INFORMATION FROM ON-LINE LIVE WEIGHT ASSESSMENT FOR OPTIMAL SELECTION OF SLAUGHTER PIGS FOR MARKET

Anders Ringgaard Kristensen, ark@dina.kvl.dk
Department of Animal Science and Animal Health
The Royal Veterinary and Agricultural University
Grønnegårdsvej 8, DK-1870 Frederiksberg C
Copenhagen, Denmark

One of the most labor intensive tasks of traditional slaughter pig production is the weighting of pigs for market. In an on-going project, methods for automatically assessing live weights through image analysis are being developed. Below, we present the framework of a farm level model to optimize selection of slaughter pigs for market. The optimization tool used in this study is a hierarchical Markov process. The state space is defined in such a way that observations from the on-line live weight estimation may serve as input for the decision support system. In that way the precision of the underlying predictions of future growth is improved. The aim is create a decision support being able to inform the farmer about the number of pigs that are ready for market per pen. The MLHMP software system will be used to implement the model.

1. Introduction

There is a growing need for computer-based methods for management support in European pig production, as the average herd size is increasing. When constructing new production systems, one of the main objectives has been to reduce the input of human labor. The number of animals managed per person is therefore also increasing. Simultaneously, the possibilities for real-time monitoring of production have improved, as a consequence of modern computer technology and the development of improved statistical methods. One of the most labor-intensive tasks of traditional slaughter pig production is the weighing of pigs for market. In an on-going project, methods for automatically assessing live weights through image analysis are being developed. When the equipment is installed in a production system, online estimation of live weights of the pigs is available throughout the fattening period. If the estimates are processed appropriately, they serve as valuable input for prediction of growth. They may therefore be used as observations for a decision support tool for optimally selecting slaughter pigs for market.

The problem of optimal selection of pigs for market can be regarded as a sequential decision problem involving decisions at two different levels: the animal level and the batch level. At the animal level, decisions comprise of selecting individual pigs for market based on observation of live weight of live weight and, at batch level, the decision when to terminate the batch (slaughter the remainder of the batch and insert a new batch of weaners). At batch level, the decision is based partially on observations of the number of remaining pigs and their live weight distribution, and partially on the operational constraints of the slaughter pig unit. (An example of an important constraint is the weaner supply. Although aspects of the optimal marketing problem have been dealt with in the literature (Jørgensen, 1993; Kure, 1997; Toft, 2000), online live weight assessment is a new aspect, which is included in the present study.

The framework of a farm level model for optimization of slaughter pig marketing is presented below. The optimization tool used in this study is a hierarchical Markov process. The state space is defined in such a way that observations from the on-line live weight estimation may serve as input for the decision support system. In that way the precision of the underlying predictions of future growth is improved. The aim is create a decision support being able to inform the farmer about the number of pigs that are ready for market per pen. The MLHMP software system[1] (Kristensen, 2003) will later be used to implement the model.

2. The monitoring system

The precise design of the monitoring system is not available yet. We shall therefore make some assumptions concerning the output, based on the goals of the project.

The basic assumption is that if a pen with n pigs is monitored, the system is able to supply n weight estimates $\hat{w}_i(t) = w_i(t) + e_{it}, i = 1,\ldots,n$, where $\hat{w}_i(t)$ is the observed weight of the ith pig at time t, $w_i(t)$ is the true live weight, and $e_{it} \sim N(0, s^2)$ is the measurement error, where we assume that the precision (i.e. $1/s^2$) is known. It must be emphasized that the pigs are not identified. We only know that the observed weight distribution is given by $\hat{w}_1,\ldots,\hat{w}_n$. In principle, we are not able to identify a particular weight estimate with a specific pig. Note that through the weight estimates, we are also able to answer questions of the type: "How many pigs in the pen have an estimated weight higher than d ?". This is important for the implementation of a delivery policy based on threshold weights, as described below.

3. A model for description of live weight over time for slaughter pigs

In order to describe the growth of the pigs in a specific pen, we use a dynamic linear model (DLM) with Kalman filtering, as described by West & Harrison (1997).

We shall assume that an estimated average growth curve $\bar{y}(t)$ has been estimated from herd specific data. In the weight interval relevant for slaughter pigs, the average curve is more or less linear, but we shall not in this paper make any specific assumptions concerning the shape of the growth curve. It is only assumed that a known herd specific growth curve $\bar{y}(t)$ exists. Even though the general curve is known, the growth potential of the pigs in a particular pen may deviate from the average value given by $\bar{y}(t)$. To be more specific, we assume the true weights $w_i(t)$ to be distributed as $N(\bar{y}(t)L, s_w^2)$, where L is a scaling factor related to the pigs currently occupying the pen. The value of L is in principle unknown, but the initial belief is that true value is distributed as $N(1, s_L^2)$, where s_L reflects the variation between pens. As observations are done for the pen, we may update our belief in the true value of L. Even though the true mean of $w_i(t)$ is assume to be $\bar{y}(t)L$, the average value of the true live weights (i.e. $\bar{w}(t) = \left(\sum_i w_i\right)/n$) will deviate from that value because of the sample uncertainty. We may express this as

$$\bar{w}(t) = \bar{y}(t)L + e(t) \quad (1)$$

where $e(t) \sim N(0, s_e^2)$ represent the sample uncertainty. If we further add the measurement error we obtain the following observation equation for the DLM:

$$\hat{\bar{w}}(t) = \bar{y}(t)L + e(t) + \bar{e}_t \quad (2)$$

[1] http://www.prodstyr.ihh.kvl.dk/software/mlhmp.html

The true value of L is assumed to be a permanent trait of the present flock, but as concerns the sample uncertainty $e(t)$, it is assumed to be auto correlated over time. If we, for instance, assume weekly intervals, we have

$$e(t) = ae(t-1) + h_t, \qquad (3)$$

where a is an auto regression coefficient, and $h_t \sim N(0, s_h^2)$ is an independent random term. Eq. (3) will be one of the system equations of the DLM.

As described so far we have only used the average observed weight as a source of information. Since we have actually observed $\hat{w}_1(t), \ldots, \hat{w}_n(t)$ we have much more information available about the distribution. We could for instance calculate the sample variance, but the problem is that the distribution of the sample variance is not normal. It is therefore difficult to include it in a DLM. Instead we shall use the 0.16 sample quantile $\hat{w}_{(0.16)}$ as a further description of the underlying distribution of live weights. Since the distribution of a sample quantile *is* (asymptotically) normal around the true quantile, we have

$$\hat{w}_{(0.16)}(t) = \bar{y}(t)L + e(t) - r(t) + t_t \qquad (4)$$

where $r(t)$ is the standard deviation of $\hat{w}_i(t)$, and $t_i \sim N(0, s_t^2)$ expresses the sample uncertainty. In Eq. (4) we use the well-known fact that in a normal distribution, the 0.16 quantile is the mean minus the standard deviation. As concerns the development of $r(t)$ (the standard deviation) over time, we assume that it increases linearly (in case of linear growth this assumption corresponds to constant coefficient of variation). This leads to the following relation

$$r(t) = \frac{t}{t-1} r(t-1) \qquad (5)$$

We are now ready to specify the full DLM. The two-dimensional observation vector is $Y_t = \left(\hat{\bar{w}}(t) \quad \hat{w}_{(0.16)}(t)\right)'$ and the parameter vector is $q_t = \left(L \quad e(t) \quad r(t)\right)'$. The observation equation is

$$\begin{bmatrix} \hat{\bar{w}}(t) \\ \hat{w}_{(0.16)}(t) \end{bmatrix} = \begin{bmatrix} \bar{y}(t) & 1 & 0 \\ \bar{y}(t) & 1 & -1 \end{bmatrix} \begin{bmatrix} L \\ e(t) \\ r(t) \end{bmatrix} + \begin{bmatrix} \bar{e}_t \\ t_t \end{bmatrix} \qquad (6)$$

or in short notation

$$Y_t = F_t q_t + V_t \qquad (7)$$

where the matrix F_t and the random vector V_t are easily identified by comparing Eqs. (6) and (7). The system equation of the DLM is

$$\begin{bmatrix} L \\ e(t) \\ r(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & \frac{t}{t-1} \end{bmatrix} \begin{bmatrix} L \\ e(t-1) \\ r(t-1) \end{bmatrix} + \begin{bmatrix} 0 \\ h_t \\ 0 \end{bmatrix} \qquad (8)$$

or in short notation

$$q_t = G_t q_{t-1} + x_t \qquad (9)$$

where the matrix G_t and the random vector x_t are identified by comparing Eqs. (8) and (9).

In addition to the observation and system equations we also need to specify the initial distribution of the parameter vector $q_0 \sim N(m_0, C_0)$. In other words, we must specify an initial mean vector m_0 and an initial variance covariance matrix C_0.

Having defined the DLM, we may at regular intervals use the Kalman filter as described by West & Harrison (1997) for updating of the distribution of the parameter vector. The updated mean vector m_t will depend on the observed values whereas the updated variance covariance matrix C_t will be independent of the observed values (but it will depend on the *number* of observations. We shall not in this paper describe the updating algorithms, but only conclude that a well-described method is available. Neither shall we discuss how to estimate the underlying observation and system variances, the initial distribution of q_0 or the auto regression coefficient a .

4. Optimization of delivery policies

An optimization model based on a multi-level hierarchical Markov process (Kristensen & Jørgensen, 2000) will be built. In this particular case it is, however, expected that two (or perhaps three) levels will suffice. We shall give a brief description of the framework of such a model. The system being modeled is a pen with a given capacity expressed by the number of pigs, *n*, it may contain. It is assumed that the weaner supply is constant and inflexible in the sense that the pen must be completely emptied before the new batch is inserted. Delivery of pigs will occur in a fixed number of possible deliveries starting a predetermined number of weeks (e.g. 4) before termination. After the first potential delivery, pigs are selected and delivered once a week

For examples of working models of this type in other application areas, reference is made to Lien et al. (2003), Nielsen & Kristensen (2002) or Kristensen & Søllested (2002).

Founder level

The founder process is an infinite stage Markov decision process, where each stage corresponds to a group of pigs occupying the pen. It has not been decided yet whether or not to include some kind of auto regression between batches. In that case, the initial estimate for *L* given the previous batch may be defined as a state variable. Otherwise, there will be only one dummy state and one dummy decision at the founder level.

Child level

A child process is initiated when a new batch of pigs is inserted into the pen, and it is terminated when the last pig of the pen has been delivered to the slaughterhouse. The duration of the first stage is from insertion to the first potential date of delivery. No observations are done in the Markov process during the first stage. Accordingly, only one dummy state and one dummy decision are defined here.

From the first potential delivery until the date of termination, stage length will be one week. The state variables defined for stage $t > 1$ will be the estimated updated parameter values, i.e. $\hat{L}_t$ for *L*, $\hat{e}(t)$ for $e(t)$ and $\hat{r}(t)$ for $r(t)$ in combination with the number of remaining pigs n_t in the pen. Even though no observations are made in the initial stage $(t = 1)$ of the Markov process, the monitoring system will concurrently provide estimates for the live weights in the pen. Those estimates are used for regular updating of the parameter estimates, so that in the second stage $(t = 2)$ of the Markov process, the observed values $\hat{L}_2, \hat{e}(2), \hat{r}(2)$ are based on (in principle) all observations from the time of insertion to the first potential delivery.

The decision to be made at stage $t > 1$ is to deliver pigs with an observed live weight exceeding d_t, where d_t is a threshold weight taken from a finite set $\Delta = \{d^1, \ldots, d^m\}$. The lowest possible value, d^1, is assumed to be so low that if it is chosen, it corresponds to delivering *all* remaining pigs of the pen.

Parameters of the child processes

The rewards of a stage are calculated as the economic net returns during the stage given the state observed and the decision made.

The transition probabilities are calculated from the underlying model for live weight described in Section 3. At any stage, we have estimates, $\hat{L}_t + \hat{e}(t)$ and $\hat{r}(t)$ for the current sample mean and standard deviation, respectively. For a given selected threshold weight d_t we may calculate the probability of delivering $1,\ldots,n$ pigs for slaughter as well as the conditional probability distributions for the estimates, $\hat{L}_{t+1}, \hat{e}(t+1), \hat{r}(t+1)$, at next stage. A selection bias will occur as soon as some of the pigs have been delivered from the pen. This means that the observed sample mean $\hat{\bar{w}}(t)$ will be biased. We will have to adjust for this bias under the assumption that the heaviest pigs are always delivered first and that the weight rank of a pig in the pen does not change during the relative few weeks where pigs are delivered. When there are only very few pigs left in the pen also the sample quantile $\hat{w}_{(0.16)}(t)$ will be biased.

A possible extension

As the model has been formulated until now, it is assumed, that the estimates $\hat{L}_t, \hat{e}(t), \hat{r}(t)$ are only updated once per stage (i.e. once a week). Since, however, the monitoring system produces observations more or less continuously, it would be possible to update the estimates more often – for instance daily. Because the number of pigs is constant between two deliveries, it would be possible to extend the model to a second child level representing the time period between two deliveries. The stage length at this new level would be one day, and the state variables would be the three estimates $\hat{L}_t, \hat{e}(t), \hat{r}(t)$. The decision concerning the selection of the threshold weight d_t would have to be moved to this level at the stage where the delivery decision is made (a few days before the actual delivery, because the slaughterhouse must be notified in advance). There would, in principle, be a separate process for each state at child level 1, but since only the initial distribution of $\hat{L}_t, \hat{e}(t), \hat{r}(t)$ differs between states having the same number of pigs left, we only need a process for each value of n_t, by use of the "sharing action" facility of the MLHMP software (Kristensen, 2003).

5. Discussion

It is not yet known how the logistics of the monitoring system will be implemented. In the previous sections, it has been assumed that for the pen considered, we have weight estimates. It is probably not realistic to expect that a monitoring system is available in all pens. The most likely solution is that only a few representative pens are monitored so that we have to draw inference on the remaining pens given the observations from those being monitored. More consideration is needed in order to implement a model for a pen not being monitored.

Recent research (Nilsson & Kristensen, 2002) describes how to integrate graphical models like Bayesian networks and influence diagrams into hierarchical Markov decision processes. It should also be considered whether such a combined technique could be of use for a problem like the one presented in this paper.

REFERENCES

Jørgensen, E. 1993. The influence of weighing precision on delivery decisions in slaughter pig production. *Acta Agriculturæ Scandinavica, Section A, Animal Science 43,* 181-189.

Kristensen, A.R. 2003. A general software system for Markov decision processes in herd management applications. *Computers and Electronics in Agriculture.* In press.

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

Kure, H. 1997. *Marketing Management Support in Slaughter Pig Production.* PhD thesis. Department of Animal Science and Animal Health, Royal Veterinary and Agricultural University, Copenhagen. *Dina Research Report No. 57.*

Lien, G., A.R. Kristensen, A. Hegrenes & J.B. Hardaker. 2003. Optimal length of leys in an area with winter damage problems. Grass and Forage Science. In press.

Nielsen, B.K. & A.R. Kristensen. 2002. A model for simultaneous optimization of feeding level and slaughtering policy of organic steers. *First European Workshop on Sequential Decisions under Uncertainty in Agriculture and Natural Resources*. Toulouse, France, September 19-20.

Nilsson, D. & A.R. Kristensen. 2002. Markov LIMID processes for representing and solving renewal problems. *First European Workshop on Sequential Decisions under Uncertainty in Agriculture and Natural Resources*. Toulouse, France, September 19-20.

Søllested, T.A. & A.R. Kristensen. 2002. A sow replacement model using Bayesian updating in a 3-level hierarchic Markov process. *First European Workshop on Sequential Decisions under Uncertainty in Agriculture and Natural Resources*. Toulouse, France, September 19-20.

Toft, N. 2000. *Elements of decision support systems in pig production.* PhD thesis. Department of Animal Science and Animal Health, Royal Veterinary and Agricultural University, Copenhagen. *Dina Research Report No. 90.*

West, M. & J. Harrison. 1997. *Bayesian forecasting and dynamic models.* 2nd edition. Springer-Verlag, New York.