Ecological Economics 104 (2014) 107–118

Contents lists available at ScienceDirect

Ecological Economics

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

Analysis

Modeling experiential learning: The challenges posed by threshold dynamics for sustainable renewable resource management

Emilie Lindkvist *, Jon Norberg

Stockholm Resilience Centre, Stockholm University, SE-106 91 Stockholm, Sweden

ARTICLE INFO

Article history:
Received 22 May 2013
Received in revised form 16 April 2014
Accepted 25 April 2014
Available online 21 May 2014

Keywords:
Adaptive management
Learning by doing
Renewable resources
Thresholds
Neural networks
Reinforcement learning
Agent based modeling

ABSTRACT

Adaptive management incorporates learning-by-doing (LBD) in order to capture learning and knowledge generation processes, crucial for sustainable resource use in the presence of uncertainty and environmental change. By contrast, an optimization approach to management identifies the most efficient exploitation strategy by postulating an absolute understanding of the resource dynamics and its inherent uncertainties. Here, we study the potential and limitations of LBD in achieving optimal management by undertaking an analysis using a simple growth model as a benchmark for evaluating the performance of an agent equipped with a 'state-of-the-art' learning algorithm. The agent possesses no a priori knowledge about the resource dynamics, and learns management solely by resource interaction. We show that for a logistic growth function the agent can achieve 90% efficiency compared to the optimal control solution, whereas when a threshold (tipping point) is introduced, efficiency drops to 65%. Thus, our study supports the effectiveness of the LBD approach. However, when a threshold is introduced efficiency decreases as experimentation may cause resource collapse. Further, the study proposes that: an appropriate amount of experimentation, high valuation of future stocks (discounting) and, a modest rate of adapting to new knowledge, will likely enhance the effectiveness of LBD as a management strategy.

1. Introduction

Exploiting a renewable resource sustainably involves two fundamental constraints: first, the difficulties of agreeing on appropriate actions (Dietz et al., 2003; Hardin, 1968) and second, the limitations in understanding the dynamics of the resource system (Allen et al., 2011; Armitage et al., 2008). This paper will focus on the latter constraint.

1.1. Adaptive Resource Management—The Learning Challenge

Ecosystems are complex adaptive systems (Levin et al., 2012) and thus will always be subject to uncertainty, unknown and unknowable phenomena (Duit and Galaz, 2008; Folke et al., 2005; Levin, 2003). These phenomena include, among others: inherent stochastic events, non-convex interactions, scale dependent dynamics and incomplete information (Crépin et al., 2012; Norberg and Cumming, 2008). Using optimal management (see e.g. Clark, 2010) and predefined models to describe complex resource dynamics with the purpose of deriving an optimal path of action for management or exploitation will always ignore parts of these phenomena and may give a false sense of accuracy in the derived solutions (Allen et al., 2011; Holling and Meffe, 1996; Walters and Hilborn, 1978). The shortcomings of optimization-based ‘command and control’ management led to the development of adaptive management (Allen et al., 2011; Holling, 1978; Holling and Meffe, 1996). This approach tries to navigate the system by more or less targeted trial and error and by building up a reservoir of knowledge through a continuous learning process (Arrow, 1962; Kolb, 1984; Walters and Holling, 1990). However, adaptive management also faces constraints such as: limitations in building an understanding of the resource dynamics based on iterative and locally based experiences, the cost of experimenting in order to learn about the system and, recognizing and responding to knowledge of changing conditions within the system (Olsson and Folke, 2001).

In this paper we study adaptive management using ‘optimal control management’ as a benchmark, to explore the limits of learning-by-doing (LBD) when managing a renewable resource exhibiting two levels of non-linear dynamics.

1.2. Thresholds Dynamics

A particularly ‘wicked problem’ (Jentoft and Chuenpagdee, 2009) in natural resource management is the fact that many ecosystems are subject to threshold dynamics, so called critical transitions or regime shifts (Scheffer, 2009). Threshold effects entail abrupt changes in the dynamics of an ecosystem, where the effect of passing a threshold (or tipping point) switches the dominant feedbacks within the system and can change a particular resource's provisioning rate. Recovery of the resource is then constrained by the degree of lock-in to the new system domain

* Corresponding author. Tel.: +46 73 9835269.
E-mail addresses: emilie.lindkvist@su.se (E. Lindkvist), jon.norberg@su.se (J. Norberg).

http://dx.doi.org/10.1016/j.ecolecon.2014.04.018
0921-8009/

i.e., the hysteresis effect (Scheffer et al., 2001). Thresholds may be seen as undesirable properties, if the change in provisioning from the system has a significantly negative impact on human well-being (Rockström et al., 2009; Stern, 2007). Moreover, research suggests that regime shifts in human-nature systems are likely to increase as human pressure on the planet accrues (Schlüter et al., 2012).

Depending on the severity of lock-in, thresholds provide a particularly difficult challenge for LBD approaches, should a threshold be crossed while experimenting to learn the dynamics of the system.

Recent research is focused on anticipating regime shifts (Biggs et al., 2009; Scheffer et al., 2012), optimal management of systems with regime shifts (Brock and Starrett, 2003; Horan et al., 2011; Polasky et al., 2011), or studying adaptive management in relation to threshold dynamics, using either multi-agent based models for studying emergent properties (Janssen and Carpenter, 1999; Janssen et al., 2000) or laboratory experiments with human subjects (Lindahl et al., 2012). However, the role of threshold dynamics for the LBD processes itself, is in dire need of further exploration.

1.3. Introducing Reinforcement Learning and Artificial Intelligence

Decision theories are strongly connected to the learning process, such as the *expected utility theory* from economics (Neumann and Morgenstern, 1947) and *prospect theory* from psychology (Kahneman and Tversky, 1979). Reinforcement learning (RL) (Sutton and Barto, 1998) is a computational approach to problems concerning goal-directed learning. It is defined as an agent's ability to learn a behavior through trial-and-error by interacting with a dynamic environment (Kaelbling et al., 1996), and allows for incorporating different decision theories. Due to its compatibility with adaptive management, RL is perceived as a latent approach for dealing with natural resource management problems (Fonnesbeck, 2005). RL attracts researchers from diverse disciplines such as psychology, control theory, artificial intelligence, and neuroscience (Sutton and Barto, 1998). Notably, Kable and Glimcher (2009) and Niv and Montague (2008) show that it can reveal the neurobiological basis for learning subjective values, which ultimately underlies all decision-making.

RL and neural networks are situated between artificial intelligence and conventional engineering, and "… extend ideas from optimal control theory and stochastic approximation to address the broader and more ambitious goals of artificial intelligence" (Sutton and Barto, 1998). Control theory has contributed to a profound understanding of why complexity in natural systems creates trade-offs between robustness and resilience, and fragility at different scales (Anderies et al., 2007; Csete and Doyle, 2002; Folke, 2006; Levin and Lubchenco, 2008). Furthermore, it provides a well-founded mathematical representation of 'feedback' as a process. The RL approach shifts the main focus from control to learning, and accentuates highly theoretical but essential parts of the LBD process—moving it closer to how humans handle information. By combining the component ideas of temporal difference learning (from RL) and neural networks, we incorporate features such as hindsight, planning horizon, exploration vs. exploitation and generalization (further described in Section 2.2).

Learning is a principal aspect of adaptive management in helping to deal with uncertainty and change (Armitage et al., 2008). Social learning has been extensively studied, but the learning process at an individual level has been scarcely addressed in resource management (see e.g. Fazey et al. (2005), Garavito-Bermúdez et al. (in press) and Marschke and Sinclair (2009) for empirical studies). In addition, literature on how the human brain accumulates knowledge through interacting with its environment is not explicitly found in resource management. However, inspiration can be gained from the relatively new field of *neuroeconomics*, which tries to understand the neurological basis of how the human reward system affects behavior (Rangel et al., 2008).

1.4. Research Questions

In this study we undertake an analysis letting an AI agent learn to manage a renewable resource with two levels of complexity (with or without threshold dynamics). By using this method we can parameterize the LBD process and its related components, and thus we are able to evaluate key learning parameters in relation to optimal control performance. Hence, for studying the limitations and possibilities of LBD for sustainable management of renewable resources, we probe the following questions;

- How does the LBD process respond to different levels of complexity of a renewable resource?
- How do key learning parameters of the LBD process—such as mental model update rate, discounting, hindsight, and experimentation—influence management outcomes?
- Is there a discrepancy between the optimal values of the key learning parameters, depending on the level of complexity of the resource?

The key learning processes are selected in accordance with learning literature on LBD (e.g. Kolb, 1984), within natural resource management literature (e.g. Armitage et al., 2008; Duit and Galaz, 2008; Ostrom, 1990), and the outline of the RL method. Obviously, learning is a much richer phenomenon than we are able to depict here. However, our aim is to analyze the core of the LBD-process, which confines learning in this study. A similar work, using RL, is conspicuously lacking and we aim to provide useful insights to the discourse on the role of learning and decision-making in natural resource management.

2. The Model

To make the model less abstract we can envision the setting to be a fishery management problem, where the agent represents a fishing unit (such as a fisherman or an organization with full ownership rights having the ability to exclude other actors, i.e., not a common pool situation), that interacts with a fish stock (where the fish is assessed as a single-species unit). For such situations, ample theory has been developed (Clark, 2010). A list of terminology is provided in Table 1. No a priori understanding of the resource system was given to, nor built into, the agent in order to study the complete learning process (which would theoretically be the case when exploring a new system). Instead, for each fishing event, the agent could set its harvest effort, observe the harvest, and learn from this experience. The agent was analyzed interacting with either a resource characterized by a logistic growth rate, or a similar resource but with a threshold in its regeneration rate. The two scenarios will further be referred to as the *logistic function* and the *threshold function*, to state which dynamics are in focus.

2.1. Resource Dynamics and Agent's Maximization Problem

First, let us describe the agent's goal function and the two resource functions. The goal of the agent was to find the effort resulting in the maximum economic yield (MEY) over time. The agent acted as a price-taker in a competitive market, and thus lacked market power.

Table 1
Terminology and selected parameters used interchangeably throughout the paper.

Model term	Fishery term
Agent	Fisherman/fishing unit[a]
Action (a)	Effort/harvest effort
State (s)	Biomass/stock
Reward (r)	Net income of harvest
Time step (t)	Fishing event

[a] A centrally organized unit of fishermen.