DECISION SUPPORT SYSTEMS IN ANIMAL PRODUCTION : A BAYESIAN FUTURE

E. JØRGENSEN

Danish Institute of Agricultural Science, P.O. Box 50, DK-8830 Tjele, Denmark
E-mail: Erik.Jorgensen@agrsci.dk

ABSTRACT

Within decision theory the use of Bayesian methodology is well established. However, primarily due to complexity problems the use in applied Decision Support Systems within animal production has been very limited. Recently, methodological developments have removed many of these problems, and the time seems to be ripe for wide-scale applications. Bayesian methods handles decision making under uncertainty especially in situations, where the amount of information varies. The case for use of the Bayesian framework will be presented, and its potential illustrated with on-going research work within herd monitoring, diagnosis, and decision support.

INTRODUCTION

The Bayesian methodology is well-established as a theoretical framework for decision making, see e.g. De Groot (1970), Anderson et al. (1977). However, primarily due to complexity problems the use in applied Decision Support Systems within animal production has been very limited. Recently, methodological developments have removed many of these problems, i.e., . techniques such as Bayesian networks, (Jensen, 1996; Lauritzen & Spiegelhalter, 1988), Dynamic Linear Models (West & Harrison, 1997), Markov Chain Monte Carlo Simulation (Gilks et al., 1996) and Hierarchic Markov Processes, (Kristensen, 1993b). Thus the time seems ripe for wide-scale applications.

As an introduction, a general framework is shown in Fig. 1 of the tasks within production planning and control.

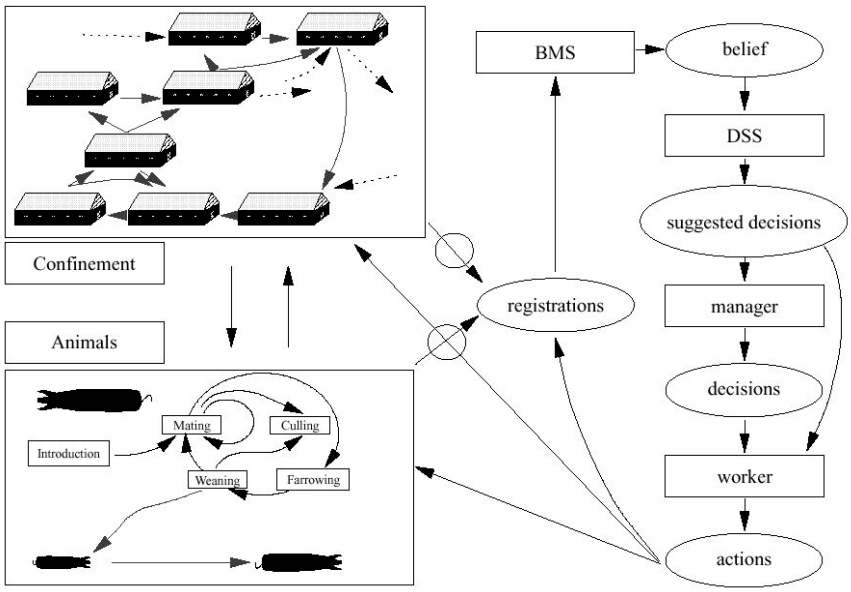

FIGURE 1: The Animal Production and Information Processing System

The figure illustrates the three main components of the system. Housing/confinement, animals and information processing. A natural starting point in the present context is the *registration*. A necessary prerequisite for any benefit of such registrations is that they are used as a basis for decision making, typically at the operational and tactical level. This requires an appropriate data processing, a communication of some kind of belief in the true state of the animal(s) and/or one or more suggested actions. The term *Belief Management System*, (BMS), is used for the system that handles the *data processing*. The purpose of the BMS is to process the data such that the resulting information is suitable for decision making. This belief may then be used for decision support, i.e., support for selecting appropriate actions. Obviously, the term Decision Support System (DSS) is appropriate for systems handling this part.

The case for use of the Bayesian framework will be presented, and its potential illustrated with on-going research work within herd monitoring, diagnosis, and decision support.

BELIEF MANAGEMENT SYSTEMS

With the appearance of sensors and electronic equipment in general, a whole new range of registrations are potentially available to the farmer. The possibilities comprise electronic identification, automatic weighing, temperature and activity measurements, automatic registration of feed and water intake, geographic positioning of individual animals and several registrations through video recordings using image analysis. The more or less continuous data-flow makes techniques based on Kalman filtering methods an obvious choice. Recent examples are heat detection (de Mol, 1999), daily gain (Madsen & Ruby, 2000), water consumption (Madsen, 2000), and monitoring of pregnancy/farrowing rates (Jørgensen & Toft, 1999; Thysen & Enevoldsen, 1994). The purpose of these systems is to allow the farmer to detect changes in the monitored variable.

Within the present framework a different formulation is more appropriate. The system (herd, animal,...) may exists in different states. The BMS calculates the probability that the system is in a given state depending on the observed registrations. Thus the BMS improves our knowledge (increases the precision) of the state of the system. Depending on the state it may be worthwhile to put some efforts into changing the state.

The *belief* returned from a BMS is a probability distribution, i.e. for each possible state of the system, the BMS returns a probability.

The major challenge is to improve the representation and utilisation of prior knowledge, e.g. heat detection should take the time since weaning or last observed heat into account when calculating the prior probability of heat signs. Another challenge is to utilize the hierarchic structure of the data. Data originates from an animal, in a pen, in a section in a herd, in a population of herds. Even though we have no registrations on the individual animal, we may be able to utilize other sources of information from the hierarchy. However, in this process we have to take the reason for the missing data into account. To illustrate as a result of selection, short lactation length indicates lactation yield, low age at slaughter indicates high daily gain and few parities indicates low litter size. Toft & Jørgensen (2001) show how information may be used while taken this into account, and they illustrate how herd population data may be used when within herd information is insufficient.

DECISION SUPPORT

The belief may then be used for decision support, i.e., support for selecting appropriate actions. Strictly speaking, the DSS should help the farmer to select the *action* with highest expected utility. Because the *belief* reflects the probability distribution, the expected utility of the *actions* increases with the precision in the *belief*.

However, when reviewing existing DSS systems the impression is that many model parameters are point estimates either from populations studies or from expert assessment. Using the Bayesian approach, such parameters should have *beliefs* with low precision.

Three different approaches towards decision support is considered. The use in error detection (*Diagnosis/ Analysis*), the use within a Markov process setting, and finally, the use within Monte Carlo simulation models. The first point may be considered as a static problem, while the latter two are dynamic, i.e., needs to consider the dynamic acquisition of new information.

Error Detection, Diagnosis/Analysis

When the monitoring systems indicates that the state of the system has changed, it is of interest to make an analysis of the herd, i.e., to find the most plausible causes of the change. This error detection problem may be considered as an expert system problem, and rule-based or knowledge-based systems abound in the literature, e.g., Huirne et al. (1991), Enting et al. (2000) and chapter 1.6 in Stärk (1998). However, Bayesian network techniques (Jensen, 1996) offer full probabilistic formulation and inference, and thus, a more natural approach in the agricultural domain. Examples are Hogeveen et al. (1994) and a current project concerning respiratory disease in slaughter pigs (Jørgensen, 2000b). Using such techniques the probability of the possible causes will be found, and the appropriate action may be taken to remove the cause. The network handles the situation, where several plausible causes exists, and in contrast to the rule-based approaches handles the *explaining away* phenomenon. The problem may be due to both cause *A* and cause *B,* but if new evidence indicates *A* the probablity of *B* is reduced. Of course, the selected action should reflect both the probability of the cause and the cost of removing the cause

In addition the Bayesian network may *learn* from the observations and thus improve over time. This corresponds to parameter estimation in ordinary statistical models.

Markov Decision Processes

The Markov process (MP) approach has information acquisition as an inherent element in the optimisation procedure, and may be framed to take information flow into account, e.g. Kristensen (1993a). The introduction of hierarchic Markov processes (Kristensen, 1988) has allowed quite complex models for replacement decisions such as Houben et al. (1995), and the recent possibility to include several hierarchic levels (Kristensen & Jørgensen, 2000) looks promising with respect to the inclusion of other decisions and planning horizons.

However, the formulation is based on a complete knowledge of the current state and the future transition probabilities. Often, we need to consider decision support within domains, where this assumption is not fulfilled. Without the assumption, the complexity becomes very large even for unrealistic small problems. On-going research has indicated, how we may proceed with approximate techniques, such as Limited Memory Influence Diagrams

(Lauritzen & Nilsson, 2000; Höhle et al., 2000). Using such techniques, we may obtain lower and upper bounds on the optimal solutions.

Monte Carlo simulation models

Monte Carlo simulation models of animal herds have been used for decision support to a certain extent. These stochastic models are natural candidates for use within a Bayesian framework. However, two major issues needs to be addressed. Firstly, the system parameters (*"State of nature"*) is often assumed known. Secondly, the current state of the herd and the decision makers current knowledge of the state need to be considered separately. With respect to the first problem, the necessary methodology for representing the precision of the parameters is available, both when considering independent parameters, and when considering parameters that are dependent, because of supplementary evidence as discussed in Jørgensen (2000c). The idea is to sample the parameter values from the appropriate probability distribution. The supplementary evidence may be handled within a Bayesian network, e.g. an expert system, and the Bayesian network is able to generate random samples from the joint distribution of the state variables, or within a hierarchical model, where Markov Chain Monte Carlo simulation may generate the requested samples. However, a major problem remains. When supplementary evidence corresponds to output from the simulation model the calibrating of model parameters is *not* straightforward. Attempts to solve this problem is discussed in Givens (1993) and Jørgensen (2000a). To the authors knowledge there are currently no simulation models that consequently distinguish between the state of the system and the state of the knowledge, when simulating the decisions in the herd.

DISCUSSION

The Bayesian framework seems ideally suited to the purpose of decision support within animal production, even though only few and isolated attempts have been made until now. Most important, the Bayesian framework enables combination of information from different sources in a coherent and reproducible manner. Currently, the most important task is to improve the assessment of the parameters used in the decision support systems, that is to establish belief management systems that handles the current registrations in the herds, and to ensure that the results from the BMS are used directly in the DSS.

Despite recent improvements in methodology, we may still face complexity problems, when considering systems for decision support. However, it is not a valid solution to this problem, just to ignore the inherent uncertainty that faces the decision maker. Rather, we should develop approximate methods that presents the farmer with near-optimal solutions.

REFERENCES

Anderson, J.R., J.L. Dillon, & J.B. Hardaker (1977). *Agricultural decision analysis.*. The Iowa State University Press, Ames, Iowa.

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

Enting, J., R.B.M. Huirne, A.A. Dijkhuizen, & M.J.M. Tielen (2000). Zovex, a knowledge-integrated computer system to support health management on pig farms. *Computers and Electronics in Agriculture*, 26 pp.13-35.

Gilks, W.R., S. Richardson, & D.J. Spiegelhalter (1996). *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London.

Givens, G.H. (1993). *A Bayesian Framework and Importance Sampling Methods for Synthesizing Multiple Sources of Evidence and Uncertainty Linked by a Complex Mechanistic Model. Ph.D. dissertation*. Department of Statistics, University of Washington, Seattle.

Höhle, M., E. Jørgensen, & D. Nilsson (2000). Modeling with LIMIDs - Examplified by Disease Treatment in Slaughter Pigs. In *Proc. International Symposium on Pig Herd Management Modeling and Information Technology Related. September, 18-20 Lleida, Spain.*, pp. 17-26.

Hogeveen, H., B.Thiesson, I.Thysen, E.N. Noordhuizen-Stassen, H. Maurice, S.K. Andersen, T.J.G.M. Lam, & T. van Werven (1994). Mastitits pathogen diagnosis using Bayesian belief networks. In H. Hogeveen, *Knowledge-based methods for automated mastitis diagnosis on dairy farms. Ph.D. thesis.*, pp. 129-144. Faculteit Diergeneeskunde, Universiteit Utrecht, The Netherlands, Utrecth.

Houben, E.H.P., R.B.M. Huirne, A.A. Dijkhuizen, & A.R. Kristensen (1995). Optimal replacement of mastitis cows determined by a hierarchic Markov process. *Journal of Dairy Science*, 77 pp.2975-2993.

Huirne, R.B.M., A.A. Dijkhuizen, & G.B.C. Backus (1991). Validation of an integrated decision support and expert system for analysis of individual sow-herd performance. *Computers and Electronics in Agriculture*, 6 pp.71-86.

Jensen, F.V. (1996). *An introduction to Bayesian Networks*. UCL Press.

Jørgensen, E. (2000a). Calibration of a Monte Carlo Simulation Model of Disease Spread in Slaughter Pig Units. *Computers and Electronics in Agriculture*, 25 pp. 245-259. URL: `http://www.jbs.agrsci.dk/~ejo/dinapig/compefit.ps`.

Jørgensen, E. (2000b). Elements of Bayesian network specification in an animal health research project. *Internal report, Biometry Research Unit, Danish Institute of Agricultural Sciences*, 2000-2 pp.1-24.URL: http://www.jbs.agrsci.dk/ `~ejo/dinapig/diag1504a.pdf`.

Jørgensen, E. (2000c). Monte Carlo Simulation Models: Sampling from the joint distribution of 'State of Nature'-parameters. In *Proc. "Economic modelling of Animal Health and Farm Management". November 23-24, 2000 Wageningen*.

Jørgensen, E. & N. Toft (1999). A Bayesian network based monitoring system for sow management. *Computers and Electronics in Agriculture*, subm. URL: `http://www.jbs.agrsci.dk/~ejo/dinapig/compsow.ps`.

Kristensen, A.R. (1988). Hierarchic Markov processes and their applications in replacement models. *European Journal of Operations Research*, 35 pp. 207-215.

Kristensen, A.R. (1993a). Bayesian updating in hierarchic Markov processes applied to the animal replacement problem. *European Review of Agricultural Economics*, 20 pp.223-239.

Kristensen, A.R. (1993b). *Markov Decision Programming Techniques Applied to the Animal Replacement Problem. Doctoral Dissertation*. Dina KVL, Institut for Husdyrbrug og Husdyrsundhed, København.

Kristensen, A.R. & E. Jørgensen (2000). Multi-level hierarchic Markov processes as a framework for herd management support. *Annals of Operations Research*, 94 pp.69-89.

Lauritzen, S.L. & D.Nilsson (2000). Representing and Solving Decision Problems with Limited Information. *Management Science*, subm.

Lauritzen, S.L. & D.J. Spiegelhalter (1988). Local computations with probabilities on graphical structures and their application to expert system (with discussion). *Journal of the Royal Statistical Society, series B*, 50 pp.157-224.

Madsen, T.N. (2000). A model for detection of changes in drinking behaviour of young pigs. In *Proc. International Symposium on Pig Herd Management Modelisation. Lleida, Spain.*, pp. 1-10. URL: `http://labor/etsea.udl.es/symposium/Madsen.ps`.

Madsen, T.N. & V. Ruby (2000). An application for early detection of growth rate changes in the slaughter-pig production unit. *Computers and Electronics in Agriculture*, 25 pp.261-270.

de Mol, R.M. (1999). Description of a detection model for oestrus and disease in dairy cattle based on time series analysis combined with a Kalman filter. *Computers and Electronics in Agriculture*, 22 pp.171-185.

Stärk, K.D.C. (1998). *Systems for prevention and control of infectious diseases in pigs. Ph.D. thesis*. Massey University, Palmerston North, New Zealand.

Thysen, I. & C. Enevoldsen (1994). Visual Monitoring of Reproduction in Dairy Herds. *Preventive Veterinary Medicine*, 19 pp.189-202.

Toft, N. & E.Jørgensen (2001). Estimation of farm specific parameters in a longitudinal model for litter size with variance components and random dropout. .*Livestock Production Science*, subm.

West, M. & J. Harrison (1997). *Bayesian forecasting and dynamic models*. Springer-Verlag, New York, second edition.