N_t, K_t), the posterior distribution for $\overline{S}$ following Bayes' Theorem (see Gelman et al., 2004, p. 8) is

$$f(\overline{S}|\mu_t,\omega_t;Z_t,K_t)=\frac{f_{\text{Beta}}(\overline{S}|\mu_t,\omega_t)\cdot L(K_t|Z_t,\overline{S})}{\int_0^1 f_{\text{Beta}}(\overline{S}|\mu_t,\omega_t)\cdot L(K_t|Z_t,\overline{S})d\overline{S}}, \tag{6}$$

where the probability (likelihood function) of observing K_t successes given $Z_t(N_t)$ trials under mean productivity $\overline{S}$ is given by

$$L(K_t|Z_t,\overline{S})=\int_0^1 f_{\text{Binomial}}(K_t|Z_t,S_t)f_{\text{Beta}}(S_t|\overline{S})dS_t. \tag{7}$$

The integral in the likelihood function above follows from taking the expectation of the (binomial) density of K_t with respect to (beta distributed) survivorship, S_t.

For economy of notation we define the posterior parameter vector $\lambda_{t+1}=(\mu_t,\omega_t,Z_t,K_t)$ and the true posterior as $f(\overline{S}|\lambda_{t+1})$. The integral in the likelihood function in Eq. (7) induces a non-conjugate posterior that is not precisely beta. Exact parameterization of the posterior is given by the vector λ_{t+1}, which grows in dimension by two each period. In a dynamic programming context, over time the increasingly high-dimensional belief space becomes intractable.

We gain tractability for the inconvenient hierarchical model by approximating the posterior distribution with a "nice" approximate function that is "close to" the true function in the sense of Judd (1998), p. 195. We formalize the idea of "nice" by seeking a function that allows us to maintain a low-dimensional belief space, as well as obtain a posterior distribution in the same distribution family as the prior (Zhou et al., 2010). The concept of "close to" is implemented by finding a function that minimizes the Kullback–Leibler (KL) divergence between the true and approximate function.[12] While this so-called density projection approach has been applied in the Bayesian empirical literature (e.g. Chen and Shao, 1997) and broader optimal control literature (e.g. Maybeck, 1982; Zhou et al., 2010) to our knowledge it is novel to the settings of adaptive resource management, resource economics and hierarchical modeling.

The density projection approach involves selecting the parameters $(\hat{\mu}_{t+1},\hat{\omega}_{t+1})$ of an approximate function $\hat{f}_{Beta}(\overline{S}|\hat{\mu}_{t+1},\hat{\omega}_{t+1})$ to minimize the KL divergence between the true and approximating functions.[13,14] The KL divergence, D_{KL}, is defined as the expected log-likelihood ratio of the true and candidate densities, where the expectation is over the domain of $\overline{S}$

$$D_{\text{KL}}(f(\overline{S}|\lambda_{t+1})||\hat{f}(\overline{S}|\hat{\mu}_{t+1},\hat{\omega}_{t+1}))\equiv\int_0^1 f(\overline{S}|\lambda_{t+1})\ln\frac{f(\overline{S}|\lambda_{t+1})}{\hat{f}(\overline{S}|\hat{\mu}_{t+1},\hat{\omega}_{t+1})}\,d\overline{S}. \tag{8}$$

The fitting optimization problem is given by

$$\min_{\hat{\mu}_{t+1},\hat{\omega}_{t+1}} D_{\text{KL}}(f(\overline{S}|\lambda_{t+1})||\hat{f}(\overline{S}|\hat{\mu}_{t+1},\hat{\omega}_{t+1})). \tag{9}$$

Zhou et al. (2010), p. 1104 show that for candidate distributions that belong to an exponential family (e.g. beta distribution) the KL divergence has a unique minimum which is achieved when the first order optimality condition is satisfied.

2.6. Implications of density projection

To implement density projection using KL divergence minimization, we solve for the first-order conditions from Eq. (9), which simplify to

$$\begin{aligned} &E_f(\ln(\overline{S})|\lambda_{t+1})=\psi(\hat{\omega}\hat{\mu})-\psi(\hat{\omega})\\ &E_f(\ln(1-\overline{S})|\lambda_{t+1})=\psi(\hat{\omega}(1-\hat{\mu}))-\psi(\hat{\omega}), \end{aligned} \tag{10}$$

[11] We set ρ=25. If for example $\overline{S}$= 0.5, the assumed scale parameter implies that the 95% confidence interval is (0.31, 0.69).

[12] An alternative approach for approximating the posterior in Bayesian inference is to use a normal approximation to the posterior as motivated by Bayesian versions of the central limit theorem (see Berger (1985), p. 224 for a discussion). This technique is unsatisfactory for our purposes since we are interested in learning implications for instances of few observations (as well as many) and we know that $\overline{S}$ belongs to a discrete domain on the unit interval.

[13] The KL divergence can be thought of as the *relative* entropy between the true and candidate densities. The concept of *maximum* entropy has been applied to problems in resource or pollution management to estimate unknown parameters and unobserved variables for ill-posed problems where the number of unknowns exceeds the number of observations (e.g. Golan, Judge, 1996; Kaplan et al., 2003). Alternatively, *minimizing* entropy has been explored as an objective itself in dynamic stochastic control (e.g. Wang, 2002).

[14] Fitting $\hat{f}_{Beta}(\overline{S}|\hat{\mu}_{t+1},\hat{\omega}_{t+1})$ to minimize the KL divergence, the expected log-difference between densities, is analogous to finding maximum likelihood estimates of $\hat{\mu}_{t+1}$ and $\hat{\omega}_{t+1}$ (Eguchi and Copas, 2006). Specifically, if the KL divergence is minimized over a discrete set of points randomly selected from the domain of $\overline{S}$, then the D_{KL}-minimizing parameter levels are identical to the maximum likelihood estimates determined using a beta density evaluated at the same set of points.

where time subscripts on the hyperparameters of interest have been suppressed and ψ is the digamma function.[15] Note that $\ln(\bar{S})$ and $\ln(1-\bar{S})$ are the sufficient statistics for the beta distribution. Each side of the equations in (10) corresponds to the expected value of $\ln(\bar{S})$ or $\ln(1-\bar{S})$. On the left hand side this expectation is evaluated *with respect to the true posterior* ($f(\bar{S}|\lambda_{t+1})$). On the right hand side the expectation is evaluated *with respect to the approximate posterior* ($\hat{f}_{Beta}(\bar{S}|\hat{\mu}_{t+1},\hat{\omega}_{t+1})$), which takes a known closed form based on the digamma function.

Zhou et al. (2010) show that such concise analytical expressions exist for determining belief dynamics for a wide range of approximate posterior models. Specifically, for exponential family distributions,[16] minimizing the KL divergence is equivalent to matching the expected value of the sufficient statistics for the true and approximate posteriors. This important finding greatly simplifies application of the density-projection method, ensuring a relatively concise set of known analytical expressions for evaluating density projection belief state dynamics—analogous to the two conditions in (10)—for an extensive selection of common distributions.

Density projection represents a substantial advance relative to previous simulation-based methods for approximating the posterior. Brooks and Williams (2011) simulate the posterior in a dynamic learning problem by generating a (weighted) set of samples from the posterior using the technique of particle filtering. As noted by Zhou et al. (2010), simulation “is not very helpful for reducing the dimensionality of the belief space” (p. 1105). To allow for a continuous belief state over the uncertain variable, it is still necessary to represent beliefs parametrically, e.g. with sufficient statistics from a particular distribution. While Brooks and Williams simply assume that the posterior can be reasonably approximated with a normal distribution (with a mean and spread given by the simulated posterior), they acknowledge, “unimodal Gaussians are unable to capture the precise shape of all distributions which may occur” and are not always appropriate (p. 215). Zhou et al. (2010) emphasize the limitations of the simulation approach of Brooks and Williams, observing that specification for how to compute approximate posterior parameters has to date only been explored for the Gaussian case.

In contrast, density projection via KL divergence minimization provides explicit guidance on how to select the parameters of the approximate posterior, as outlined above. Conditional on this guidance, simulation methods could still be used to evaluate necessary integrals—e.g. the left hand side terms in (10). However, we use numerical quadrature for this step since it is more efficient when feasible. The KL minimization framework is quite flexible, accommodating, for example, distributions that are bimodal (e.g. beta as used here) or fat-tailed.[17]

Using Monte Carlo analysis, we find that beliefs updated using density projection converged over time to a neighborhood of the truth as the number of observations grew (see details in Appendix). For example, strict convergence would result in a deviation of 0% and across 1000 simulations and seventy-five periods, we found an average divergence of 5%. The divergence is due to learning in a “noisy world”, i.e. the variability in the system that allows in finite time for instances in which the series of “actual” stochastic outcomes are predominantly above or below the true mean.

Overall this adaptive management problem involves nesting an estimation step—learning about structural parameters—within a larger stochastic dynamic programming problem. Discrete choice dynamic programming (DCDP) models are related in spirit, though inverted in construction, featuring a stochastic optimization problem nested within an econometric estimation routine (Rust, 1994).[18] Challenges from dimensionality also arise in DCDP, for example when there are many parameters to estimate or many possible values of a state variable (Aguirregabiria and Mira, 2010). While numerical integration is applied when feasible (e.g. Provencher, 1995), simulation offers a way forward when necessary (e.g. Timmins, 2002).

3. Numerical results

Using dynamic programming, we solve for the optimal value and policy function under an active-adaptive management (AAM) approach, which solves the full optimization problem in Eq. (3). Under AAM, the decision maker optimally adjusts harvest policy given the instantaneous returns from harvesting, beliefs about the productivity of the system, expectations about what can be learned over time, and expectations over future harvest levels. Also recall, in our system, the number of trials in the learning model increases with the size of the population. Therefore, there is a direct trade-off between harvesting more today and the associated level of returns with the lower population size and less information from which to learn. Details on parameter values (which generally follow Sanchirico and Springborn (2011)) and the value function iteration solution procedure are provided in the Appendix.

For many advocates of AAM, there is a presumption that optimal policies under AAM are more precautionary than under the status quo, which is often assumed to be a case with no learning.[19] We investigate the potential conditions under which optimal investment in learning has a significant effect on decision-making by solving for a non-learning management (NLM) approach. Under NLM, the decision maker's belief state is held fixed over time: $(\mu_t,\omega_t)=(\mu_0,\omega_0)$.

[15] The logarithmic derivative of the gamma function.

[16] Exponential family distributions include normal, exponential, gamma, χ^2, beta, Dirichlet, Bernoulli, categorical, Poisson, Wishart, Inverse Wishart, etc.

[17] Note that Student's *t*-distribution, featured prominently in recent discussions of climate change under fat-tailed uncertainty (e.g. Weitzman, 2009), is not in the exponential class of distributions. KL minimization is still feasible, though the requisite first order conditions will not necessarily simplify as cleanly as for exponential family distributions.

[18] We thank a reviewer for pointing out this parallel.

[19] If one considers experimental action itself to pose a substantial risk, then AAM would not be considered precautionary. However, if moderate risks are tolerable and one is concerned with avoiding extreme risk stemming from structural uncertainty, then AAM might be thought of as precautionary. Given this potential ambiguity we use the term sparingly.

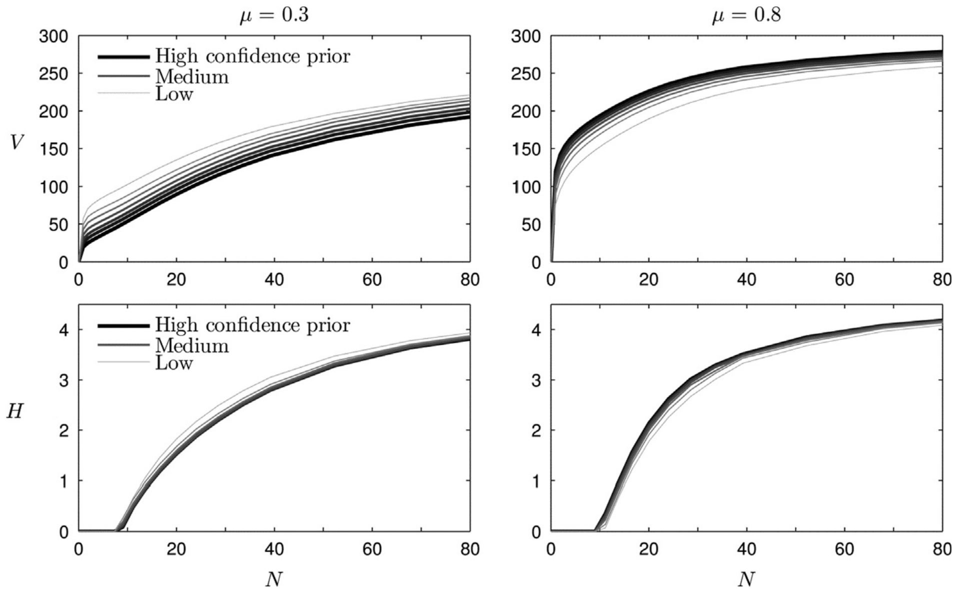

Fig. 2. Optimal value and policy functions under AAM.

We also supplement our analysis of the value and policy functions of NLM and AAM with a set of Monte Carlo simulations that permit us to introduce another "learning model", which we denote as *passive adaptive management* (PAM). Under PAM, the decision maker utilizes the NLM policy rule but takes advantage of observations made on population dynamics to update beliefs. For Monte Carlo simulations under PAM and AAM, information dynamics were updated using the same density projection approach used in the solution to the dynamic programing problem.

3.1. AAM, beliefs, and confidence

The concentration parameter, ω_t, in our belief state density $f_{Beta}(\bar{S}|\mu_t, \omega_t)$ is a measure of the decision-maker's *confidence.* As ω_t grows, the spread in the belief state declines and confidence increases. When ω_t is relatively low, updated beliefs will be relatively sensitive to new information. Conversely, when ω_t is already relatively high, new information will have less of an effect on the belief state (see Appendix and Fig. A1 for an illustration).

We investigate within AAM how the degree of confidence the manager holds affects the optimal value and policy functions, as their expectations on the value of the system depend on the degree of confidence. In particular, we present in Fig. 2 optimal value and policy function solutions under two belief states that include low expected productivity ($\mu = 0.3$) and high expected productivity ($\mu = 0.8$) as well as a range of confidence (ω) levels. Solutions for the NLM/PAM case are visually similar and are presented in the Appendix.

We find that the value function increases with confidence when the decision-maker believes productivity is relatively high. On the other hand, we find that the value function decreases with confidence when productivity is believed to be low. The latter result is surprising, especially given an expectation that reducing the spread of beliefs is valuable. There is, however, an intuitive explanation for this effect. When expected productivity is low ($\mu_t = 0.3$), a low level of confidence means that the decision maker still believes there is a significant (if minority) chance that the true productivity will be higher.[20] Of course a lack of confidence also raises the prospect that true productivity is lower, but the fact that expected productivity is low constrains this downside uncertainty. A high level of confidence in the low productivity belief means that the decision maker *does not* believe that there is a significant chance of higher productivity and therefore the value function is lower, all else being equal.

The value function effect described above is driven by changes in the relative likelihood the decision-maker allocates to more or less preferable true states. It is important to separate that effect from the value of "learning" or reducing error, i.e. the gain from tailoring policy ever-closer to the true state of the system. Conditional on any particular true state of the system, *if* confidence increases *due to the accrual of observations on the actual system*, then we would expect performance to go up as confidence increases. However, this second effect is not discernible from the value function, which conveys only the decision-maker's expected net profits *conditional on their belief state,* rather than conditional on the true state of the system. We return to this second effect below using Monte Carlo simulations.

[20] Recall in our system the productivity associated with a change in the survivorship rate comprises both a change in the growth rate of the stock from one period to the next *and* the long-run equilibrium size of the population.

The AAM policy functions for optimal harvest follow a similar pattern as the value functions with respect to the change in the level of confidence (see lower panels in Fig. 2). For example, when expected productivity is relatively high (bottom right panel of Fig. 2) decreasing confidence leads to less harvest, even though there is a reduction in expected profits from fishing.

For intuition into what drives the optimal (harvest) policy it is instructive to decompose the effects of uncertainty. For a decision problem of similar structure, Bar-Shalom (1981) develops an approximate value function that can be decomposed into three components. The first is the *deterministic* component, the value along the nominal trajectory in which all uncertainties are ignored. The second is the *stochastic* component associated with all uncertainty that is (or is assumed to be) irreducible. In the NLM/PAM learning models, these two effects comprise the value function. The value function under the AAM strategy, however, includes a third *experimentation* component that captures the expected future value of information.[21] From the NLM/PAM policy functions (Appendix) we observe that the stochastic effect from lower confidence can drive *greater* harvest when productivity is believed to be low (e.g. $\mu_0=0.3$) and *lower* harvest when productivity is believed to be high (e.g. $\mu_0=0.8$). When the belief is that productivity is high, for example, there is more opportunity for true productivity to be lower. Consider an extreme case in which a manager believes productivity is extremely high, $\mu=0.99$. To the extent that the manager is uncertain about $\overline{S}$, the opportunity for (and implications of) a downside surprise ($\overline{S}<0.99$) is greater than that of an upside surprise ($\overline{S}>0.99$).

3.2. AAM and harvest levels

To disentangle the stochastic and experimentation effects, we consider the relative levels of optimal harvest from AAM and NLM. Given the structure of the survivorship and belief models, we have that less harvest results in greater production of juveniles in the future, in greater number of juveniles observed, and thus a higher rate of learning about survivorship (faster reduction of structural uncertainty). Since forgoing some degree of harvest in the current period generates an opportunity cost, a lower harvest level under AAM relative to NLM ($H_{\mathrm{AAM}}/H_{\mathrm{NLM}}<0$) would be consistent with the notion of costly immediate investment to enhance the rate of learning.

In Fig. 3, we plot the ratio of $H_{\mathrm{AAM}}/H_{\mathrm{NLM}}$ for a low and high productivity system as a function of confidence (ω) at various levels of the fish population, *N*. We plot the ratio, because the policy functions for the AAM and NLM/PAM strategies are not identical but visually similar. We find that accounting for the opportunity to learn can lead to differences in optimal harvest policy, apparent in the figure when the ratio departs from 1. Under the high productivity case (right panel of Fig. 3), we also find the intuitive result that AAM harvest levels relative to NLM are low when the manager's confidence on the state of the system is low (low ω). The manager invests in learning by reducing the harvest level.

In the low productivity case (left panel of Fig. 3), we observe that the optimal harvest under AAM can be greater than under NLM, particularly for low population and low confidence levels. In this case, for both the NLM and AAM strategies, the raw levels of optimal harvest increase as confidence falls, i.e. as ω becomes smaller (see Fig. 2 and Fig. A2 in the Appendix). For NLM, this is entirely due to the stochastic effect. For AAM we might expect the learning incentive (which increases as ω falls) to temper this increase in harvest. Instead harvest under AAM is higher in this particular case.

We find therefore that even if AAM typically motivates actions that increase the rate of learning relative to NLM/PAM, this is not necessarily universally the case. A similar result, in which the optimal PAM action generates a higher rate of learning than an optimal AAM action in a subset of the belief space, is also found by both Hauser and Possingham (2008) and Rout et al. (2009). In Rout et al. (2009), each period the manager decides how to allocate two captive bred individuals between two separate wild populations, one of which has an unknown annual mortality rate. The rate of learning about the uncertain mortality rate parameter is increasing the number of individuals allocated to that population since only captive bred individuals are monitored. In the model of Hauser and Possingham (2008), increased harvest of a generic species increases the likelihood of a population collapse. A collapse interrupts harvest but also provides an opportunity to learn about the key uncertain parameter of interest: the rate of recovery from a collapsed to partially-restored state. Thus, increased harvest enhances the rate of learning about the uncertain system process.

Hauser and Possingham (2008, p. 78) attribute the surprising finding to a “precautionary effect” since in their example the more informative control action selected under PAM but not AAM involves increased risk of a collapse of the harvested population. This explanation does not appear to hold in general since in our problem the result involves *increased* harvest.

For intuition on this counterintuitive result, we return to the decomposition of Bar-Shalom (1981) in which he observes that if the control can reduce structural uncertainty via experimentation (as is the case here) *then it can change the stochastic effect*. This implies that the stochastic effect operating under the NLM/PAM strategy will not necessarily be the same as the stochastic effect operating under the AAM strategy since in the latter case the stochastic effect is modified by the

[21] We summarize Bar-Shalom's (1981) decomposition using terminology consistent with our application to bioeconomics and adaptive management. Bar-Shalom, writing within the engineering *dual control* literature, refers to the three components of the cost-to-go (value) function as deterministic, caution, and probing. We use the more general term *stochastic* instead of *caution* given that in our application greater stochasticity does not necessarily imply a more cautious policy as it does in Bar-Shalom's problem. Furthermore, while learning in Bar-Shalom's problem occurs via *probing* (perturbing the system from its current state), we use the more general term *experimentation* since in our application the rate of learning increases with the stock size. In the spirit of this decomposition from the dual control literature, Bond (2010) also develops a related decomposition of the value function into several components including: certainty equivalent, uncertainty, passive learning, and active learning.

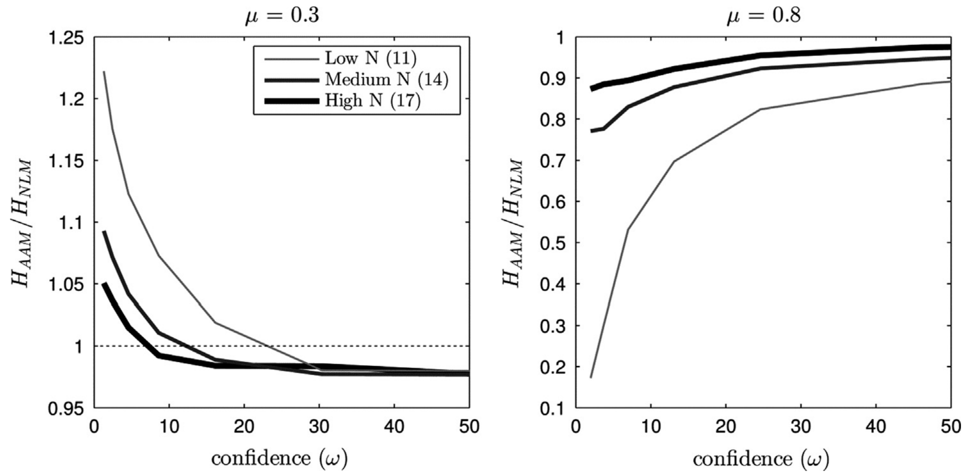

Fig. 3. Relative optimal harvest (H_{AAM}/H_{NLM}) as a function of confidence (ω) for various levels of *N. Note*: The left panel considers low productivity, and the right panel is high productivity. The lines represented by different thickness correspond to various population levels: thin ($N=11$), medium ($N=14$) and thick ($N=17$).

opportunity to experiment.[22] In our case (left panel of Fig. 3), the interaction between the experimentation and stochastic effect is strong enough in some (albeit extreme) cases to lead to an optimal AAM action that is *less* informative than the optimal NLM/PAM action.[23]

3.3. Relative performance of AAM, PAM and NLM

While the calls to implement AAM management are becoming ubiquitous in the natural resource management literature (Doremus, 2010), there remain questions on the efficiency gains of AAM relative to more passive learning approaches. These questions remain, because addressing them requires formalizing AAM and being able to solve for the optimal solutions under realistic conditions, such as hierarchical models of uncertainty. Intuitively, we would expect that any difference between the present discounted value of profits averaged over the simulations between AAM and PAM strategies would be small if the beliefs are unbiased and confidence is high. High confidence implies that beliefs have low spread, there is little to learn about the system, and there is a low expected return to investments in information acquisition. Therefore, in what follows, we will focus on cases with low confidence.

We investigate the relative performance of AAM to passive adaptive management (PAM) and a no learning model (NLM) using Monte Carlo simulations. In the simulations, we consider a range of initial conditions for the fish stock (N_0) and belief state space (μ_0, ω_0). In the belief state space, we assume that the decision maker begins with low confidence ($\omega_0=1.3$), but the starting points differ in the expected value of mean productivity, μ_0. For each initial population-information state considered (N_0, μ_0, $\omega_0=1.3$), one thousand Monte Carlo simulations were run ($M=1000$). For each simulation $m\in\{1, \dots, M\}$, the "true" mean productivity $\overline{S}^m$ was drawn from a beta distribution with parameters (μ_0, $\omega_0=1.3$), which is consistent with *imprecise but unbiased* (or accurate) initial beliefs.[24] Conditional on the draw of $\overline{S}^m$, a set of annual productivity parameters, S_t, were drawn for $t\in\{1, \dots, T\}$, where $T=75$. Finally, for each period t in each simulation m, we drew the binomial variable determining the "actual" number of survivors.[25]

We find not surprisingly that NLM performs worse in terms of present discounted value than AAM and PAM. For example, for a low productivity case ($\mu_0=0.2$) the advantage of AAM over NLM in terms of increased average PV ranged from a high of 6.5%, for a degraded stock ($N_0=11$), to a low of 1.8% for a large stock ($N_0=86$). In the case of high productivity ($\mu_0=0.8$) the relative value of learning is lower—the advantage of AAM over NLM ranges from 0.4% to 0.2%. We find in general that the importance of uncertainty declines as the initial fish population increases since at high population levels the cost of imprecise beliefs is reduced and density dependent recruitment induces lower incremental returns to more finely tuned management.

[22] In other words, the NLM and PAM learning models are not necessarily nested within AAM.

[23] Future work in a simpler setting is needed to tease out more explicitly the general conditions under which this result may emerge. As already observed, the stochastic effect under the NLM/PAM strategy leads to an increase in harvest. This occurs because under uncertainty, when expected productivity is low, there is more potential for an upside surprise than down. Tempering this effect is the prospect that productivity is as expected or worse in which case damage that is difficult or impossible to recover from could be done to the stock. We speculate that because the decision maker never expects to know whether this is truly the case, the influence of the tempering may be strong. However, under AAM it may be the case that this tempering is not as strong given that the decision maker anticipates learning further about productivity and making adjustments accordingly, including avoiding substantial harm from overharvest at truly low productivity.

[24] For example, this implies a 95% confidence interval of (10^{-6}, 0.90) when $\mu_0=0.2$ and (0.10, $1-10^{-6}$) when $\mu_0=0.8$.

[25] We set the binomial random component of the population model in advance using the inverse cumulative distribution function method (see Gelman et al., 2004, p. 25). This allows us to impose the same binomial shock (i.e. of the same cumulative density) in any period t for the four different management strategies being evaluated, even after the states under different strategies have diverged.

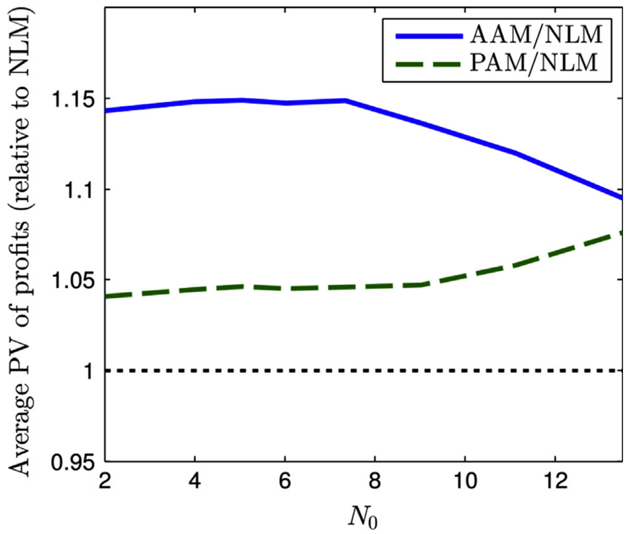

Fig. 4. Average present value (PV) of profits across Monte Carlo simulations for AAM and PAM relative to NLM, as a function of different starting population levels, N_0.

Less obviously, we find that in certain cases PAM can perform essentially as well as AAM, even though in expectation AAM is the ex ante optimal approach. In our application, this is more likely to be the case, for example, when the initial population level is high or when the system is highly productive (i.e. high N_0 and/or μ_0). In the case of high population levels, learning occurs relatively quickly even under the passive strategy (PAM) due to the large sample size and additional investments in reducing harvests to increase sample size (and the learning rate) are not worth very much. Similarly, when the system is highly productive (high μ_0) large sample sizes are apparent either immediately if N_0 is high or relatively quickly since recovery from lower levels of N_0 is quick.

Walters (1986) observed that it is possible for a passively adaptive policy to "produce essentially as informative a sequence" as AAM (p. 249), i.e. there would be very little difference between PAM and AAM when they were not "differentially informative" (p. 260). In the relatively limited literature on the value of AAM over PAM, the result is also not unique. For example, in the context of a shallow lake pollution model, Bond and Loomis (2009) estimate the expected incremental value of AAM above PAM is less than 0.13%, while Rout et al. (2009) find that the incremental value is generally 1–3% in an application to species translocation.

But what if the two approaches are differentially informative, that is, there is a structural discontinuity in how information is acquired in the system that can be anticipated by AAM and not by PAM? All else being equal, we would expect according to Walters (1986) that AAM might lead to potentially large gains over PAM. We explore this potential by considering a situation in which observations on population dynamics are not available when the population level is below a specific threshold. In particular, we assume observations do not accrue unless the juvenile population, Z, is greater than or equal to 7.5, which occurs when $N-H\geq 10.5$. Below the threshold no learning occurs regardless of the strategy (e.g. PAM or AAM) but that does not imply that the policy functions are identical. The potential difference arises, because AAM anticipates the value and ability to learn in the near future (e.g., setting harvest lower initially to lead to a faster recovery) while PAM does not.[26]

Fig. 4 illustrates the average present value (PV) of profits for AAM and PAM—each relative to a NLM baseline—as a function of the initial population level, N_0. For lower stock levels, below $N_0=8$, PAM and AAM each outperform NLM, by approximately 5% and 15% respectively. Over this population range, we see that the average AAM return is approximately 13% higher than for PAM. However, when N_0 is above the threshold, a substantial amount of learning begins to occur without altering harvest and the returns to AAM and PAM begin to converge, which is consistent with our findings without the threshold. Under optimal endogenous learning (AAM) the manager balances the opportunity cost of reduced fish harvest today against the anticipated returns from future learning.

Using the same observation-threshold assumptions, we compare the dynamic returns to management under AAM, PAM and NLM to illustrate the investment dynamics. The pattern in Fig. 5 (left panel) is intuitive. Under AAM there is an initial period of "investment" in learning which involves costly deviations from the NLM and PAM strategies. Initially AAM profits are essentially zero since optimal harvest is at or near zero. As the payoff for AAM crosses the payoff curves for PAM and NLM, the relative dynamic benefits of an active learning strategy begin to accrue.

[26] We also assume in the analysis that the true mean productivity parameter for each Monte Carlo simulation ($\overline{S}^m$) is drawn from a beta distribution with a mean of 0.4 and concentration of 100 (i.e. 95% of the cumulative density lies in the range (0.3, 0.5)).

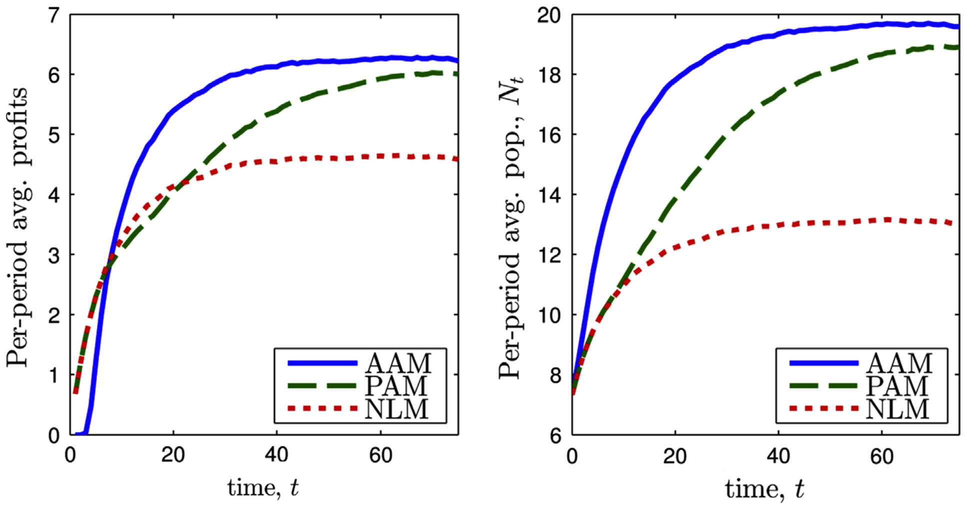

Fig. 5. Per-period average profits and population across Monte Carlo simulations for AAM, PAM and NLM. *Note*: We assume that ($\mu_0 = 0.05$, $\omega_0 = 1.3$, $N_0 = 7.4$).

In the right panel of Fig. 5, the average population level in each period illustrates how the AAM policy anticipates the value associated with learning once the stock recovers and thus speeds the recovery, even though each policy involves rebuilding stocks (on average). Before the observation threshold is reached the NLM and PAM belief states are identical and there is no difference in management. Around period 10, stock recovery under the PAM policy passes the observation threshold and the belief states of the NLM and PAM strategies diverge. In the long run, the passive learning approach arrives at a stock recovery similar to that of the AAM policy, albeit with substantial delay.

Finally, our results are sensitive to the assumed level of variation in the stochastic draw of the survivorship parameter S_t. Recall that as the concentration parameter (ρ) goes to infinity, the hierarchical model simplifies to the non-hierarchical model (S_t is constant) and learning is trivial (happens almost immediately). At the opposite extreme, under highly variable S_t, learning takes time and can be less valuable given the dominance of stochasticity in S_t. For example, in the case of $\rho = 1$ we do not find substantial differences in optimal AAM and PAM strategies.

4. Conclusion

We contribute to the application of AAM in natural resource economics by demonstrating the use of a hierarchical model for flexible representation of uncertainty, and by solving an otherwise intractable control problem using a density projection approximation method and standard dynamic programing techniques. We do so while at the same time relaxing constraints imposed in the existing AAM literature that limit the action space or belief space to a few discrete alternatives.

While we demonstrate the density projection (Kullback–Leibler divergence minimization) approach using a particular fishery model, the approach is general. Density projection has substantial flexibility, facilitating arbitrary belief states such as multimodal or heavy tail distributions. We have shown explicitly in our setting how implementing the KL divergence minimization dramatically simplifies to satisfying a pair of moment conditions. These concise conditions emerge not just for our example but for a broad set of applications from exponential family distributions.

In our setting the application of a hierarchical model of uncertainty resulted in intractable belief state dynamics, but there are many other ways in which similar problems of non-conjugacy and dimensionality can arise. For example, the density projection methodology could also be used to make tractable models in which learning occurs about an unknown resource stock (see, e.g., Kling and Sanchirico, 2012). As computational capacity expands, an important area for future exploration is the development of a fully general model of joint learning over an unknown stock *and* structure at the same time. We expect that this approach will expand the range of environmental problems for which credible policy analysis with learning can be conducted.

The AAM framework identifies the ideal level of investment in learning. We have shown conditions under which AAM provides minimal gains relative to a passive approach (PAM), as well as situations in which AAM is more likely to provide substantial returns. The latter case was driven by a differentially informative learning process. We also find that these returns will likely follow a period of costly short-run investments that could result in the costs and benefits of AAM being realized by different stakeholders. This potential equity issue is often not appreciated by advocates for AAM. It is important to also highlight that implicitly in our analysis the manager updates their beliefs based on monitoring and data collection. While we did not consider the cost of monitoring directly in this analysis, an interesting area of future work would be to consider optimal allocation of monitoring and fishing effort simultaneously. AAM, for example, might require more costly monitoring methods.

We also confirm previous counterintuitive findings that there exist some states in which the AAM action involves *less* learning than a non-learning approach (NLM and PAM). We provide a general explanation of this surprising result based on a decomposition of the stochastic and experimentation effects, as well as the interactions between the two effects.

Recent development of models for adaptive management, including this paper, have largely focused on a marginal learning process in which a decision maker attempts to fine-tune information about a particular uncertain component within a larger economic and/or biological structure that is fixed and given. In this setting, a fully active learning strategy does not always convey a substantial advantage relative to the passive adaptive management approach. A driving factor in relative performance will be whether the active adaptive management strategy generates information at either a substantially different rate or in a fundamentally different way. We conjecture that the additional returns to AAM will shift from significant but moderate in the present context to something much more substantial when the AAM methodology is trained on discerning between fundamentally different models of systems, or on generating entirely new hypotheses about how systems work.

Understanding when and where AAM will yield substantial returns raises an interesting political-economy question with respect to the current implementation of AAM. For example, despite the theoretical appeal of adaptive management, scholars argue that the approach has found influence more as "an idea than as a practical means of gaining insight" (Lee, 1999) and that the concept is "often appealed to but rarely achieved" (Doremus, 2007 p. 568). The degree to which such failures represent situations in which the returns to passive learning are unlikely to be exceeded versus opportunities for substantial improvement through active learning is an open question for further study.

Appendix A. Ecological and economic parameters used in the numerical dynamic programming solution

See Appendix Table A1

Appendix B. Validity of the KL divergence minimization approximation

We evaluated the validity of the D_{KL} minimization by verifying through Monte Carlo simulation that beliefs (updated using the approximation) converged over time to the truth as observations accrued. For each Monte Carlo simulation, $m=\{1,\ldots,M\}$, the true mean productivity was set equal to the same constant "true" level, $\overline{S}^m=\overline{S}^0$. For each simulation m, a series of annual productivity parameters, S_t, were drawn for $t=1,\ldots,T$, where $T=75$. For each time period t, actual survivorship was drawn from a binomial distribution given S_t and $Z_t=10$. For each simulation m, the decision-maker's initial beliefs were randomly drawn from a beta distribution with mean $\overline{S}_0$ and a wide spread (concentration of 1.3).

Over the time horizon, we find as expected that beliefs about mean productivity converge towards the true value. However, because learning takes place in a "noisy world" convergence is not perfect. For example, given a true level of $\overline{S}^0=0.3$ the average absolute value of deviations between the truth and beliefs (in percentage terms) declines from 85% to 5.5% over the 75 periods (or 0.25 to 0.02 in percentage point terms). Perfect convergence to the truth would be reflected by a deviation of 0%. We found by examining the deviation between $\overline{S}^0$ and the average of "actual" simulated survivorship (K_t/Z_t) over time that the lack of near-perfect convergence was driven by the variability in the system. In infinite time, this difference would converge to zero. Across 1000 simulations this difference was declining over time—after 75 time periods it was 5.1%. Thus, the key driver of imperfect convergence in finite time (e.g. 5.5% deviation) is the potential for instances in which the "actual" stochastic outcomes are significantly above or below the true mean (5.1% deviation).

Appendix C. Belief state dynamics and confidence

Fig. A1 presents an example of how the mean of the (approximate) posterior, $\hat{\mu}_{t+1}$, varies as a function of prior confidence (ω_t) and new information (K_t, Z_t). For example, with high confidence on the survivorship rate, new information has very little influence on the posterior belief on the mean survivorship rate. On the other hand, with low confidence, new information has a much larger impact on the posterior belief on the mean survivorship.

Table A1
Ecological and economic parameters.

	Parameter	*Level*	*Notes*
Ecology	b_1	1	Survival rate of juvenile recruits at low density
	b_1/b_2	10	Saturation rate of recruitment in each t
	Natural mortality rate, ψ	0.1	10% mortality of the adult standing stock in each t, $\overline{S}\in[0,1]$
	Scaling factor, θ	0.7	
	Larval production per adult, γ	1	If γ is greater than one, then larval production is increasing in the adult standing stock
Economics	Choke price, a_0	7	Vertical intercept of the demand curve
	Slope of demand curve, a_1	0.75	Slope of the demand curve, when harvest equals to a_0/a_1 the price is zero
	Harvesting costs, c	30	Cost per unit of harvesting, when holding the stock size constant
	Discount factor, δ	1/(1.05)	

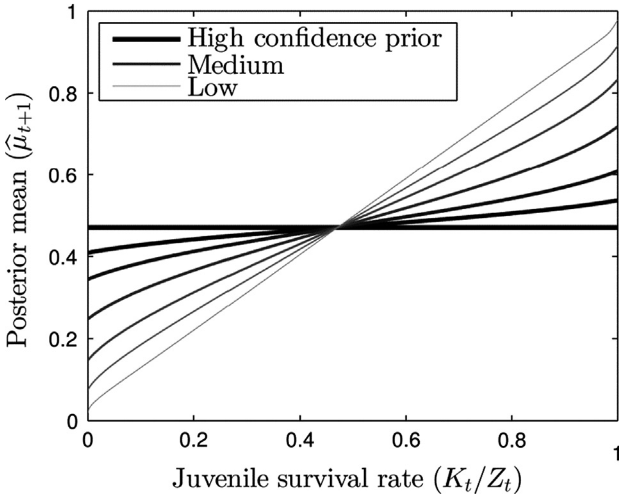

Fig. A1. Bayesian updating example: the mean of (approximate) posterior beliefs, $\hat{\mu}_{t+1}$, on the survivorship rate conditional on the proportion of juvenile survivors and different levels of prior confidence (ω_t) in the prior beliefs (Z_t=61, μ_t=0.47) .

Appendix D. Implementation and coding

The dynamic programming problem for each management strategy was solved using value function iteration (Judd, 1998) in Matlab (release 2012a) on a PC running Microsoft Windows 7.[27] The solution procedure was implemented as follows:

1. Using density projection, generate and store belief state transition arrays characterizing updated beliefs $\left(\hat{\mu}_{t+1}(\lambda_{t+1}), \hat{\omega}_{t+1}(\lambda_{t+1})\right)$ at nodes of the discretized space of the belief-outcome vector $\lambda_{t+1} = (\mu_t, \omega_t, Z_t, K_t)$. The run time for this one-time operation is approximately one month.
   a. Discretize the domain of the belief-outcome vector
      i. Discretize the belief state space (μ, ω). We use 28 non-uniform nodes in each dimension.
      ii. Discretize the observed outcome space (Z, K). We use for Z a set of integer nodes between zero and max(Z)=61 and for K a set of 128 non-uniform nodes between zero and max(Z).[28]
   b. For each node in the 4-dimensional belief-outcome-space, solve the non-linear system of two equations in (10) for the projected posterior parameters ($\hat{\mu}_{t+1}$, $\hat{\omega}_{t+1}$). We execute this loop using parallel processing.
      i. Use numerical quadrature to calculate the integrals$E_f\left(\ln(\overline{S})|\lambda_{t+1}\right)$ and $E_f\left(\ln(1-\overline{S})|\lambda_{t+1}\right)$, which are the left hand side terms in (10).
2. Use value function iteration to solve for the value and policy functions. We found that all models converged within 180 iterations or less.
   a. Setup:
      i. Specify exogenous parameter values and functional forms.
      ii. Discretize variables
         1. Belief state space (μ, ω). We use 28 non-uniform nodes in each dimension.
         2. Adult population state variable, N. We use a set of 25 nodes between 0 and the carrying capacity, where the distance between nodes is increasing in N to better represent curvature of value and policy functions.
         3. Harvest control variable, H. We use a set of 13 uniform spaced nodes between 0 and 6.
         4. Observed outcome, K. We use a set of 448 non-uniform nodes between 0 and max(Z).
      iii. Use the posterior belief state transition arrays calculated at nodes in the belief-outcome space $\left(\hat{\mu}_{t+1}(\lambda_{t+1}), \hat{\omega}_{t+1}(\lambda_{t+1})\right)$ from step 1 above to identify posterior belief transition arrays over nodes in the belief-outcome-*choice* space: $\lambda_{t+1}^{Z(N,H),K} = (\mu_t, \omega_t, Z_t(N_t, H_t), K_t)$.
      iv. Calculate array $N_{t+1}(N_t, H_t, K_t)$ using assumed functional forms.
      v. Calculate the expected probability of state transitions $\Pr(K_t|\mu_t, \omega_t, Z_t(N_t, H_t))$using numerical quadrature to integrate over beliefs. Developing the state transitions takes approximately 2 weeks of computer time.

[27] The PC operated with two quad-core processors (3.06GHz) and 96GB RAM.
[28] We generalize the binomial distribution to allow for non-integer (but still discrete) outcomes for K.

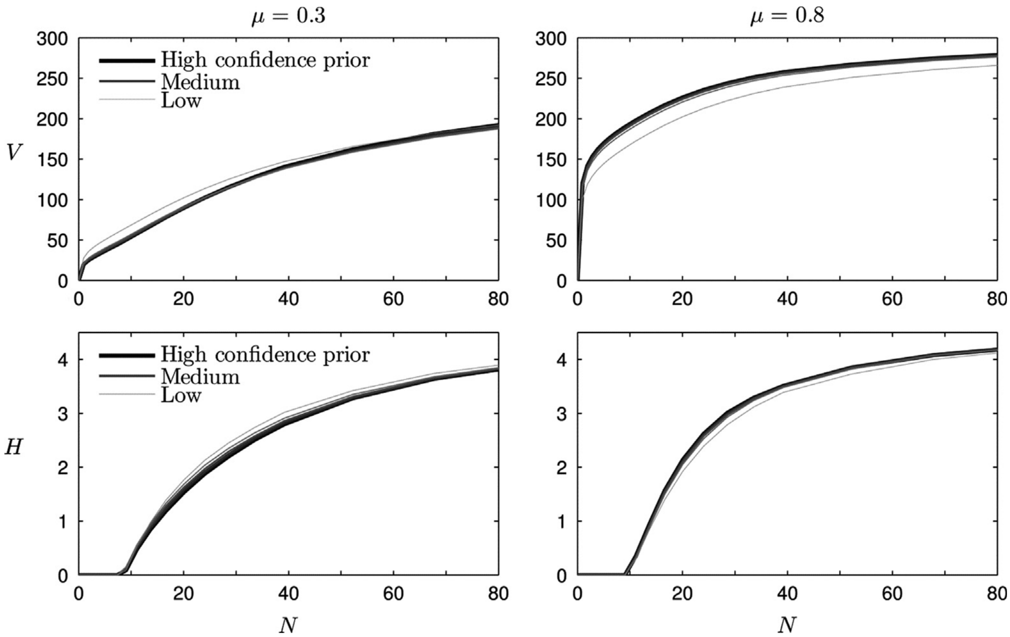

Fig. A2. NLM/PAM value (V) and policy (H) functions given beliefs about productivity—low μ_t=0.3 in left column, and high (μ_t=0.8) in right column—and different levels of confidence (ω) .

   vi. Specify initial guess for the value function $V(N_t, \mu_t, \omega_t)$. For the deterministic case the initial guess is set to zero and we use the deterministic solution as the initial guess for the NLM strategy. Finally, the NLM solution is used as the initial guess for the AAM strategy.

b. Value function iteration: Let V represent the starting estimate of the value function at the beginning of the iteration and V' the updated estimate at the end of the iteration. We iterate until max$\{|V - V'|\} \leq 2\text{x}10^{-4}$. Depending on the decision strategy, the run time for this step ranges from 2-14 hours.

   i. Interpolate to find V over a finer grid for K

   ii. Calculate the present expected maximized future value across the choice vector,$\delta E[V|\mu_t, \omega_t, Z(N_t, H_t)]$, using V and $\Pr(K_t|\mu_t, \omega_t, Z_t(N_t, H_t))$.

   iii. For each node in the state space (N, μ, ω)

      1. Calculate $v(H_t) = \pi(H_t|N_t) + \delta E[V|\mu_t, \omega_t, Z(N_t, H_t)]$ across the discretized domain of H.
      2. Interpolate (piecewise polynomial cubic spline) to find$v(H_t)$over the domain of H to identify the optimal continuous choice, H_t^*.
      3. Store the updated value for V' given H_t^*.

Appendix E. NLM/PAM value and policy functions

We present in Fig. A2 optimal value and policy function solutions for the NLM strategy under two belief states that include low expected productivity ($\mu = 0.3$) and high expected productivity ($\mu = 0.8$) as well as a range of confidence (ω) levels. The PAM strategy, by definition, employs the NLM strategy and therefore shares the same policy and value functions. While beliefs are updated over time under the PAM strategy, this updating is not anticipated by the decision maker. Conditional on a particular belief state, the expectations of the NLM and PAM decision maker are the same.

References

Aguirregabiria, V., Mira, P., 2010. Dynamic discrete choice structural models: a survey. Journal of Econometrics 156 (1), 38–67.
Bar-Shalom, Y., 1981. Stochastic dynamic programming: caution and probing. IEEE Transactions on Automatic Control 26, 1184–1195.
Berger, J.O., 1985. Statistical decision theory and Bayesian analysis. Springer, New York.
Bertsekas, D.P., 1995. Dynamic programming and optimal control. Athena Scientific, Belmont, MA.
Bond, C.A., Loomis, J.B., 2009. Using numerical dynamic programming to compare passive and active learning in the adaptive management of nutrients in shallow lakes. Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie 57, 555–573.
Bond, C.A., 2010. On the potential use of adaptive control methods for improving adaptive natural resource management. Optimal Control Applications and Methods 31, 55–66.
Brooks, A., Williams, S., 2011. A Monte Carlo update for parametric POMDPs. Robotics Research 66, 213–223.
Chen, M.H., Shao, Q.M., 1997. Performance study of marginal posterior density estimation via Kullback–Leibler divergence. Test 6 (2), 321–350.
Doremus, H., 2007. Precaution, science, and learning while doing in natural resource management. Washington Law Review 82, 547.

Doremus, H., 2010. Adaptive management as an information problem. NCL Review 89, 1455.
Eguchi, S., Copas, J., 2006. Interpreting Kullback–Leibler divergence with the Neyman–Pearson lemma. Journal of Multivariate Analysis 97 (9), 2034–2040.
Filatov, NM, Unbehauen, H., 2000. Survey of adaptive dual control methods. IEE Proceedings-Control Theory and Applications 147 (1), 118–128.
Gelfand, A.E., 2012. Hierarchical modeling for spatial data problems. Spatial Statistics 1, 30–39.
Gelman, A., Carlin, J., Stern, H., Rubin, D.B., 2004. Bayesian Data Analysis, 2 ed. Chapman and Hall/CRC, Washington, District of Columbia.
Gilboa, I., Marinacci, M., 2011. Ambiguity and the Bayesian Paradigm, Working Papers 379, IGIER, Innocenzo Gasparini Institute for Economic Research, Bocconi University.
Golan, A, Judge, G., Karp, L., 1996. A maximum entropy approach to estimation and inference in dynamic models or counting fish in the sea using maximum entropy. Journal of Economic Dynamics and Control 20 (4), 559–582.
Grimsrud, Kristine M., Huffaker, Ray, 2006. Solving multidimensional bioeconomic problems with singular-perturbation reduction methods: application to managing pest resistance to pesticidal crops. Journal of Environmental Economics and Management 51 (3), 336–353.
Halstead, B.J., Wylie, G.D., Coates, P.S., Valcarcel, P., Casazza, M.L., 2012. 'Exciting statistics': the rapid development and promising future of hierarchical models for population ecology. Animal Conservation 15 (2), 133–135.
Hartmann, K., Bode, L., Armsworth., P., 2007. The economic optimality of learning from marine protected areas. ANZIAM J 48(CTAC2006), C307–C329.
Hauser, C.E., Possingham, H.P., 2008. Experimental or precautionary? Adaptive management over a range of time horizons. Journal of Applied Ecology 45, 72–81.
Hudomiet, P., Willis, R.J., 2012. Estimating second order probability beliefs from subjective survival data (No. w18258. National Bureau of Economic Research.
Johnson, F.A., 2011. Learning and adaptation in the management of waterfowl harvests. Journal of environmental Management 92, 1385–1394.
Judd, K.L., 1998. Numerical methods in economics, Cambridge, MassMIT Press.
Kaplan, J.D., Howitt, R.E., Farzin., Y.H., 2003. An information-theoretical analysis of budget-constrained nonpoint source pollution control. Journal of Environmental Economics and Management 46 (1), 106–130.
Kling, David, and Sanchirico J.N.,2012. Taming the Lionfish. Working Paper. Department of Agricultural and Resource Economics, UC Davis. Available upon request.
Kruschke, John K., 2011. Doing Bayesian data analysis: a tutorial with R and BUGS. Academic Press, Elsevier.
Lee, K.N., 1999. Appraising adaptive management. Conservation Ecology 3 (2), 3.
Maybeck, P.S., 1982. Stochastic Models, Estimation and Control. Academic Press, NY, USA.
McDonald-Madden, E., Probert, W.J.M., Hauser, C.E., Runge, M.C., Possingham, H.P., Jones, M.E., Moore, J.L., Rout, T.M., Vesk, P.A., Wintle, B.A., 2010. Active adaptive conservation of threatened species in the face of uncertainty. Ecological Applications 20, 1476–1489.
Massey, D.M., Newbold, S.C., Gentner, B., 2006. Valuing water quality changes using a bioeconomic model of a coastal recreational fishery. Journal of Environmental Economics and Management 52 (1), 482–500.
Moore, A.L., McCarthy, M.A., 2010. On valuing information in adaptive-management models. Conservation Biology 24, 984–993.
Paté-Cornell, Elisabeth, M., 1996. Uncertainties in risk analysis: six levels of treatment. Reliability Engineering and System Safety 54, 95–111.
Provencher, B., 1995. Structural estimation of the stochastic dynamic decision problems of resource users: an application to the timber harvest decision. Journal of Environmental Economics and Management 29 (3), 321–338.
Reed, W.J., 1979. Optimal escapement levels in stochastic and deterministic harvesting models. Journal of Environmental Economics and Management 6, 350–363.
Rout, T.M., Hauser, C.E., Possingham, H.P., 2009. Optimal adaptive management for the translocation of a threatened species. Ecological Applications 19, 515–526.
Roy, N., Gordon, G.J., Thrun, S., 2005. Finding approximate POMDP solutions through belief compression. Journal of Artificial Intelligence Research (JAIR) 23, 1–40.
Royle, J.A., Dorazio, R.M., 2008. Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities. Academic Press, San Diego, California, USA.
Rust, J., 1994. Structural estimation of Markov decision processes. Handbook of Econometrics 4, 4.
Sanchirico, J.N., 2005. Additivity properties in metapopulation models: implications for the assessment of marine reserves. Journal of Environmental Economics and Management 49 (1), 1–25.
Sanchirico, J.N., Mumby, P., 2009. Mapping ecosystem functions to the valuation of ecosystem services: implications of species-habitat associations for coastal land-use decisions. Theoretical Ecology 2, 67–77.
Sanchirico, J.N., Springborn, M., 2011. How to get there from here: ecological and economic dynamics of ecosystem service provision. Environmental and Resource Economics 48 (2), 243–267.
Schaub, M., Kéry, M., 2012. Combining information in hierarchical models improves inferences in population ecology and demographic population analyses. Animal Conservation 15 (2), 125–126.
Sethi, G., Costello, C., Fisher, A., Hanemann, M., Karp, L., 2005. Fishery management under multiple uncertainty. Journal of Environmental Economics and Management 50, 300–318.
Seo, K., 2009. Ambiguity and second-order belief. Econometrica 77, 1575–1605.
Siniscalchi, M., 2008. Ambiguity and ambiguity aversion. The New Palgrave Dictionary of Economics. Palgrave Macmillan, New York.
Skyrms, B., 1980. Higher order degrees of belief. In: Mellor, D.H. (Ed.), Prospects for Pragmatism. Essays in Memory of F.P., Cambridge University Press, Ramsey. Cambridge, pp. 109–137.
Smith, M.D., Zhang, J., Coleman, F.C., 2008. Econometric modeling of fisheries with complex life histories: Avoiding biological management failures. Journal of Environmental Economics and Management 55 (3), 265–280.
Springborn, M., Costello, C., Ferrier, P., 2010. Optimal random exploration for trade-related non-indigenous. Bioinvasions and globalization: ecology. In: Perrings, C., Mooney, H., Williamson, M. (Eds.), Economics, Management, and Policy, Oxford University Press, Oxford, pp. 127–144.
Tahvonen, O., 2009. Economics of harvesting age-structured fish populations. Journal of Environmental Economics and Management 58 (3), 281–299.
Timmins, C., 2002. Measuring the dynamic efficiency costs of regulators' preferences: Municipal water utilities in the arid west. Econometrica 70 (2), 603–629.
Walters, C., 1986. Adaptive Management of Renewable Resources, Pub. CoMacMillan, New York, NY.
Wang, H., 2002. Minimum entropy control of non-Gaussian dynamic stochastic systems. IEEE Transactions on Automatic Control 47 (2), 398–403.
Weitzman, M.L., 2009. On modeling and interpreting the economics of catastrophic climate change. The Review of Economics and Statistics 91 (1), 1–19.
Wieland, V., 2000. Learning by doing and the value of optimal experimentation. Journal of Economic Dynamics and Control 24, 501–534.
Wikle, C.K., 2003. Hierarchical models in environmental science. International Statistical Review 71, 181–199.
Zhou, Enlu, Fu, M.C., Marcus, S.I., 2010. Solving continuous-state POMDPs via density projection. IEEE Transactions on Automatic Control 55 (5), 1101–1116. (May 2010).