Kristensen, A. R. 1993. *Markov Decision Programming Techniques Applied to the Animal Replacement Problem*. Dissertation for the dr. agro. degree. Dina KVL, Department of Animal Science and Animal Health, The Royal Veterinary and Agricultural University, Copenhagen. ISBN 87 7026 332 9.

Chapter VII. Bayesian updating in hierarchic Markov processes applied to the animal replacement problem (p 85-104)

The chapter has also been published as a journal paper, so please cite it as:

Kristensen, A. R. 1993. Bayesian updating in hierarchic Markov processes applied to the animal replacement problem. *European Review of Agricultural Economics 20*, 223-239. DOI: 10.1093/erae/20.2.223

VII

Bayesian updating in hierarchic Markov processes applied to the animal replacement problem*

ANDERS RINGGAARD KRISTENSEN
Royal Veterinary and Agricultural University, Copenhagen

(Received January 1992, final version received June 1992)

Summary

The observed level of milk yield of a dairy cow or the litter size of a sow is only partially the result of a permanent characteristic of the animal; temporary effects are also involved. Thus, we face a problem concerning the proper definition and measurement of the traits in order to give the best possible prediction of the future revenues from an animal considered for replacement. A trait model describing the underlying effects is built into a model combining a Bayesian approach with a hierarchic Markov process in order to be able to calculate optimal replacement policies under various conditions.

Keywords: replacement, animal, Bayesian updating, Markov decision programming.

1. Introduction

In any production based on the operation of an asset of significant value, the determination of an optimal lifetime of the asset is important in order to maximise the profit from the production, which in this paper is assumed to be the overall objective of the manager. The considerations are relevant no matter whether the asset is a dairy cow, a farm building or some kind of industrial equipment, but the way of solving the problem may vary considerably, depending on the individual situation. Most often the asset will be replaced by a new asset of the same or at least a similar kind. In that case the present asset is only a link in a chain of assets. Then we have

* This research was carried out as part of Dina, Danish Informatics Network in the Agricultural Sciences.

European Review of Agricultural Economics
20 (1993), 223–239

0165–1587/93/0020–0223 $2.00

to consider what kind of objective function to use in order to meet the overall objective of profit maximisation.

In some cases it is relevant to maximise the total profit of the asset during the entire lifetime. That applies, for instance, when the asset itself is a scarce resource. A case in point is the determination of the optimal age at slaughter of fattening bulls in a dairy herd, where the main activity is milk production. If the dairy farmer does not buy bull calves at the market (for instance because of infection risk) he will only have the bull calves provided by the dairy cows of the herd. If the housing capacity is sufficient, the bull calves are a scarce resource, and the total profit is maximised if the net returns per animal during its whole lifetime are maximised. We shall refer to this situation as the *single asset* situation.

Another situation is when new assets are permanently available at the market. In that case the total profit is *not* maximised by maximising total net returns per asset. A more relevant criterion is here the maximisation of either average net returns over time or the total discounted net returns (i.e. the *present value*) of the entire chain of assets. In both cases the time horizon may be finite or infinite, whichever is relevant. An infinite horizon is just an abstraction indicating that the time of termination (the last link of the chain) is unknown, but at least 'far' ahead. We shall refer to this situation as the *asset chain* situation.

Finally, we shall consider a situation where a restriction is imposed on the production. It may be in the form of either a production quota or a limited supply of an input factor. In that case the total profit is maximised by the maximisation of average net returns per product or factor unit. We shall refer to this situation as the *quota* situation.

A more formal discussion of possible objective functions is given by Kristensen (1992a). The choice of objective function depends only on the conditions of production. It does not matter what kind of asset we are dealing with. If we turn to the *method* used in the maximisation of the objective function, it will depend very much on the nature of the asset. The classical replacement theory developed by pioneers like Preinrich (1940) and Terborgh (1949) typically assumes that all functions and parameters describing the problem are completely known in advance, and that no random variation is involved. The implicit items considered comprise machinery and other kinds of industrial equipment. The functions and parameters representing the problem are used for deduction of *general* replacement rules based on variants of the marginal net revenue approach.

As opposed to machinery or industrial equipment, the present study is part of a larger research project dealing with the *animal* replacement problem in agricultural production. A relevant question to ask is, therefore, in what way the animal problem differs from the general set-up. Based on a study by Ben-Ari et al. (1983) the main difficulties of the animal replacement problem may be summarised as:

(1) *Uniformity.* The traits of an animal are difficult to define and measure.
(2) *Variability.* The random variation of each trait is relatively large.
(3) *Reproductive cycle.* The production of for instance cows and sows is cyclic. It has to be decided *in which* cycle to replace as well as *when* to replace inside a cycle.
(4) *Herd restraints.* Animal production is performed in herds. There may be restraints that apply to the herd as a whole and not to the individual animal. Examples are a limited supply of heifers or gilts, limited housing capacity or a milk quota.

Because of the variability, Markov decision programming had already been applied to the dairy cow replacement problem by Giaever (1966). In an evaluation of techniques van Arendonk (1984) concluded that in dairy cow replacement this method should be used in preference to the marginal net revenue approach. Also, the method directly solves the problems caused by the reproductive cycle as shown by Kristensen and Østergaard (1982) as well as van Arendonk (1985b). The only problem concerning variability and cyclic production is that in order to cover the variability in traits, the state variables (traits) have to be represented by many levels, and to deal with the cyclic production, a state variable representing the stage of the cycle has to be included. Both aspects contribute significantly to an explosive growth of the state space. Therefore, we face a *dimensionality* problem. Though all necessary conditions of a Markov decision process are met, the solution is prohibitive in practice even on modern computers.

The problem of herd restraints is important. In dairy cattle two particular restraints should be considered. One is a limited supply of heifers when the dairy farmer only uses home-grown heifers as replacements. In that case a simple comparison of the animal in production with a replacement is no longer valid. Instead we face a much more difficult problem of choosing the optimal composition of animals from the available population of cows in production and heifers for replacement. The other restraint is the milk quota which is imposed on all dairy herds of the EC.

The overall objective of the animal replacement research project is to adapt the Markov decision programming techniques in order to be able to cope with the problem in a satisfactory way. The problems to be solved (totally or partially) have been identified as the dimensionality problem, herd restraints and uniformity. In order to circumvent the dimensionality problem, a new notion of a hierarchic Markov process was introduced by Kristensen (1988) and applied to the dairy cow replacement problem by Kristensen (1987, 1989). The technique may be applied in the single asset situation, the asset chain situation and the quota situation.

Both herd restraints mentioned above have been studied in the project. The milk quota restraint was discussed by Kristensen (1989) based on a technique described in Kristensen (1991) of maximising average net returns per kg milk produced. A limited supply of replacement heifers was discussed

by Ben-Ari and Gal (1986), who developed a technique called parameter iteration. The method was modified and further developed by Kristensen (1992b). In both studies on the latter restraint, a herd model was based on an underlying single animal model, which in the study by Ben-Ari and Gal (1986) was an ordinary Markov decision process and in the study by Kristensen (1992b) was a hierarchic Markov process.

Only a solution of the uniformity problem concerning the definition and measurement of traits remains to complete the project. The solution of that problem is the objective of the present study. In any replacement problem a good prediction of the future net revenues (or *rewards* as they are called in Markov decision programming) is essential. This prediction is based on the observed traits of the animal, but since the traits are subject to random variation, we do not know to what extent the observed value represents a permanent characteristic of the animal or just a temporary fluctuation. In order to be able to give the best possible prediction of future rewards, a method has been designed which combines the ability of knowledge updating known from causal probabilistic nets, as described for example by Pearl (1988), and hierarchic Markov processes. In order to keep the presentation simple we shall only consider examples where the animal in production is compared to a standard replacement. However, the technique may just as well be applied in a herd model under some restraint as discussed above.

For an empirical application of the updating technique in a sow replacement study, reference is made to Jørgensen (1992).

2. A model describing a trait of an animal

In this section we shall describe a general model of an animal trait to be used in replacement studies. The model will form the basis of the further considerations of the paper.

Assume that the *state* of an animal is observed at regular intervals called *stages*. The state is defined by the values of a number of *state variables* each representing a trait of the animal. We assume that one of the traits (Y_n) is described by the following relation at stage n:

$$Y_n = m(\cdot) + X + e_n, \quad n = 0, \ldots, N, \tag{1}$$

where m is a known function expressing the expected value of Y_n under the circumstances in question. The circumstances are represented by the arguments of m which may be the value of other state variables (e.g. the age of the animal, the season, etc.) and/or the average herd level concerning the trait. The trait itself may, for instance, be the milk yield of a dairy cow or the litter size of a sow. The symbol X is the combined effect of genetic level and permanent environment. We assume X to be normally distributed with the expected value zero and a certain variance σ_x^2 representing the variation

across the population. The symbol e_n represents the random variation caused by temporary environmental effects. We assume that e_n is normally distributed with the expected value zero and a certain variance σ_e^2. The random variables X and e_n are assumed to be independent. Furthermore, the variables $e_1, \ldots, e_N$ of a particular animal are assumed to be independent of the corresponding variables of the other animals of the herd. Thus all systematic effects influencing all animals of the herd (e.g. seasonal effects) are assumed to be included in the function $m(\cdot)$. The sum $I_n = X + e_n$ forms the value of a state variable.

The relation over time of $e_1, \ldots, e_N$ is described by a first order autoregressive process, i.e.

$$e_n = ae_{n-1} + \varepsilon_n, \quad n = 1, \ldots, N, \tag{2}$$

where $0 < a < 1$ and $\varepsilon_1, \ldots, \varepsilon_N$ are independent and normally distributed with the expected value zero and the variance $(1 - a^2)\sigma_e^2$. Furthermore, ε_n is assumed to be independent of e_{n-1} and X for $n = 1, \ldots, N$.

As appears from equations (1) and (2), the permanent effect X varies only between animals, whereas the temporary random effect e_n varies over time for the same animal. It is obvious that the value of the permanent effect is very important in the decision of which animals to keep in the *long run* (e.g. whether a cow should be kept for an additional lactation) and that the current value of the random effect e_n is important in the decision concerning the optimal replacement time in the *short run* (e.g. when to replace a cow inside a lactation). Thus, if the value of X is high, we would probably ignore a low current value of e_n which just represents a temporary crisis. On the other hand, a sufficiently high current value of e_n might lead to postponed replacement of an animal of low permanent value X.

These fundamental observations illustrate that the differentiation of variation between animals and over time for the same animal is important for the replacement decision, because it directly influences our expectations concerning the future net revenues from the animal. The only problem is that neither the permanent effect X nor the current random effect e_n are directly observable. What we observe are the resulting numerical values $Y_1, \ldots, Y_N$ of the trait in question, but since the systematic effect m is assumed to be known, this is equivalent to the sums $I_1, \ldots, I_N$, where $I_n = X + e_n$. On the other hand we may have a prior belief in X based on an estimated variance among animals and possible observations concerning the animal of characteristics correlated with X. Further, as observations of the sums I_n are gathered they will increase our knowledge of X. If, for instance, all sums are relatively large for an animal, it implicitly indicates a high value of X and vice versa.

The trait model (1) and (2) may easily be extended to cover several traits, each being influenced by several unobservable effects. Assume for instance that Y_{1n} and Y_{2n} are the milk yield and weight of a dairy cow at stage n and

that each trait is influenced by two unobservable effects. In matrix notation we may express the relations as follows:

$$Y_n = \begin{pmatrix} m_1(\cdot) \\ m_2(\cdot) \end{pmatrix} + \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} \begin{pmatrix} 1 \\ c_n \end{pmatrix} + \begin{pmatrix} e_{1n} \\ e_{2n} \end{pmatrix}, \tag{3}$$

where $Y_n = (Y_{1n}, Y_{2n})'$, and

$$\begin{pmatrix} e_{1n} \\ e_{2n} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} e_{1,n-1} \\ e_{2,n-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1n} \\ \varepsilon_{2n} \end{pmatrix}. \tag{4}$$

If, at a previous stage, the cow has been suffering from a specific disease (e.g. mastitis) that permanently influences the milk yield and weight, the value of c_n is -1. Otherwise it is zero. Equation (3) expresses that the observed milk yield and the weight of the cow are determined partly by permanent animal-specific effects (X_{11} and X_{21}) as in the single trait model (1), partly by possible permanent negative effects of a previous disease (X_{12} and X_{22}) and finally by temporary random effects (e_{1n} and e_{2n}). The variables X_{11}, X_{12}, X_{21} and X_{22} may be mutually correlated, and e_{1n}, e_{2n} may be correlated, whereas $(X_{11}, X_{12}, X_{21}, X_{22})'$ and $(e_{1n}, e_{2n})'$ are assumed to be independent.

In the following, only the single-trait model of equations (1) and (2) will be discussed, but all results may be directly extended to cover the multi-trait model of equations (3) and (4). For the derivation of the extended results, a Kalman filter approach as described by, for example, Harrison and Stevens (1976) is a relevant tool.

We refer to X in equation (1) as the *basic state* of the animal and to the sum I_n as the *current state*. Thus the current state is directly observable and therefore known at any stage, whereas the basic state is unknown. In accordance with common practice in dynamic programming we shall consider both kinds of states to be discrete, i.e. only a finite number of levels are considered for each kind. Both X and I_n are random variables, which will be referred to in upper-case letters. Transformed realisations of X and I_n, on the other hand, will be denoted by lower-case letters $x \in \Omega_X$ and $i \in \Omega_I$, respectively, where Ω_X and Ω_I are finite sets. In other words, if the basic state is x, it means that X is in the interval $]x^-; x^+]$ defined as $\{y \mid x^- < y \leq x^+\}$, where x^- and x^+ are the lower and higher limit, respectively, of the x'th level of X. In addition to the states defined by levels of I_n, the state space Ω_I includes a *replacement state* representing a situation where the animal has been culled.

From the assumptions made concerning the normal distribution of X, we are able to calculate the prior probability $p_x(0)$ of any basic state x. At any stage n we may select an *action* $d \in \{1, 2\}$ that influences the system. We shall interpret $d = 1$ as 'keep' and $d = 2$ as 'replace'. For given basic state x and current state i at stage n, we know the conditional probability $p^d_{xij}(n)$ of the

current state to be j at stage $n+1$ if action d is taken. For $d=1$ we have approximately:

$$\begin{aligned} p^1_{xij}(n) &= P(j \in]j^-; j^+] \mid x \in]x^-, x^+] \wedge i \in]i^-; i^+]) \\ &\approx P(x^m + e_{n+1} \in]j^-; j^+] \mid X = x^m \wedge I = i^m) \\ &= P(e_{n+1} \in]j^- - x^m; j^+ - x^m] \mid X = x^m \wedge e_n = i^m - x^m) \\ &= \Phi((j^+ - x^m - a(i^m - x^m))/(1-a^2)^{1/2}\sigma_e) \\ &\quad - \Phi((j^- - x^m - a(i^m - x^m))/(1-a^2)^{1/2}\sigma_e), \end{aligned} \tag{5}$$

where Φ is the distribution function of the standard normal distribution. The symbol x^m denotes the conditional expectation $\mathrm{E}(X \mid X \in]x^-; x^+])$, and analogously for i^m. If $d=2$, the process transfers to an absorbing replacement state with probability 1.

At stage 0, the marginal probability of a transition from current state i to j at stage 1 under the action d is calculated as

$$p^d_{ij}(0) = \sum_{x \in \Omega_x} p^d_{xij}(0) p_x(0). \tag{6}$$

At each stage the current state i is observed, each time increasing our knowledge of the basic state x. Our current belief at stage n concerning the basic state is represented by the probability distribution given by $p_x(n)$. If the current state is i and at stage $n+1$ we have observed a transition from state i to state j following the action d, we may use Bayes' theorem to update our belief concerning the basic state. The new probability distribution at stage $n+1$ is calculated as

$$p_x(n+1) = p_x(n) p^d_{xij}(n) / p^d_{ij}(n), \quad x \in \Omega_X,\ n = 0, \ldots, N-1. \tag{7}$$

If the current state i is observed at stage n, and the action d is taken, a *reward* depending on the basic state x is gained. This reward is denoted as $r^d_{xi}(n)$. We also assume that some kind of physical output $m^d_{xi}(n)$ is produced during the stage. In a replacement model, the reward is usually defined as the net revenue, and the physical output may be defined as the amount of milk produced by a cow, the litter size of a sow, etc.

3. Causal probabilistic nets

A trait described as in section 2 may be modelled by a *causal probabilistic net* (sometimes also referred to as an *inference diagram* or a *Bayes belief net*) as shown in Figure 1. Using the terminology of Tatman and Shachter (1990), the elements of the net are *decision nodes* representing variables under the control of the decision maker, *chance nodes* corresponding to random variables or random events, *value nodes* together representing the arguments of

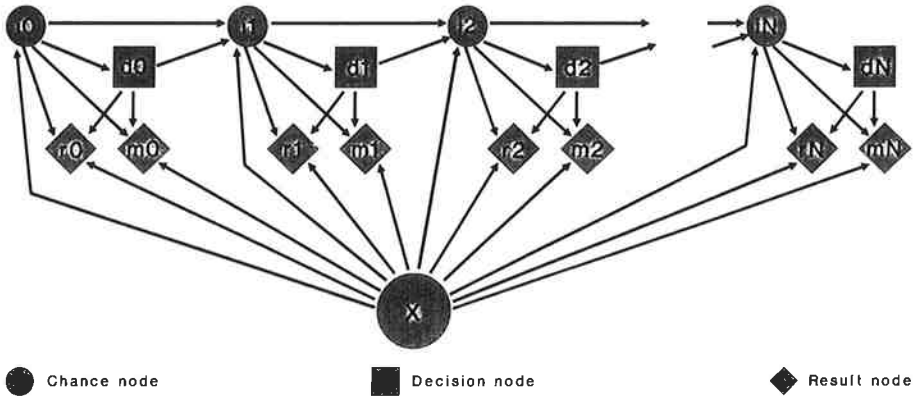

Figure 1. *The animal model described as a causal probabilistic net*

X = basic state
in = observable state at stage n
dn = action at stage n
rn = reward at stage n
mn = physical output at stage n

the objective function of the model and, finally, *directed arcs* representing the causal relationships among the nodes. Arcs into a decision node indicate the information which will be known to the decision maker at the time of decision. Arcs into a chance node indicate which variables condition the probability distribution of the associated random variable. Arcs into a value node indicate which variables condition the associated expected value (arcs from a value node to another node are not allowed).

An advantage of causal probabilistic nets is that they provide a graphical modelling language very close to ordinary human reasoning, and at the same time they are mathematically well defined and, therefore, suitable for analyses derived from traditional probabilistic theory. The main idea is that the knowledge of unobservable nodes (state variables) is updated each time the value of any other node is observed. Thus, in the example, the model is learning by successive observations, and step by step knowledge concerning the value of X is increased. By assuming a decision policy describing which decision to make for given values of the chance nodes it is possible to calculate the expected value of the objective function under the policy. This value may be compared to the expected values under alternative policies, making possible the identification of an optimal (or at least a satisfactory) policy. During the last few years there has been extensive research into causal probabilistic nets, and many results have been obtained on how to collect and distribute evidence over the net (e.g. Pearl, 1988; Jensen et al., 1990).

Possible objective functions are the expected sum of all rewards under a policy (ignoring the physical output) or the expected sum of all discounted rewards under a policy (if the physical output is defined to be the stage length). Thus, the causal probabilistic net of Figure 1 may be used directly

for solving the problem described above, as long as the time horizon is restricted to the N stages. In other words, the causal probabilistic net directly solves the problem of determining the optimal lifetime of an animal in the single asset situation as defined in section 1. If, on the other hand, the system at the end of the N stages is replaced by a new system described in the same way (and a third system will ultimately replace the second one and so on) the causal probabilistic net approach will run into trouble, because all time steps explicitly have to be in the model. Thus, the method does not cover the asset chain and quota situations mentioned in section 1.

The infinite stage problem is dealt with appropriately by a hierarchic Markov process, but the problem of that method in relation to the current problem is that all states must be observable and all parameters must be known. The possibility of learning from the successive observations is not directly present. Therefore, a hybridisation of a hierarchic Markov process and a causal probabilistic net is desirable in order to cover the asset chain and quota situations as described in section 1.

4. Hierarchic Markov processes

If we want to describe the system by a Markov decision process instead of a causal probabilistic net, we may take at least two different views. One possibility is to define it as a Markov decision process with unobservable states. Such processes are called *partially observable Markov decision processes*, and they have been discussed by Monahan (1982). Another view to take is to define the process in such a way that the state space is directly observable, but with unknown parameters. This kind of process is called an *adaptive Markov decision process*. It has been discussed in detail by Wessels (1968) and later by van Hee (1978).

However, in this paper we shall consider the system in the context of a hierarchic Markov process. A hierarchic Markov process is a series of finite stage Markov decision processes called *subprocesses* built together in one Markov decision process called the *main process*. The basic formulation, including an optimisation cycle for the discounting (present value) criterion, is given by Kristensen (1988). This criterion covers the asset chain situation. A similar cycle intended for the quota situation, using a criterion maximising the average rewards/output ratio, is described by Kristensen (1991). The hierarchic technique has been developed as a way of circumventing the dimensionality problem of Markov decision programming, i.e. that practical problems have a tendency to become very large and, therefore, difficult to analyse by the usual techniques as mentioned in section 1. In the description of hierarchic Markov processes it has been assumed that all parameters of the model are known and that all states in the subprocesses as well as in the main process are directly observable.

If, in the specification of traits in section 2, the basic state X had been directly observable, the formulation as a hierarchic Markov process would have been straightforward. In that case the state space of the main process would have been the finite set Ω_X, and the state space of the subprocesses would be Ω_l. The parameters of the x'th subprocess would be $p^d_{xij}(n)$, $r^d_{xi}(n)$ and $m^d_{xi}(n)$. Finally, the x'th element of the y'th row of the main process transition matrix would be $p_x(0) = \Phi(x^+/\sigma_x) - \Phi(x^-/\sigma_x)$. Under these assumptions we would be able to determine an optimal solution for the asset chain situation as well as for the quota situation.

Since, however, the basic state x is unobservable, we may conclude that we have a hierarchic Markov process with *unobservable main state*. But at each stage we observe the state transition in the subprocess and use the observation for updating our knowledge of the state of the main process. At the end of the subprocess, however, the learning stops, because the knowledge of the old system cannot be used on the new one. In other words, the specific traits of the present animal will not improve our ability to predict the future revenues from the replacement.

In the following we shall describe how the updating of knowledge may be incorporated into the hierarchic process. First, we should notice that, for a given transition from state i to j in a subprocess, the new probability distribution of the main stage x is uniquely defined according to equation (7). Therefore, the imperfect knowledge of the main state does not add any further random elements to the transitions of the subprocesses. Next, we should consider whether we know that the distribution of x always belongs to a certain class of distributions so that it may be sufficiently described by one or a few parameters. In that case we may replace $p_1(n), \ldots, p_v(n)$, where v is the number of elements in Ω_X, by these few parameters without losing any information. The prior distribution of X is normal with known mean and variance. In the following we shall investigate the posterior distribution after observations of current states in the subprocesses.

Having observed state i through the sum $I_n = X + e_n$ at stage n, we are ready to observe the state j defined by the sum $I_{n+1} = X + e_{n+1}$ at the next stage. Recalling that X has a *fixed* (but unknown) value, we find that the distribution of I_{n+1} is normal, having the expected value

$$\mathrm{E}(I_{n+1} \mid X, I_n) = \mathrm{E}(X + a(I_n - X) + \varepsilon_{n+1} \mid X, I_n) = X + a(I_n - X) \tag{8}$$

and variance

$$\mathrm{V}(I_{n+1} \mid X, I_n) = \mathrm{V}(X + a(I_n - X) + \varepsilon_{n+1} \mid X, I_n) = (1 - a^2)\sigma_e^2. \tag{9}$$

From equations (8) and (9) we observe that I_{n+1} has unknown mean but known variance. Our prior knowledge of the mean is that it is normally distributed with the expected value

$$\mu_{n+1} = \mathrm{E}(X + a(I_n - X)) = (1 - a)\mathrm{E}_n(X) + aI_n \tag{10}$$

and the variance

$$\sigma_{n+1}^2 = \mathrm{V}(X + a(I_n - X)) = (1-a)^2 \mathrm{V}_n(X), \tag{11}$$

where the index n on the expectation and variance of X refer to the distribution of X at stage n.

Having taken the observation of I_{n+1} we may update the distribution of the mean according to the following equations taken from DeGroot (1970):

$$\mu'_{n+1} = (\mu_{n+1}\sigma_{n+1}^{-2} + I_{n+1}(1-a^2)^{-1}\sigma_e^{-2})/(\sigma_{n+1}^{-2} + (1-a^2)^{-1}\sigma_e^{-2}) \tag{12}$$

and

$$\sigma'^2_{n+1} = \sigma_{n+1}^2(1-a^2)\sigma_e^2/(\sigma_{n+1}^2 + (1-a^2)\sigma_e^2). \tag{13}$$

Furthermore, the posterior distribution of I_{n+1} is still normal according to DeGroot (1970). By combining equations (10) and (11) with (12) and (13) we are able to calculate the parameters of the new distribution of X at stage $n+1$:

$$\begin{aligned}\mathrm{E}_{n+1}(X) &= (\mu'_{n+1} - aI_n)/(1-a) = (1-a^2)\sigma_e^2 \mathrm{E}_n(x) \\ &\quad + (I_{n+1} - aI_n)(1-a)\mathrm{V}_n(x))/((1-a^2)\sigma_e^2 + (1-a)^2 \mathrm{V}_n(x))\end{aligned} \tag{14}$$

and

$$\begin{aligned}\mathrm{V}_{n+1}(x) &= \sigma'^2_{n+1}/(1-a)^2 \\ &= (1-a^2)\sigma_e^2 \mathrm{V}_n(x)/((1-a)^2 \mathrm{V}_n(x) + (1-a^2)\sigma_e^2).\end{aligned} \tag{15}$$

Under the assumptions made we find that if the prior distribution of X is normal, it will remain normal at all stages. Only the expectation and variance change over stages, and furthermore the change in variance does not depend on the observed value of I_{n+1}! For given prior variance $\mathrm{V}_0(x) = \sigma_x^2$ we are able to calculate the variance at all future stages in advance according to the recurrent equation (15). If instead of the variance we consider the reciprocal value (sometimes referred to as the *precision*), we easily have

$$1/\mathrm{V}_n(x) = n(1-a)^2/(1-a^2)\sigma_e^2 + 1/\mathrm{V}_0(x) \tag{16}$$

showing that the precision increases linearly with n. Thus the variances should be considered as known in advance, and only the changes in expected value depend on the observations made. It will not be necessary to keep the probabilities $p_x(n)$. It is sufficient to keep the expected value of X, and the probabilities $p_x(n)$ may at any stage be reproduced by the relation

$$p_x(n) = \Phi((x^+ - \mathrm{E}_n(x))/(\mathrm{V}_n(x))^{1/2}) - \Phi((x^- - \mathrm{E}_n(x))/(\mathrm{V}_n(x))^{1/2}). \tag{17}$$

As a consequence of this finding, we now redefine the state spaces of the hierarchic Markov process so that the state space of the main process holds

only one element, and the state space of the only possible subprocess becomes $\Omega_2 = \{\mu_1, \ldots, \mu_m\} \times \Omega_I$, where the set $\{\mu_1, \ldots, \mu_m\}$ represents alternative values of the expectation of X. In the following $\mu(i)$ will denote the expected value belonging to state $i \in \Omega_2$.

Put $E_n(x) = \mu(i)$ and $E_{n+1}(x) = \mu(j)$ for any $i, j \in \Omega_2$. If $E_n(x)$ and $E_{n+1}(x)$ satisfy equation (14), the transition probabilities of the subprocess are calculated as

$$p_{ij}^1(n) = \sum_x p_{xij}^1(n) p_x(n) \tag{18}$$

and otherwise

$$p_{ij}^1(n) = 0. \tag{19}$$

The expected production and reward given stage, state and action are calculated as

$$m_i^d(n) = \sum_x m_{xi}^d(n) p_x(n), \tag{20}$$

and

$$r_i^d(n) = \sum_x r_{xi}^d(n) p_x(n). \tag{21}$$

We have now arrived at an ordinary hierarchic Markov process that may be solved by usual methods as described by Kristensen (1988, 1991). Thus we are able to solve the asset chain situation as well as the quota situation.

5. Benefits from updating: a numerical example

In order to illustrate the benefits of updating, we shall consider a numerical example. Suppose that the expected production of an animal decreases linearly with age from an initial level at stage 0 according to the following relation:

$$m(n) = c_1 - c_2 n. \tag{22}$$

The observed total production of an animal during stage n is calculated as Y_n in equation (1), and we define the physical output as $m_{xi}^d(n) = Y_n$ for $d \in \{1, 2\}$, since a replacement is assumed to take place (and to be decided) at the end of a stage. In the replacement state, the physical output is zero. The reward gained at stage n is defined as

$$r_{xi}^d(n) = c_3 m_{xi}^d(n) - c_4(n) + c_5^d(n), \tag{23}$$

where c_3 is the unit price of the product, $c_4(0)$ is the price of a new animal for replacement and $c_4(n) = 0$ for $n > 0$, $c_5^1(n) = 0$ for all n, and $c_5^2(n)$ is the

value of an animal being replaced at stage n. In the replacement state, however, the reward is zero. We shall assume the value of an animal to decrease linearly over stages from an initial carcass value c_6 at stage 0 according to the relation

$$c_5^2(n) = c_6 - c_7 n. \tag{24}$$

A set of numerical values were chosen for σ_x, σ_e, a, N and the c-constants of equations (22) to (24). The selected values are summarised in Table 1. For X, the nine levels $]-\infty; -3.5]$, $]-3.5; -2.5], \ldots,]3.5; \infty[$ were distinguished. The levels are referred to as $1, \ldots, 9$, respectively. For the sums $I_n = X + e_n$, the 13 levels $]-\infty; -5.5]$, $]-5.5; -4.5], \ldots,]5.5; \infty[$, referred to as $1, \ldots, 13$, were considered. For the current *expected* value of X, the same levels are used as for X. Thus the total number of states in the subprocess becomes $9 \times 13 + 1 = 118$ (the last state added is the replacement state).

In order to be able to evaluate the benefit of updating, two alternative hierarchic models were formulated. In one model it was assumed that X was directly observable. In that case the nine levels of X were defined as states of the main process, and the 13 levels of I_n plus the replacement state were defined as 14 states in the subprocesses.

The second alternative represents a situation where X is not observable, and no updating of the belief in X is performed. In other words, the prior distribution of X is used during the whole lifetime of the animal. The same hierarchic design was used as in the updating situation, but in equations (18), (20) and (21) the *initial* state distribution $p_1(0), \ldots, p_9(0)$ of x was used at all stages instead of $p_1(n), \ldots, p_9(n)$. In all three models, optimal policies under the discounting criterion referring to the asset chain situation were calculated using the optimisation cycle of Kristensen (1988), and the economic results were measured by the present value of the entire infinite process calculated just before purchase of a new animal. The results are compared in Table 2.

Table 1. *Selected values for the parameters of the numerical example in section 5*

Parameter description	Symbol	Value
Standard deviation of basic state X	σ_x	2
Standard deviation of temporary effect e_n	σ_e	2
Autoregression coefficient	a	0.5
Maximum age (stages) of an animal	N	10
Expected production of an animal at stage 0	c_1	20
Expected reduction per stage in production	c_2	0.5
Unit price of product	c_3	10
Price of new animal for replacement	$c_4(0)$	200
Carcass value of an animal at stage 0	c_6	120
Reduction per stage in carcass value	c_7	4
Discount factor per stage		0.95

Table 2. *Present values under optimal policies in three alternative situations, representing different levels of knowledge on the basic state X*

Level of knowledge	Present value	Relatively
Only prior knowledge available	3120	100
Updating of knowledge as observations are done	3435	110
Complete knowledge of the value of X	3449	111

As it appears, the updating of knowledge increases the economic result by 10% compared to a situation with no updating. Furthermore, the result under updating is very close to the result under complete knowledge. The autoregression coefficient a is a measure of the constancy of the random effect e_n. In the extreme case $a = 0$, the variation of e_n is just noise in the observation of X, whereas in the opposite situation with $a = 1$, e_n will be constant over time, making X and e_n measure exactly the same, namely a permanent characteristic of the animal. In order to study the effect of a on the benefit of updating the value was varied from 0.1 in steps of 0.1 to a value very close to 1.

It appears from the results that there is practically no benefit of updating when a is close to 1, whereas the benefit is considerable for small values of a. The reason is that for values of a close to 1, X and e_n express almost the same, and in that case only the directly observable sum is of interest. Further, it appears from equation (16) that the precision of the belief concerning X increases only very little over stages when a is close to 1, since the increase per stage is proportional to the factor $(1-a)^2/(1-a^2)$ which decreases for a increasing towards 1. Therefore, the benefit from updating is very small for such high values of a.

In the consideration of the influence of a on the benefit of updating, two arguments lead in the same direction: (1) the precision increases only a little for values of a close to one, and (2) the economic significance of distinguishing X and e_n vanishes as a converges towards 1. We shall now consider the influence of combined values of the standard deviations σ_x and σ_e. Concerning this question we are less fortunate than when we considered the influence of a. On the one hand, we know from equation (16) that the increase in precision concerning the belief in X is small when σ_e is big. On the other hand, if σ_e is small compared to σ_x, the directly observed sum $I_n = X + e_n$ expresses almost the same as X. Therefore, we must expect the benefit of updating to be relatively small, since we are told something that we (almost) knew in advance. Thus, we have conflicting views, and only the results may show us the true influence of σ_n.

It was found that for a fixed value of σ_x the benefit *increases* with σ_e. We may therefore conclude that even though the information acquired is rather vague for high values of σ_e, it is at least *new* and therefore more valuable

than for low values of σ_e. It also appeared that the benefit from updating increases even more with σ_x. This is not surprising, since a great variation in a variable automatically increases the economic value of information concerning the true level of the variable.

6. Discussion

The numerical example in section 5 has shown that the benefit of updating may be considerable. In order to evaluate the method more carefully, we shall now compare it to the methods typically used for modelling trait variations in replacement studies in the literature. In most cases the state variables have been defined from the directly observable variables $I_1, ..., I_N$, but the variables have typically not been regarded as sums of underlying unobservable effects. Examples in dairy cattle are Giaever (1966), Smith (1971), McArthur (1973), Kristensen and Østergaard (1982), van Arendonk (1985a, 1985b) and Kristensen (1986, 1987). An example in sows is the work of Huirne et al. (1988).

All authors mentioned have been aware that part of the observed value of I_n is due to a permanent property of the animal, even though it has not been formulated directly as is done in equations (1) and (2). Without such a model, the ideal way to take the permanent effect into account is to use *all* previous observations $I_1, ..., I_n$ in the prediction of I_{n+1}. Thus all observed values should be kept as state variables in the model. Therefore, the size of the state space becomes prohibitive if an appropriate number of levels is defined for each of them.

The most common way of dealing with this problem in the literature is to assume that the last two or three observations $(I_{n-2})\ I_{n-1}, I_n$ are sufficient in the prediction of I_{n+1}. Thus we only have to keep two or three state variables instead of all n observations. This method was used in dairy cows by Smith (1971), van Arendonk (1985b) and Kristensen (1986, 1987), all keeping two observations of I_n. In sows, Huirne et al. (1988) used the same approach keeping three observations of litter size.

In the approach taken in this paper, I_n is assumed to be the sum of two (or more) underlying unobservable effects as defined in equation (1). Under these conditions, it is easily shown that the expected value and variance of I_{n+1} given $I_1, ..., I_n$ are calculated as

$$\mathrm{E}(I_{n+1} \mid I_1, ..., I_n) = (1-a)\mathrm{E}(X \mid I_1, ..., I_n) + aI_n \tag{25}$$

and

$$\mathrm{V}(I_{n+1} \mid I_1, ..., I_n) = (1-a)^2\,\mathrm{V}(X \mid I_1, ..., I_n) + (1-a^2)\sigma_e^2. \tag{26}$$

From equations (25) and (26) we are able to conclude that by keeping only the current expectation of X and the most recent observation I_n, the predic-

tion of I_{n+1} is exactly the same as if all previous values $I_1, \ldots, I_n$ were kept and used in the prediction! As shown in equation (16), the conditional variance of X is independent of the observations made and known in advance.

We are able to conclude that if the model represented by equations (1) and (2) is true, we only need two state variables in the model to obtain the same precision as if all previous observations of the trait in question were kept as state variables. Since the dimension of the model is the more limiting restriction in practical applications of Markov decision processes, this is an important contribution to the problem of reducing the state space without loss of precision. The generalisation of this conclusion to the multi-trait model sketched in section 2, equations (3) and (4), is that the number of necessary state variables equals the sum of directly observed traits and the number of unobservable permanent effects.

It must be emphasised that the multi-trait formulation in equations (3) and (4) is not necessary in all cases where several random traits are observed. If, an addition to Y_n, another trait Z_n is observed, and this trait only affects Y_n through the function $m(\cdot)$ so that I_n and Z_n are independent there is no problem in treating such a case within the single-trait model. The transition probabilities concerning I_n described in this paper simply have to be multiplied by those of Z_n.

The definition of a directly observable trait as a sum of underlying unobservable effects provides a framework for prediction of the future revenues in an optimal way given the information available at the time of decision. Thus the conclusion of the present study is that the developed knowledge updating technique seems to be an appropriate solution to the uniformity problem of defining and measuring the traits of an animal considered for replacement in the asset chain situation as well as in the quota situation.

References

Ben-Ari, Y., Amir, I. and Sharar, S. (1983). Operational replacement decision model for dairy herds. *Journal of Dairy Science* 66, 1747–1759.

Ben-Ari, Y. and Gal, S. (1986). Optimal replacement policy for multicomponent systems: an application to a dairy herd. *European Journal of Operational Research* 23, 213–221.

DeGroot, M. H. (1970). *Optimal Statistical Decisions.* New York: McGraw-Hill.

Giaever, H. B. (1966). *Optimal Dairy Cow Replacement Policies.* Ann Arbor, MI: University Microfilms.

Harrison, P. J. and Stevens, C. F. (1976). Bayesian forecasting. *Journal of the Royal Statistical Society* B 38, 205–247.

Huirne, R. B. M., Hendriks, T. H. B., Dijkhuizen, A. A. and Giesen, G. W. J. (1988). The economic optimization of sow replacement decisions by stochastic dynamic programming. *Journal of Agricultural Economics* 39, 426–438.

Jensen, F. V., Olesen, K. G. and Andersen, S. K. (1990). An algebra of Bayesian belief universes for knowledge-based systems. *Networks* 29, 637–659.

Jørgensen, E. (1992). Sow replacement: reduction of state space in dynamic programming model and evaluation of benefit from using the model. Unpublished manuscript.

Kristensen, A. R. (1986). A model to describe and predict the production and food intake of the dairy cow. *Acta Agriculturæ Scandinavica* 36, 305–317.

— (1987). Optimal replacement and ranking of dairy cows determined by a hierarchic Markov process. *Livestock Production Science* 16, 131–144.

— (1988). Hierarchic Markov processes and their applications in replacement models. *European Journal of Operational Research* 35, 207–215.

— (1989). Optimal replacement and ranking of dairy cows under milk quotas. *Acta Agriculturæ Scandinavica* 39, 311–318.

— (1991). Maximization of net revenue per unit of physical output in Markov decision processes. *European Review of Agricultural Economics* 18, 231–244.

— (1992a). Markov decision programming techniques applied to the animal replacement problem. Unpublished manuscript.

— (1992b). Optimal replacement in the dairy herd: a multi-component system. *Agricultural Systems* 39, 1–24.

Kristensen, A. R. and Østergaard, V. (1982). Optimalt udskiftningstidspunkt for malkekoen fastlagt ved en stokastisk model. (In Danish with English summary). *Beretning fra Statens Husdyrbrugsforsøg* 533.

McArthur, A. T. G. (1973). Application of dynamic programming to the culling decision in dairy cattle. *Proceedings of the New Zealand Society of Animal Production* 33, 141–147.

Monahan, G. (1982). A survey of partially observable Markov decision processes: theory, models and algorithms. *Management Science* 28, 1–16.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* San Mateo, CA: Morgan Kaufmann.

Preinrich, G. (1940). The economic life of industrial equipment. *Econometrica* 8, 12–44.

Smith, B. J. (1971). The dairy cow replacement problem – an application of dynamic programming. *Bulletin Florida Agricultural Experiment Station* 745.

Tatman, J. A. and Shachter, R. D. (1990). Dynamic programming and influence diagrams. *IEEE Transactions on Systems, Man, and Cybernetics* 20, 365–379.

Terborgh, G. (1949). *Dynamic Equipment Policy.* New York: McGraw-Hill.

van Arendonk, J. A. M. (1984). Studies on the replacement policies in dairy cattle. I. Evaluation of techniques to determine the optimum time for replacement and to rank cows on future profitability. *Zeitschrift für Tierzüchtung und Züchtungsbiologie* 101, 330–340.

— (1985a). A model to estimate the performance, revenues and costs of dairy cows under different production and price situations. *Agricultural Systems* 16, 157–189.

— (1985b). Studies on the replacement policies in dairy cattle. II. Optimum policy and influence of changes in production and prices. *Livestock Production Science* 13, 101–121.

van Hee, K. M. (1978). *Bayesian Control of Markov Chains.* Mathematical Centre Tract 95. Amsterdam: Mathematical Centre.

Wessels, J. (1968). Decision rules in Markovian decision processes with incompletely known transition probabilities. Unpublished dissertation, University of Technology, Eindhoven.

Anders Ringgaard Kristensen
Department of Animal Science and Animal Health
The Royal Veterinary and Agricultural University
Rolighedsvej 23
DK-1958 Frederiksberg C
Denmark