Optimal Monitoring of Agri-environmental Schemes

Ben White

Paper presented to the Australian Association of Agricultural and Resource Economics Conference, Canberra, February 12 to 15, 2002.

School of Agricultural and resource Economics,University of Western Australia, Perth,

Crawley WA 6009. E-mail: bwhite@agric.uwa.edu.au

Agri-environmental schemes are found in most European countries and now account for approximately 4 per cent part of EU expenditure on UK agriculture. A significant part of that expenditure is the cost of monitoring farmer compliance with input restrictions. This paper analyses the design of monitoring schedules for long duration agri-environmental schemes where the aim is to reinstate preferred ecosystems using a Partially Observed Markov Decision Process (POMDP). The approach has much in common with the Arrow-Fisher-Henry model of irreversible land development where there is uncertainty over environmental value. Uncertainty in the Partially Observed Markov Decision Process (POMDP) model analysed here, relates to the current vegetation state, stochastic transitions between vegetation states and monitoring accuracy. It applies to both irreversible changes and changes subject to varying degrees of reversibility. Results from this model present a scheme for monitoring which depends upon the regulators prior probability of vegetation states. Over time, monitoring resolves the uncertainty the regulator has about the vegetation state. For some prior probabilities monitoring is repeated for a number of periods for others no monitoring or one period of monitoring is optimal

Keywords: biodiversity, conservation, irreversibility, MDP, POMDP

JEL Classification: Q0, Q2, C6

1. Introduction

The last twenty years have seen a rapid expansion of the agricultural area in the UK entered into agri-environmental schemes. These include national schemes such as the Sites of Special Scientific Interest and schemes partly funded by the EU under Regulations 2078/92 and 746/96 (European Union, 1992, 1996). Over the period 1993 to 1997 compensation payments totalled £378 million for the 410,000 ha entered into the Environmentally Sensitive Areas (ESA) Scheme and the 92,600ha in the Countryside Stewardship Scheme (Falconer and Whitby, *1999*, p64). Both schemes compensate farmers for restricting input use and reducing the extent of certain land uses. Public accounts for the ESA and Countryside Stewardship schemes include monitoring costs under the general administrative costs schemes. A recent study (*op cit*, p94) estimates administrative cost in the UK at 48 per cent of compensation payments, an average of £45 million per annum for 1993 to 1996. A significant part of this cost is due to monitoring (National Audit Office, 1997).

The principle aim of these schemes is to increase biodiversity in semi-natural vegetation: as such agri-environmental schemes procure public capital goods in the form of increased biodiversity. However, measuring progress towards a 'target' vegetation is costly and prone to errors (Hooper, 1992). Monitoring plays two roles: first it ensures that the vegetation is following the expected transition: second it acts as an incentive for farmers to comply with input restrictions as they run the risk of losing income if they are ejected from the scheme.

The next section describes the model in a number of stages. The first sub-section reviews the Markov processes of vegetation transitions and how they might be used to measure the degree of reversibility associated with a vegetation change. Second the Markov Decision Process (MDP) is described as a component of the partially observed Markov decision process (POMDP). Section 3 applies the POMDP to agri-environmental schemes. Example 1 is for a two period planning horizon and it establishes a range of different valuations for a scheme depending upon assumptions about the time when uncertainty about the vegetation state resolves. Example 2 is a monitoring study for a farm in the Cambrian Mountains Environmentally Sensitive Area in Wales it uses an approximate infinite horizon solution to present the results. Section 4 concludes.

2. The Model

The literature on irreversible environmental change under uncertainty commences with Arrow and Fisher (1974) and Henry (1974) (also see Fisher 2000 for a recent review). The Arrow, Fisher and Henry (AFH) model identifies the option value generated by delaying an irreversible development so that uncertainty about the environmental value of land resolves. The decision problem analysed here has similarities with the AFH model, but also introduces a difference in emphasis and a number of generalisations. First the uncertainty in the model derives from the stochastic process which governs vegetation change not uncertainty over environmental value. In most cases there is uncertainty over the value of enhanced plant biodiversity, but this is not the focus here. Second, unlike the change in land use in the AFH model, changes in natural vegetation due to agriculture may vary from immediately reversible to completely irreversible depending upon the stochastic process which governs transitions between vegetation communities. Third an implicit assumption of the AFH model is that uncertainty resolves without cost through time, in fact the regulator is uncertain, without actively monitoring, as to which vegetation state the land is in. Fourth, ecological monitoring tends to be inaccurate, in that the vegetation state can be misclassified (Hooper, 1992). Fifth, the original AFH model was for two-periods. Changes in the vegetation state can be very gradual, for instance reinstating heather moorland may take ten to twenty years to reinstate (Ball *et al*, 1978). For this reason agri-environmental schemes are typically for five years often with provision for further extensions. Over the planning horizon there are opportunities for the regulator to discontinue agreements, monitor, start or restart agreements.

This paper shows how the regulation of an agri-environmental scheme can be represented by a Partially observed Markov Decision Process (POMDP). The key assumption is that the transition between vegetation states is a Markov process. The advantage of this assumption is that Markov transition matrices are widely used in applied ecology (Horn 1976; Balzer, 2000) and can be estimated from field studies (Rushton *et al,* 1996) and algorithms are available to solve Multi-period POMDP problems (Smallwood and Sondik 1973; Monahan 1982 and Cassandra, 1995).

2.1 Markov Process for Vegetation Dynamics

This Section gives a brief review of those aspects of Markov processes relevant to describing vegetation dynamics (see Ross (1999) for a general review and Balzar (2000) for a review of applications to ecology). A Markov model of vegetation dynamics defines a finite number of vegetation *states*. If $s_t=i$ the process is in state i at time t and there is a fixed transition probability p_{ij} that it will be in state j at $t+1$. The transition probabilities satisfy:

$$p_{ij} \geq 0 \qquad \sum_{j=1} p_{ij} = 1 \qquad i,j = 0,1,...,I$$

Where P is a matrix of transition probabilities. Given the initial ($1xI$) vector of state probabilities π_0 we can calculate the probability for m-steps ahead:

$$\pi_m = \pi_0 (P)^m.$$

Vegetation states are *recurrent*, if a process is in state i and will return to i, *transient* if the process is in state i there is a positive probability of the process never returning to i or *absorbing* where $p_{ii}=1$. Define f_{ij}^m as the probability of the first passage from state i to state j after m-steps. This can be calculated recursively as:

$$f_{ij}^m = p_{ij}^m - \sum_{k}^{m-1} f_{ij}^k p_{ij}^{m-k}$$

If we define:

$$f_{ij} = \sum_{m=1}^{\infty} f_{ij}^m$$

then f_{ij} denotes the probability of ever making a transition into state j, given that the process starts in i. Thus a state is recurrent if $f_{jj} = 1$ and transient otherwise. In ecology the absorbing state represents a climax vegetation community where there is no further succession, while a transient vegetation state is a part of a succession towards the 'climax' community.

Transition probability matrices also provide a measure of the degree of reversibility for a vegetation transition. If the first transition time is infinite then a state change is irreversible, if it is finite then a state change is reversible. Mathematically:

$$\mu_{ij} = \begin{cases} \infty & if \sum_{n=1}^{\infty} f_{ij}^{n} < 1 \\ n\sum_{n=1}^{\infty} f_{ij}^{n} & if \sum_{n=1}^{\infty} f_{ij}^{n} = 1 \end{cases}$$

where μ_{ij} denotes the expected time from state i to state j and is used here as a measure of reversibility.

2.2 Partially Observed Markov Decision Process

A POMDP model is a generalisation of a Markov decision process (MDP), originally developed by Bellman (1957) and Howard (1960). A MDP consists of a set of states s_i, a set of actions a_i and a reward structure defined for each state-action pair. Actions form policy vectors δ_n where each element is a function returning the action to be performed in each state. Thus $\delta_n(s_i)$ gives the action to take in state s_i. For example, actions may include *continue* an agri-environmental scheme or *stop*. Net social benefits depend upon the vegetation state and the action: the action *continue* generates a non-market benefit in terms of preserving a valued vegetation state and cost in terms of the social cost of farmer compensation and the net value of lost agricultural production; the action *stop* involves benefits in terms of increased agricultural production and reduced public expenditure and a non-market cost in terms of land switching to a less valued vegetation state. Section 3.1 describes the objective function in more detail.

Selecting the optimal policy is a problem is stochastic dynamic programming Bellman's equation is:

$$V_n^{\delta_n}(s_i) = \sum_j p_{ij}^{\delta_n(s_i)}[w_{ij}^{\delta_n(s_i)} + \beta V_{n-1}^{\delta_{n-1}}(s_j)]$$

$V_n^{\delta_n}(s_i)$ gives the present-value of following a policy δ_n over the stages remaining in the planning horizon, *n*. The term $w_{ij}^{\delta_n(s_i)}$ gives the reward for following action $\delta_n(s_i)$ while in state s_i and moving to state s_j where $p_{ij}^{\delta_n(s_i)}$ is the probability of moving to state *j* from state *i* given action $\delta_n(s_i)$. Note that the transition probabilities depend upon the action taken. The term β is a discount factor. The optimal policy δ_n^* is one where for any state $V_n^{\delta_n^*}(s_i) \geq V_n^{\delta_n}(s_i)$.

The partially observed Markov decision process, POMDP, generalises a MDP to allow for uncertainty about the current state. For an agri-environmental scheme, this means the regulator bases decisions upon the expected vegetation state and engages in monitoring to reduce uncertainty. POMDP extends the MDP model by including a set of observations Θ and an observation probability matrix R^a, where $r_{j\theta}^a$ gives the probability that we observe θ when in state s_j at n when our last action was *a*. Thus the matrix R^a gives the accuracy of monitoring. The immediate reward for taking an action in a particular state is $w_{ij\theta}^a$ this gives the reward for taking action *a* in state s_i, moving to state *j* and observing θ. In this form of the POMDP model observations are made after actions are taken. This allows time to complete monitoring.

Before Sondik (1974) developed a solution algorithm, even small POMDP proved difficult to solve. Sondik's approach was to define a *belief state* which is a vector of state probabilities $\pi = \{\pi_0, \pi_1, \ldots, \pi_I\}$ where *I* is the number of states. After each action and observation the belief state is updated by Bayes' rule:

$$\pi'_j = \frac{\sum_i \pi_i p_{ij}^a r_{j\theta}^a}{\sum_{i,j} \pi_i p_{ij}^a r_{j\theta}^a} = T(\pi \mid a, \theta)$$

where *T(.)* is the posterior probability of observing state *j* given observation θ. In a MDP model, a policy is a mapping from states to actions and since the number of states is finite, then the optimal policy can be determined by recursion. The solution to a POMDP, problem maps *belief states* into actions, but instead of having a finite number of *belief states* each element in π can take any value over the range $0 \le \pi_i \le 1$ ∀i subject to $\sum_i \pi_i = 1$. If we define:

$$q_i^a = \sum_{j\theta} p_{ij}^a r_{j\theta}^a w_{ij\theta}^a$$

The value function is given by:

$$V_n^*(\pi) = \max_a \sum_i \pi_i q_i^a + \sum_{i,j,\theta} \pi_i p_{ij}^a r_{j\theta}^a V_{n-1}^*[T(\pi \mid a, \theta)]. \qquad (1)$$

The optimal value function $V_n^*(\pi)$ over belief states is made up of the expected immediate reward, the first term on the right hand side, and the expected reward for future periods, the second term. The approach to solving this problem relies on the fact that there are a finite number of actions and each action generates a vector α^k of expected rewards across the *I* states. The solution involves finding a set of vectors which are optimal for some *belief state*. In fact $V_n^*(\pi)$ is piecewise linear and convex where the vectors α^k represent the line segments or hyperplanes (Sondik and Smallwood, 1973). Thus the value function can be represented by:

$$V_n^*(\pi) = \max_k \sum_i \pi_i \alpha_i^k(n)$$

where each vector $\alpha_i^k(n)$ gives the rewards for a particular action in state *i*. Using the vector notation (1) becomes (Sondik and Smallwood, 1973):

$$V_n^*(\pi) = \max_a \sum_i \pi_i \left[q_i^a + \sum_{i,j,\theta} \pi_i p_{ij}^a r_{j\theta}^a \alpha_j^{\iota(\pi,a,\theta)} \right]$$

where

$$\iota(\pi,a,\theta) = \arg\max_k \left[\sum_{i,j,\theta} \pi_i p_{ij}^a r_{j\theta}^a \alpha_j^k \right]$$

that is $\iota(\pi,a,\theta)$ is an index of vectors that maximises $V_{n-1}^*(T(\pi \mid a,\theta)$. The solution to POMDPs involves (Cassandra, 1995) generating vectors starting at *n=1*, and checking that each new vector is non-dominated, that is, represents an optimal solution for a subset of belief states.

3. POMDP Examples

3.1 Objective Function

This section applies the POMDP model to an agri-environmental scheme. The immediate payoff to the regulator is given by:

$$w_{ij\theta}^a = g_{ij\theta}^a + m_{ij\theta}^a - \lambda b_{ij\theta}^a - (1+\lambda) c_{ij\theta}^a (R^a)$$

where $g_{ij\theta}^a$ is the non-market benefits and $m_{ij\theta}^a$ market benefits, less the social cost of public funds as the transfer payment $b_{ij\theta}^a$, weighted by the shadow price λ and the social cost of monitoring $(1+\lambda) c_{ij\theta}^a (R^a)$.

Compensation is set to satisfy the individual rationality constraint of the farmer, in the empirical examples it is equal to the profit foregone. This form of the objective function (1) represents a cost benefit analysis framework, where transfer payments represent a cost in terms of the deadweight loss due to taxation (Laffont and Tirole, 1993). Monitoring costs are a resource cost, but must be covered by taxation so these costs are weighted by $(1+\lambda)$.

Monitoring cost is given by the function $c^a_{ij\theta}(R^a)$. The likelihood matrix R^a represents the accuracy of forecasting. Monitoring effort can result in a matrix which lies between two extremes: an identity matrix implies monitoring is perfectly accurate; conversely a uniform matrix where all elements equal *1/I* is uninformative. The cost of monitoring is defined as a function of the observation probabilities, in general C(*R*). Assume $c^a_{ij\theta}(R^a)$ takes a maximum value when *R* is an identity matrix and $c^a_{ij\theta}(R^a)=0$ when *R* is uniform distribution. Further assume that the cost of monitoring only depends upon the diagonal elements of *R*. Thus, in the absence of other information we assume that:

$$c^a_{ij\theta}(R^a) = diag(R^a)c_m$$

where $diag(R^a)$ gives the diagonal of matrix R^a and c_m is a cost coefficients. In a further simplification R^a can be represented by a single scalar value where all the diagonal elements are identical $r^a_{ii} = r \;\; i = 1,...,I$ and all off diagonal elements are give by *(1-r)/(I-1)*. This is a strong assumption as it implies that the probability of misclassification is the same for all states. In practice mistaken identification is more likely between similar vegetation states. However, this assumption simplifies the sensitivity analysis presented in Section 3.4.

3.2 Example 1 a Two-period Model

A scheme has been designed to reinstate a traditional hay-meadow (HM) from agricultural grassland (AG) by compensating farmers for reducing cattle and sheep stocking rates. The scheme runs for two years. The social net benefit of actions in the two vegetation states are given in Table 1.

Table 1 about here

The regulator can take one of three actions: *continue* ($a=1$) the scheme without monitoring, *stop* ($a=2$) the agreement and third *monitor* compliance ($a=3$). The effects of these policies is represented through changes in the transitional probabilities P^a and the observation probabilities R^a. Where the decision is to *continue* the transition probabilities are

$$P^1 = \begin{bmatrix} p_{11}^1 & p_{12}^1 \\ p_{21}^1 & p_{22}^1 \end{bmatrix} = \begin{bmatrix} 0.8 & 0.2 \\ 0 & 1 \end{bmatrix}$$

HM is denoted by subscript 1 and AG by subscript 2. If the regulator decides to stop the land converts to AG with certainty, thus the effective transition matrix is

$$P^2 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$$

in other words the subsequent state is AG with certainty. *Monitor* transition probabilities are the same as for *continue* thus $P^3=P^1$.

The matrix of observation probabilities for the *monitor* action is:

$$R^3 = \begin{bmatrix} r_{11}^3 & r_{12}^3 \\ r_{21}^3 & r_{22}^3 \end{bmatrix} = \begin{bmatrix} 0.85 & 0.15 \\ 0.15 & 0.85 \end{bmatrix}$$

For instance, if the state is HM there is a 0.85 probability of observing heather but a 0.15 probability of mistakenly observing AG. Where the actions are *continue* or *stop* then R^a is uniform. These actions provide no further information on the vegetation state than is obtained from the transition probabilities.

The solutions to the model depend upon the information available to the regulator. If we assume that the regulator knows the vegetation state at each stage, then the solutions is an MDP. If the land starts in the hay meadow state HM then we have the following problem:

$$V_2^1[s_1] = q_1^1 + \beta(p_{11} \max_a\{ q_1^a\} + p_{12} \max_a\{ q_2^a\}) \tag{2}$$

note that the regulator can chose the action which maximises the payoff in each state at state n=1. In this example this is equivalent to a POMDP model where the initial *belief state* is that $\pi=1$ (where π is defined here as the probability of the initial state being HM) and monitoring is perfect and at zero cost.

If the regulator is myopic and takes no account of the additional information revealed over the first period, or no further information is avaialble over the first period then the best the decision maker can do is:

$$V_2^2[s_1] = q_1^1 + \beta \max_a\{ p_{11}q_1^a + p_{12}q_2^a\} \tag{3}$$

this is equivalent to the case where monitoring is uninformative. The difference between (2) and (3)

$$V_2^1[s_2] - V_2^2[s_2] \geq 0$$

from the convexity of the maximum function and Jensen's Inequality. This difference is the option value (Albers *et al*, 1996). From Table 2, the option value for case *(a)* is £95=£2773-£2678

Table 2 about here

For POMDP case *(b)* the regulator can only obtain information on the vegetation state by monitoring. The decision to *continue* when there is uncertainty about the initial state is given by:

$$V_2^1[\pi] = \pi q_1^1 + (1-\pi)q_2^1 + \beta \max_a \{\pi(p_{11}q_1^a + p_{12}q_2^a) + (1-\pi)(p_{21}q_1^a + p_{22}q_2^a)\} \tag{4}$$

In case *(b)* π=1 thus (4) simplifies to (3) and the POMDP solution (*b1)* is equivalent to a myopic MDP solution (*a2)* , see Table 2. For this belief state it is optimal to continue without monitoring.

The expected value of monitoring is given by:

$$\begin{aligned} V_2^3[\pi] = \pi q_1^3 + (1-\pi)q_2^3 + \beta\pi(p_{11}r_{11}q_1^{\iota} + p_{11}r_{12}q_1^{\iota} + p_{12}r_{21}q_1^{\iota} + p_{12}r_{22}q_1^{\iota}) + \\ (1-\pi)(p_{21}r_{11}q_1^{\iota} + p_{21}r_{12}q_1^{\iota} + p_{22}r_{21}q_1^{\iota} + p_{22}r_{22}q_1^{\iota})\} \end{aligned} \tag{5}$$

where q_i^{ι} indicates that the regulator chooses the best action for the state observed. In case *b* equation (5) simplifies to (2). Notably the MDP solution with perfect information *a2* is equivalent to the POMDP solution with perfect monitoring *b2* after an adjustment for monitoring costs.

Cases *c* shows the net social benefit per ha when there is an equal probability that the land is in either state. Case *d* illustrates the point that the expected return to monitoring is reduced as the monitoring accuracy is reduced, *d1* should be compared with *c2*. In this *belief state* it is optimal to monitor, this will have the effect of reducing the uncertainty about the vegetation state. If *θ=1* it is optimal to *continue*, if *θ=1* it is optimal to *stop*.

There are two ways of describing POMDP solutions. First, where there are three or less states the initial optimal action can be shown as a plot between the initial *belief* state and the optimal intitial vector $\alpha_j^{\iota(\pi,a,\theta)}$. Second, as a sequence of decisions in a policy graph. Figure 1 gives the expected value of the agri-environmental scheme. Where the prior probability of HM is low, 0.309≥π *stop* is optimal. Over a range of probabilities

$0.309 \leq \pi \leq 0.412$ *monitor* is optimal. The value of monitoring is the difference between the expected value of monitoring and the next best policy, the value of monitoring depends upon the *belief state*. For $\pi \geq 0.412$ *continue* is optimal. Figure 2, a policy graph, gives a sequence of optimal actions. It is optimal to *continue* at n=2 it is optimal to *continue* at n=1. If it is optimal to *stop* at n=2 it is optimal to *stop* at n=1. The decision taken after monitoring at n=2 depends upon what is observed, if HM is observed (θ=1) then it is optimal to continue at n=1 if AG is observed (θ=2) it is optimal to *stop*.

Figure 1 and Figure 2

3.3 Example 2 Cambrian Mountains Environmentally Sensitive Area

The Cambrian Mountains Environmentally Sensitive Area (ESA) aims to reinstate the heather dominated vegetation community H12 in Mid-Wales from U4 (low intensity agricultural grassland) and an intermediate community H18. The classifications refer to the British National Vegetation Classifications (Rodwell, 1991, 1992) The transition matrix P^1 in Table 3 was estimated from data collected from a 5-year field scale experiment at Pwllpeiran experimental husbandry farm in Mid-Wales (Rushton et al, 1996). Without monitoring the long-run equilibrium is one where there is a probability of all vegetation states being observed, but the highest probability is that the land will revert to U4. It is assumed that monitoring influences the behaviour of farmers so that they comply with input restrictions, thus the transition probabilities are modified so that H12 is an absorbing state and the probability of transitions from H12 to H18 and from H18 to U4 are zero. The natural ecological succession at low sheep grazing intensity is from U4 to H18 to H12. Monitoring involves an annual assessment of vegetation by field survey and observing farm activities at key times during the production cycle. In practice the Ministry of Agriculture engages in monitoring using field surveys and aerial photographs to assess vegetation states and regular visits by project officers to assess compliance (Hopper, 1992). The assumption her is that a farmer knows that monitoring is taking place and during that year is compliant with the terms of the scheme. In years where there is no monitoring they are only partially compliant. The effect of monitoring is reflected by an *ad hoc* adjustment to the transition matrix.

Table 3 and Table 4

All three P^a matrices are ergodic, that is they converge to a single long-run matrix. Where there is no monitoring the expectation is, irrespective of the initial state, the U4 vegetation state will dominate. With monitoring the desirable state H12 is an absorbing state.

The economic parameters for the problem are given in Table 4, the estimates of non-market value are modified from a contingent valuation study by Garrod and Willis (1994). The measure of irreversibility μ_{ij}^{a} gives the expected first passage of time from state *i* to *j*. In this example we are interested in the expected time taken to return to H12. With no monitoring, the expected return time from U4 to H12 is μ_{13}^{1}=105.88 years and from H18 to H12, μ_{23}^{1}=55.88 years. This represents a high degree of irreversibility. With monitoring the expected return time is reduced to $\mu_{13}^{2} = 18.137$ and $\mu_{23}^{2} = 2.94$.

The solution to this POMDP problem consists of 26 vectors $\alpha^{k}(n)$ each of which is optimal for a range of belief states. The solution converged after 493 year long steps and approximates an infinite time horizon solution. If contracts with farmers are of short duration it is also possible to run the problem for a finite horizons. In fact the sub-stages of the infinite horizon problem provide the optimal decisions for finite horizon problems.

However, the advantage of an infinite horizon solution is that the optimal vectors are time dependent.

Table 5

Figure 3 represents the optimal vectors given in Table 5 as a function of *belief states* using a triangular plot, where the optimal initial action is given for a fine grid of probabilities (0.02 intervals). This locates where most vectors are optimal, but not all, some vectors are optimal for such a small range of probability values that the grid misses these points. Overall, the *belief state* is divided into three regions, a *continue* region a *monitor* region and a *stop* region. The *continue* region is where the optimal initial action is to *continue* and this includes vectors 1 to 6 and 23 to 25. From Figure 4, this region falls in the upper left hand segment of the diagram where the probability of H18 and H12 together are greater than the probability of U4. The *stop* region is defined by a single vector 26 and this action is optimal where the probability of H18 and U4 together exceed that of H12. The *monitor* region where vectors 7 to 22 are optimal is found where the probability of each state is equal, but extends back to the p1 to p3 axis.

Figure 3 and Figure 4

Figure 3 gives the initial actions which are optimal across belief space, but vectors which have the same initial actions may lead to quite different sequence of actions. The sequence of actions can be traced in Vector n-1 section of Table 3. For instance, vector 1 is followed by vector 3 then 5 and 8. Vector 8 is to monitor and the subsequent action is conditional upon the observation. If 1 is observed the optimal action is to stop if 2 or 3 are observed it is to *continue* but using different vectors. A policy graph clarifies the decision sequence (Figure 4). If vector 1 is optimal then it is optimal to *continue* for three years using vectors 1, 3 and 5, after three years it is then optimal to *monitor* using vector 8, *stop* if U4 is observed, return to 1 if H12 is observed and if H18 is observed *continue* for 4 years and then *stop*. The last of these alternatives implies that, in the short-run, it is worthwhile benefiting from the non-market value, but as the probability of U4 increases then it is optimal to stop. Some actions, for instance vector 18, indicate repeated monitoring when H18 is observed. Monitoring, by modifying the belief state, is reducing the 'entropy' associated with the vegetation state until a belief state is reached where either *continue* or *stop* are optimal actions (Kaelbling *et al,* 1998).

3.4 Sensitivity Analysis

Here we analyse sensitivity of the optimal solution to selected parameters, of particular interest is the monitoring decision. As the accuracy of monitoring increases as measured by *r*, the frequency of monitoring increases and the range of belief states over which monitoring is an optimal initial action increases. In the extreme case where *r=1 (perfect monitoring)* monitoring occurs every three years after an initial decision to *continue* and it is the action with the widest probability range. This result is represented in Figure 5 and Table 6, note that the range of probability values over which vectors that involve monitoring (1, 2, 3) is expanded compared to Figure 3 Conversly , if r is reduced to 0.6 then monitoring is never optimal and the agri-environmental scheme stops after 8 years whatever the belief state.

Table 6

The frequency of monitoring increases as the discount rate falls, this indicates that monitoring is of more value at low discount rates. Monitoring is a form of investment in that it increases the efficiency of future actions thus as the discount rate increases the value of monitoring falls.

Figure 6

4. Conclusion

Monitoring is an essential component of schemes designed to conserve or enhance biodiversity. Such schemes are now a central part of European agricultural policy, but are also found in forestry management, wildlife management and marine conservation. A characteristic of these policies is that they offer relatively long term contracts to farmers to conserve or reinstate a particular ecosystem. To date the literature has had relatively few contributions on the subject of optimal dynamic monitoring. The original Arrow Fisher Henry model is concerned with information gathering, but as a passive function of time passing and is limited to two-periods. The POMDP framework presents a flexible approach to determine actions where the stochastic process is represented by a Markov chain. It allows a link to ecology where Markov chains are applied to model transitions in vegetation states. It also allows a classification of vegetation transitions along a continuum from reversible to irreversible.

Evidence from UK agri-environmental schemes suggests that monitoring is necessary and costly (Falconner and Whitby, 1999). In this paper we propose a POMDP as a framework for analysing monitoring patterns. The results show the complexity of the monitoring decision which depends upon the prior probability of a vegetation state. A general result is that monitoring tends to be optimal where there is more uncertainty about the target state and repeated monitoring might be used before making an irreversible decision.

References

Albers, H. J., Fisher, A. C., Hanemann, W.M., 1996. Valuation and management of tropical forests implications of uncertainty and irreversibility. Environmental and Resource Economics 8, 39-61.

Arrow, K.J. and A. C. Fisher, 1974. Environmental preservation, uncertainty, and irreversibility, Quarterly Journal of Economics, 88 312-319.

Ball, D. F., J. Dale, J. Sheail and O. W. Heal, 1982. Vegetation change in upland landscapes. NERC/ITE, London

Balzer, H. 2000. Markov Chain Models for Vegetation Dynamics, Ecological Modelling, 126, 139-154.

Cassandra, A. R. 1995. Optimal Policies for Partially Observable Markov Decision Processes. Department of Computer Science Brown University.

EC. 1992. Council Regulation EEC No 2078 on agricultural production methods compatible with the requirements of the protection of the environment and the maintenance of the countryside 'The Agri-Environment Regulation', Official Journal of the European Communities L215/85.

EC. 1996. Council Regulation EEC 746/96 laying down the detailed rules for the application of 2078/92, Official Journal of the European Communities L102/96

Falconer, K. and Whitby, M.C. 1999. Transactions and Administrative Costs in Countryside Stewardship Policies: An Investigation for Eight European Member States. Newcastle University Centre for Rural Economy Research Report.

Fisher, A. C. 2000. Investment under uncertianty and option value in environmental economics. Resource and Energy Economics, 22, 197-204.

Garrod G. D., Willis K. G., 1994. Valuing biodiversity and nature conservation at a local-level. Biodiversity and Conservation 3, 555-565.

Henry, C., 1974. Investment decisions under uncertainty, American Economic Review 64,1006-1012.

Hooper, 1992. Measurement and perception of change: filed monitoring of environmental change in the Environmentally Sensitive Areas, in: M. C. Whitby ed Land Use Change and it Consequences. HMSO, London

Horn, H. S. 1976. Succession. In May R. M. Ed. Theoretical Ecology. Principles and Applications . Saunders : Philadelphia, pp 187-204.

Monahan, G. E. 1982. A Survey of Partially Observable Markov Decision Processes: Theory Models and Algorithms, Management Science, 28, 1-16.

National Audit Office 1997. Protecting Environmentally Sensitive Areas. London:HMSO.

Laffont, J-J. and Tirole, J. 1993. A Theory of Incentives in Procurement and Regulation. Cambridge, Massachusetts: MIT Press.

Rodwell J S. 1991. *British Plant Communities: Volume 2 Mires and Heaths.* Cambridge: Cambridge University Press. 628 pp.

Rodwell J S**.** 1992. *British Plant Communities: Volume 3 Grassland and Montane Communities*. Cambridge: Cambridge University Press. 540 pp.

Ross, S. M., 1996. Stochastic Processes 2nd Edition. Wiley, New York

Rushton, S.P and R. A. Sanderson, J Wildig and J. P. Byrne, 1996. The effects of grazing management on moorland vegetation: a comparison of farm unit, grazing paddock and plot experiments using a community modelling approach, Aspects of Applied Biology 44, 1-8.

Smallwood, R. D. and Sondik, E.J. 1973. The Optimal Control of Partially Observable Markov Processes Over a Finite Horizon, *Operations Research*, 21, 1071-1088.

Table 1 Immediate Rewards and Optimal Solution per ha

Current State: Action:	Agricultural Grassland (AG)	Hay meadow (HM)
Immediate rewards ($w^{a}_{ij\theta}$):		
Continue	200	1500
Stop	700	0
Monitor	0	1300
Optimal Policy Vectors (α^{k})		
n=1		
Continue	200	1500
Stop	700	0
n=2		
Continue	390	2678
Stop	1365	665
Monitor	693.75	2487.75

Table 2 Results for Example 1 For Specific Points in Belief Space

a. $\pi = 1, r_{11} = r_{22} = 1, \mathrm{c_m} = 0$	£ Net Social Benefit per ha
a1. MDP (perfect information) continue	£2773
a2. MDP (myopic) continue	£2678
a3 MDP stop	£665
b $\pi = 1, r_{11} = r_{22} = 1, \mathrm{c_m} = 200$	
b1 POMDP continue	£2678
b2 POMDP monitor	£2573
b3 POMDP stop	£665.0
c $\pi = 0.5, r_{11} = r_{22} = 1, \mathrm{c_m} = 200$	
c1 POMDP continue	1534
c2 POMDP monitor	1619
c3 POMDP stop	1015
d $\pi = 0.5, r_{11} = r_{22} = 0.85, \mathrm{c_m} = 200$	
d1 POMDP monitor	1490.75

Table 3 Markov Transition Matrix

	Markov Matrix P^a			Long-run Markov Matrix			Observation Matrix (R^a)		
Action	U4	H18	H12	U4	H18	H12	U4	H18	H12
Continue (P^1)									
U4	0.98	0.02	0	0.87	0.07	0.06	1/I	1/I	1/I
H18	0.17	0.66	0.17	0.87	0.07	0.06	1/I	1/I	1/I
H12	0.1	0.1	0.8	0.87	0.07	0.06	1/I	1/I	1/I
Monitor (P^2)									
U4	0.94	0.03	0.03	0	0	1	0.8	0.1	0.1
H18	0	0.66	0.34	0	0	1	0.1	0.8	0.1
H12	0	0	1	0	0	1	0.1	0.1	0.8
Stop (P^3)									
U4	1	0	0	1	0	0	1/I	1/I	1/I
H18	1	0	0	1	0	0	1/I	1/I	1/I
H12	1	0	0	1	0	0	1/I	1/I	1/I

Table 4 Payoff Matrix for Example 2 £ per ha

States: Actions:	*U4*	*H18*	*H12*
1. Continue	49.17	99.17	149.17
2. Monitor	-10.83	39.17	89.17
3. Stop	89.4	89.4	89.4

Parameters: Gross margin from sheep £44.7 per ewe; stocking rate 1.25 ewes pe ha for ESA 2 ewes per ha profit maximising. Non-market values U4 £0, H18 £50, H12 £100 per ha. Monitoring cost £50 per ha. The shadow price of public funds is £0.2.

Table 5 Results for Cambrian ESA Case Study

		Reward Vector (α)			Vector n-1		
Vector No	Action	1 (U4)	2 (H18)	3 (H12)	1	2	3
1	1	1566.35	1808.84	1959.28	3	3	3
2	1	1574.85	1812.36	1957.37	4	4	4
3	1	1597.03	1811.66	1955.64	5	5	5
4	1	1605.98	1816.05	1951.46	6	6	6
5	1	1629.33	1810.07	1947.01	8	8	8
6	1	1638.75	1816.57	1939.52	23	23	23
7	2	1662.52	1799.54	1933.05	26	5	1
8	2	1663.32	1802.56	1932.34	26	6	1
9	2	1665.41	1793.81	1931.72	26	7	1
10	2	1665.48	1795.30	1931.65	26	8	1
11	2	1665.66	1790.89	1931.60	26	9	1
12	2	1665.66	1791.64	1931.59	26	10	1
13	2	1665.68	1789.42	1931.58	26	11	1
14	2	1665.68	1789.80	1931.58	26	12	1
15	2	1665.68	1788.69	1931.58	26	13	1
16	2	1665.68	1788.87	1931.58	26	14	1
17	2	1665.68	1788.32	1931.58	26	15	1
18	2	1665.68	1788.41	1931.58	26	16	1
19	2	1665.68	1787.95	1931.58	26	19	1
20	2	1665.69	1791.06	1931.57	26	21	1
21	2	1665.73	1794.17	1931.48	26	22	1
22	2	1666.29	1800.42	1930.47	26	23	1
23	1	1673.24	1813.57	1919.88	24	24	24
24	1	1709.55	1807.05	1890.30	25	25	25
25	1	1747.77	1797.77	1847.77	26	26	26
26	3	1788.00	1788.00	1788.00	26	26	26

Action 1 *continue*, 2 *monitor*, 3 *stop*

Table 6 Results for Cambrian ESA Case Study – Perfect Monitoring

		Reward Vector (α)			Vector n-1		
Vector No	Action	1 (U4)	2 (H18)	3 (H12)	1	2	3
1	1	1630.26	1855.02	2005.05	2	2	2
2	1	1664.30	1856.07	2001.90	3	3	3
3	2	1700.14	1850.55	1993.97	4	2	1
4	3	1788.00	1788.00	1788.00	4	4	4

Action 1 *continue*,2 *monitor*, 3 *stop*

Figure 1 Optimal Action by Belief State

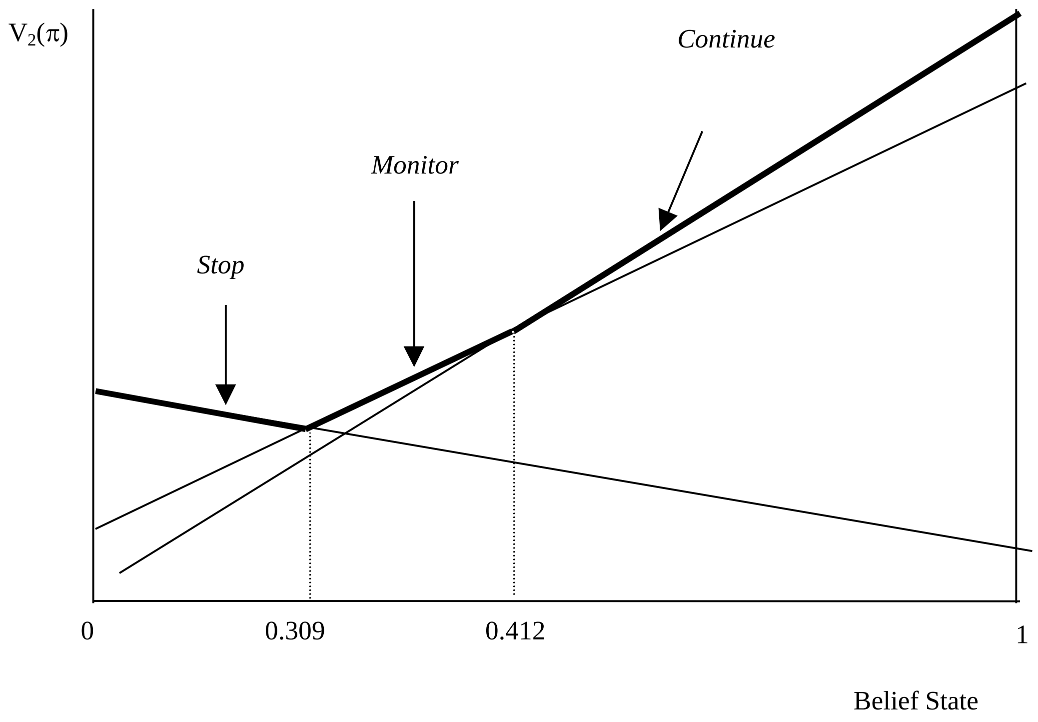

Figure 2 Policy Graph

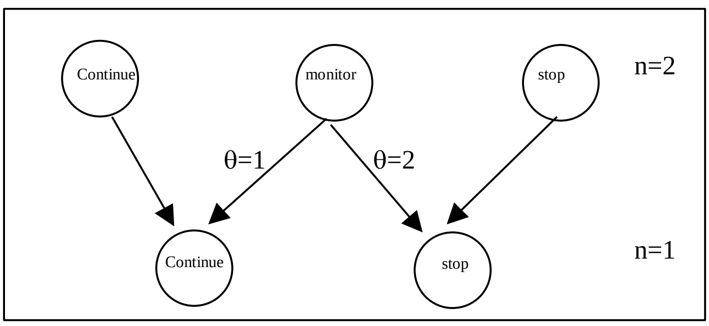

Figure 3 Example 2 Optimal Action for Different Belief States

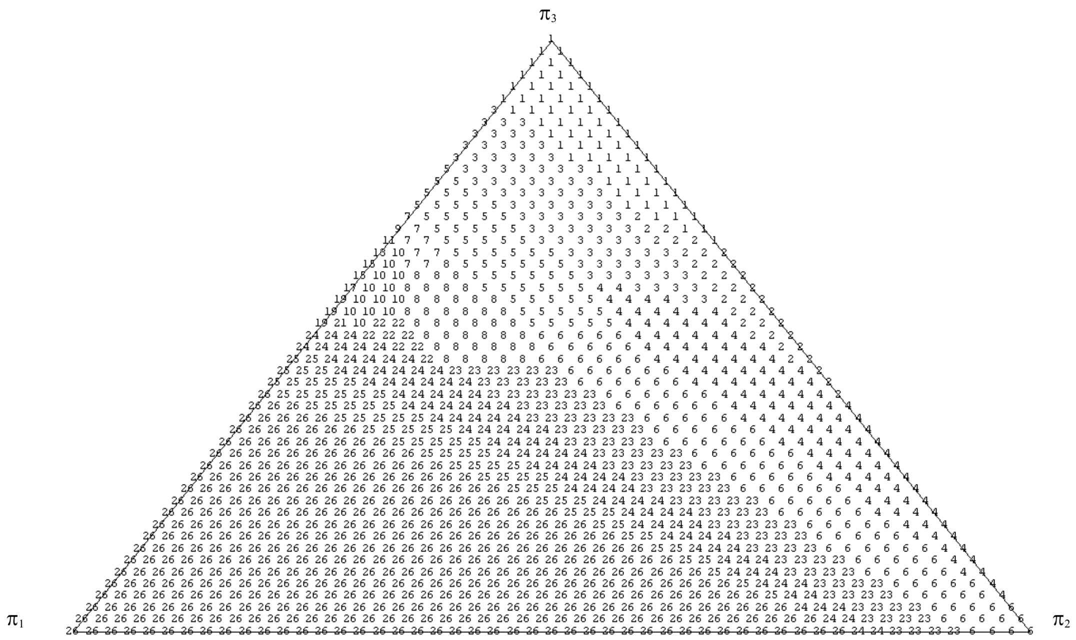

Figure 4 Example 2 Policy Graph Vector 1

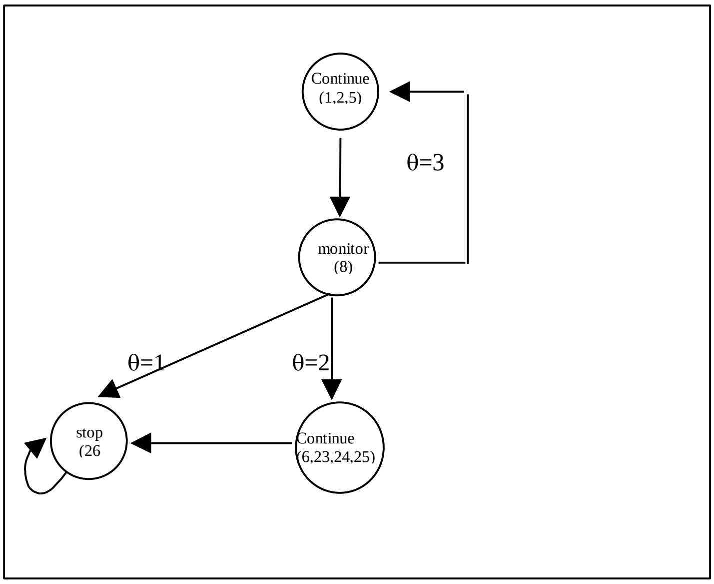

Figure 5 Optimal Initial Policy with Perfect Monitoring