Detecting oestrus by monitoring sows' visits to a boar

Final preprint (accepted version) of article published in *Computers and Electronics in Agriculture.* Please cite as:

Ostersen, T., C. Cornou & A.R. Kristensen. 2010. Detecting oestrus by monitoring sows' visits to a boar. *Computers and Electronics in Agriculture 74*, 51-58.

DOI: 10.1016/j.compag.2010.06.003

Detecting Oestrus by Monitoring Sows' Visits to a Boar

T. Ostersen[1,*], C. Cornou, A. R. Kristensen

University of Copenhagen, Faculty of Life Sciences, Department of Large Animal Science, Grønnegårdsvej 2, DK-1870 Frederiksberg C, Denmark

Abstract

This paper suggests a method for automatic detection of sows returning to oestrus in the gestation department. The detection is based on monitoring of sows' visits to a boar, where the duration and frequency of visits are modelled separately and subsequently combined. The hypothesis is that it is possible to reduce the response time and the number of false alarms compared to previously published attempts. The duration of visits to a boar is defined as seconds per hour the sow is near the boar – logarithmically transformed. The duration is modelled with a multiprocess dynamic linear model with first order Markov probabilities. The indicator of oestrus is the probability of the model describing oestrus, $P(M_{OE})$, and it is monitored with a threshold value. The frequency of visits to a boar is defined as number of visits per 6 hours. A dynamic generalised linear model with two built-in diurnal periods is applied. The indicator of oestrus is the relative deviation from the

*Corresponding author: email: tageostersen@gmail.com

Email addresses: `tageostersen@gmail.com` (T. Ostersen), `cec@life.ku.dk` (C. Cornou), `ark@life.ku.dk` (A. R. Kristensen)

[1]Present address: Danish Agriculture and Food Council, Pig Research Centre, Axeltorv 3, DK-1609 Copenhagen V, Denmark

forecasted frequency, which is monitored with a threshold value. The probability, $P(M_{OE})$, and the relative deviation from the forecasted frequency are combined by means of Bayes Theorem. The combined probability of oestrus is monitored with a threshold value as well. Results indicate that the specificity is superior compared to previously published attempts. The model describing duration alone yields the most satisfactory specificity – 99.4 % per sow day, which is considerably greater than previously published studies. Furthermore, this model detects 87.4 % of the sows entering oestrus, which is slightly lower than previous attempts. The response time of the models is one hour for the duration model and the combined model and six hours for the frequency model. This is better than previous attempts. Even though the specificity is greater, the proportion of false alarms on a day-to-day basis is still too high (91.0 %), which is due to the very large proportion of the sow days defined as non-oestrus. In order to improve the specificity of the detection method, it is suggested to combine the detection method in the present study with other information sources about oestrus.

Keywords: Automatic, Group Housing, Heat Detection, Sow, Oestrus, State space model

---

1. Introduction

Group housing for pregnant sows has become more prevalent in the EU since 2003. Group housing is often more labour intensive (Rasmussen and Duus, 2003), and the labour associated with group housing tends to be perceived as more strenuous (Backus et al., 1997). Part of the labour associated with group housed sows is reproduction management. Usually, a sow is ser-

viced in a separate mating department approximately 5 days after weaning. It is then transferred to the gestation department, where it stays until a few days before expected farrowing, which is 115 days after service. However, some sows that are transferred to the gestation department will return to oestrus either because they did not conceive at first service or because they abort during gestation. In practice, between 5 and 25 % will return to oestrus depending on the efficiency on the individual farm. Detecting those sows in the gestation department is a challenge, because the loose sows are often housed in very big groups. A well optimised reproduction management makes it possible to reduce the averaged number of non-productive days (days, where the sows are neither pregnant nor lactating) by servicing non pregnant sows in the gestation section the first time they re-enter oestrus. Reduced non productive days entail both a better utilisation of the production capacity and reduced feeding costs per produced litter. These factors combined make optimal reproduction management one of the most important means of reducing costs (Korthals, 1999). Currently, reproduction management is performed by daily routines, where the ultimate sign of oestrus is when the sow is susceptible to weight applied on the back (the back pressure test). These daily routines are time consuming and demand a well trained staff. Automation of oestrus detection is one option for improvement of labour conditions and for optimisation of reproduction management of group housed sows. Automated oestrus detection means that the sows are monitored automatically in order to inform the staff of sows entering oestrus. In a review article by Cornou (2006), it is concluded that measurements of the sow's visits to a boar pen show the best results compared to

other automated methods for oestrus detection.

Detection of oestrus by monitoring the sows' visits to a boar is an inexpensive and widely investigated method for automatic oestrus detection (Houwers, 1988; Buré and Houwers, 1989; Bressers et al., 1991, 1995; Korthals, 1999). There are two ways to monitor visits to a boar; one is to have a detection area, which means that there is an area separated from the rest by a passageway. In this detection area, the sow can obtain contact with the boar. By monitoring when the sow passes the passageway it is possible to monitor frequency and duration of the visits (Bressers et al., 1995). The second way of monitoring visits to a boar is a so-called ticket window. This method does not require a separate area for detection, but instead offers the sow a narrow window to obtain contact with the boar (Bressers et al., 1995). Bressers et al. (1995) concluded that there was only little difference in the efficiency of the two methods.

Buré and Houwers (1989) observed an increasing frequency of visits to a boar three days before peak of back pressure test score. The authors observed that the frequency reached a basic level two days after peak of back pressure test score. Bressers et al. (1991) defined a variable containing both frequency and durations of visits per day (Boar Visiting Index - BVI) and compared it to a fixed threshold value. The authors were able to detect 96 % of the oestrus cases and classified 93 % of the sow-days defined as non-oestrus correct (Bressers et al., 1995).

Korthals (1999) improved the above mentioned method by comparing BVI with a fixed value and an exponentially weighted moving average of previous levels of BVI for the individual sow. The author was able to detect

76.3 % of the sow-days defined as oestrus and classify 80 % of the sow-days defined as non-oestrus correctly. Note that the sensitivities of the methods described by Bressers et al. (1991) and Korthals (1999) are not comparable.

Only the approach described by Korthals (1999) considers both coincidental visits and the fact that the activity level of individual sows varies considerably. Another drawback of these methods is that they operate on a day-to-day level, causing the response time of the models to be rather slow. The response time of the model is important in that the sow only is in oestrus for 1 to 3 days. Furthermore, if only 80 % of the sow-days defined as non-oestrus are correctly classified, a normal gestation period of 115 days would result in 23 days with false alarms for a single sow. This indicates that the specificity of the existing methods is too low for use in a gestation section.

A way of obtaining low response time is to use shorter intervals than daily measurements. However, shorter intervals will entail greater fluctuations in the duration and frequency of the visits, creating a need for a model capable of distinguishing random fluctuations from systematic. State space models, as described by West and Harrison (1997), offer numerous filtering approaches.

In the literature, a variety of studies describe the use of automated monitoring systems based on state space models. Examples are given by Madsen et al. (2005), who implemented a dynamic linear model for modelling drinking patterns of young pigs, Cornou and Lundbye-Christensen (2008), who implemented a multiprocess dynamic linear model for modelling activity types from acceleration patterns and Thysen (1993), who implemented a multiprocess dynamic linear model for monitoring somatic cell counts in

dairy production.

The aim of this paper is to implement an alarm system for detecting oestrus in sows in the gestation section by monitoring visits to a boar. The hypothesis is that state space models can reduce the number of false alarms and reduce the response time compared to previously published attempts.

The following section describes the experimental design and the characteristics of the raw data. Sections 3 and 4 describe the model design and parameter values regarding the duration and frequency of visits to a boar, respectively. The two models are combined in Section 5. Section 6 provides evaluation methods, whereas the obtained results are presented and discussed in Sections 7 and 8.

2. Data

All data were collected from the same commercial farm on Zealand, Denmark and data analyses have been performed with the statistical software R (R Development Core Team, 2009).

2.1. Experimental Design

Two distinct data sets were used. The data used for creating the models (learning data) were from a controlled environment and consisted of measurements from 39 sows. The test data were used for testing sensitivity and specificity of the detection methods. A test period for an individual sow is here defined as a period of at least 14 days, where the sow is in the experimental gestation pens. The test data consisted of measurements from 3886

such test periods of a duration of at least 14 days. The measurements were collected in a less controlled environment than the learning data.

2.1.1. Learning Data

Data were collected in three separate experiments that were conducted in 2005 (5 sows), 2007 (12 sows) and 2008 (24 sows), and total at 41 sows. The sows chosen were in their third or fourth parity, had no leg disorders and had reproduction cycles in prior parity of 145 to 147 days. Eight days after weaning oestrus, the sows were introduced to the experimental pen.

The data analysed were from 12-14 to 31-33 days after weaning (i.e. around the expected time of a return to oestrus). All sows were tested positive for weaning oestrus around day 5 after weaning with the back pressure test (Willemse and Boender, 1966), but only 17 of the 41 sows were serviced. The remaining 24 sows were to ensure that some sows entered oestrus during the data collection period. In order to identify oestrus, and thereby establishing a golden standard for when the sows were in oestrus, the back pressure test was conducted three times a day (7 a.m./2 p.m./9 p.m.) from day 21 after weaning. Two sows entered oestrus before or after the period of back pressure testing, which led to misinterpretations. These sows were omitted from further analysis. Thus, 39 sows remained. For a more detailed description see Cornou and Heiskanen (2007).

2.1.2. Test Data

The test data were collected in the period October 2004 to June 2009. There were 3886 test periods (a period of at least 14 days in the gestation section); and 111 cases, where the sows entered oestrus and were serviced in

the gestation section. All test periods were associated to a farrowing date in order to ensure correct date of service. Sows that were serviced during the first three days in the gestation section were omitted. No additional observations were made, which means that the data quality relies on ordinary registrations based on daily observations (e.g. back pressure test) performed by the staff of the farm. Sows included in the learning data were omitted from the test data.

2.2. Housing System and Sensors

All sows were housed in a mechanically ventilated gestation section in pens containing approximately 120 sows. The boar pens were situated at the end of each pen, and contact to the boar could be obtained through a ticket window. The plan of the gestation section is depicted in Fig. 1; only sows in pens 1 and 2 (experimental pens) had access to the boar through a ticket window. The ticket windows and feeding stations were from Skiold A/S (Ikast, Denmark). Each sow carried an RFID-tag (Radio Frequency Identification) in the right ear, which allowed for individual identification when the sow attempted contact to the boar through the ticket window. A visit was recorded whenever the RFID-tag was within 0.2 m of the sensor in the ticket window. For each visit to the boar, the identification of the sow, the starting time and total time were recorded on a central computer. Fig. 2 displays the ticket window, where the sensor is located on the right side of the ticket windows.

[Figure 1 about here.]

[Figure 2 about here.]

2.3. Characteristics of the Raw Data

Because both frequency and duration have been shown to influence the detection efficiency, the characteristics of the raw data are examined in relation to frequency and duration separately.

2.3.1. Duration

Fig. 3 shows the visiting patterns of three different sows. It seems that both the level and variance during normal conditions (i.e. non-oestrus) vary considerably between sows, as the visiting pattern of sow number 9 reveals both a higher level and variance compared to the two other sows. This suggests that the model should be able to adapt to the individual level and variance of each sow during normal conditions.

[Figure 3 about here.]

Based on Fig. 3 (a-b), oestrus seems to be characterised by a shift in level of duration followed by an increase in variance compared to the normal situation. A single outlier, day 6 in Fig. 3 (c), does not necessarily entail oestrus, implying that the model should be able to distinguish single outliers from a level shift. Furthermore, the shift in level around oestrus is more identifiable if the duration is accumulated per hour, whereas accumulation per 6 hours entails a slower response time and too few observations. This enhancement of the level shift around oestrus by accumulating per hour is presumably because the number of visits per hour also increases.

Thus, the most effective response variable seems to be seconds near the boar per hour. Because the multiprocess dynamic linear model applied later

assumes normality, hours where the sow did not visit the boar are removed and duration per hour is transformed on the logarithmic scale.

To summarise: The response variable should be seconds per hour transformed logarithmically, and hours with no observations should be removed. The model should be able to adapt to an individual level and variance, be able to ignore single outliers, and detect oestrus based on a level shift followed by an increase in variance.

2.3.2. Frequency

As seen in Fig. 3, the frequency of visits to a boar is significantly affected when the sow is in oestrus. Moreover, the visits tend to follow a diurnal pattern, which implies that the model should include a diurnal effect. Fig. 4 shows the diurnal distribution of visits for sows in the learning data. It suggests that the diurnal pattern can be divided into two periods, where the high activity period is roughly from 5 a.m. until 5 p.m.

[Figure 4 about here.]

As for the duration of visits, the frequency of visits varies considerably among sows. This suggests a model that can adapt to the individual sow.
The time unit on which the frequency is measured influences the efficiency of the model. If a short time unit is chosen, the frequency will vary only little and often equal zero. On the other hand, a long time unit will cause prolonged response time. As a compromise, a period of 6 hours seems most efficient.
Finally, it is assumed that frequency is Poisson distributed, which entails a need for a generalised model.

To summarise: The response variable should be visits per 6 hours, and each 24-hour-period should be divided into two diurnal periods (daytime and nighttime). The model should be capable of adapting to the individual sow's level, and capable of handling Poisson distributed data.

3. Detecting Oestrus via the Duration of Visits to a Boar

The requirements for the model introduced in Section 2.3.1 indicate that implementing a multiprocess dynamic linear model is an obvious choice, as the model should be able to recognise both level shifts and random outliers.

3.1. Model Design

Dynamic linear models consist of a set of two equations, defined as an observation equation and a system equation, as described by West and Harrison (1997):

$$Y_t = \theta_t + v_t, \; where \; v_t \sim N(0, V), \tag{1}$$

$$\theta_t = \theta_{t-1} + w_t, \; where \; w_t \sim N(0, W). \tag{2}$$

In this case all parameters are scalars and Y_t is the response variable defined as logarithmically transformed seconds near the boar per hour. The observation equation (1) defines the sampling distribution, whereas the system equation (2) defines the time evolution of the underlying mean (θ_t). The error sequences v_t and w_t are assumed mutually independent and normally distributed.

Because the variance differs among sows, the sampling variance, V, is assumed constant but unknown. The evolution variance, W, is determined relatively to the estimated total variance by applying a discount factor, δ, as described by West and Harrison (1997, p. 51).

A multiprocess dynamic linear model consists in general of a number of these sets of equations, but with different variance parameters for each set. The multiprocess dynamic linear model implemented is a class II, which recognises single observations (West and Harrison, 1997); it consists here of four simultaneous models describing the normal situation (N), outliers (OU), level shifts (LS) and oestrus (OE). The sampling variance (V) is only estimated for the normal model, whereas the sampling variances for the other models, V_k, are calculated as

$$V_k = V \cdot c_k, \tag{3}$$

where c_k is the sampling variance factor for the k'th model. A discount factor, δ_k, is provided for each model. Thus, a large value of the discount factor makes the model less adaptive, whereas a small value of the discount factor makes the model very adaptive.

Each model is attached with a prior first order Markov probability, π, of its appropriateness (West and Harrison, 1997, pp. 444-445). This prior influences the final probability of each model. In other words, each observation is assumed to be drawn from one and only one of the four alternative models. That model is referred to as the current model. If the current model is the normal model, M_N, we denote the corresponding probabilities of observing an observation from the normal (M_N), the outlier (M_{OU}), the level shift (M_{LS}) and the oestrus (M_{OE}) models as π_N, π_{OU}, π_{LS} and π_{OE}, respec-

tively. If the current model is an outlier, the next observation is normal with probability 1 (otherwise it would not be an outlier). Similarly, if the current model is a level shift, the next observation is from the oestrus model with probability 1. The full specification of these first order Markov probabilities is given in Fig. 5. Furthermore, if the time distance, measured in hours, between two observations exceeds a time threshold, Δ, or if the observation is among the first six observations for the individual sow, the current model is always assumed to be the normal model.

[Figure 5 about here.]

The four models are defined as:

M_N — Normal model with sampling variance factor $c_N = 1$, discount factor $\delta_N = 0.99$ and prior probability $\pi = \pi_N$

M_{OU} — Outlier model with sampling variance factor $c_{OU} = 20$, discount factor $\delta_{OU} = 0.99$ and prior probability $\pi = \pi_{OU}$

M_{LS} — Level shift model with sampling variance factor $c_{LS} = 1$, discount factor $\delta_{LS} = 0.01$ and prior probability $\pi = \pi_{LS}$

M_{OE} — Oestrus model with sampling variance factor $c_{OE} = 20$, discount factor $\delta_{OE} = 0.99$ and prior probability $\pi = 0$

The values of the sampling variance factors, c_k, and the discount factors, δ_k, are determined by qualitative considerations. Thus, the models can be tolerant or non-tolerant to large fluctuation and adaptive or non-adaptive to new levels.

Fig. 6 provides an example of how the duration is modelled and monitored. At every observation, the posterior probability of each model is calculated via its prior probability, forecast error and forecast variance. This probability is updated as the estimated sampling variance is updated. For further details see West and Harrison (1997, pp. 443-488) and Kristensen et al. (2009). The indicator chosen for detection of oestrus is the posterior probability of the model describing oestrus, $P(M_{OE})$.

[Figure 6 about here.]

3.2. Estimation of Model Parameters

Before any observation is performed for the individual sow, θ_0 is assumed normally distributed with initial mean m_0 and variance C_0. The value of m_0 is set to the mean level per sow not in oestrus, whereas C_0 is set to four times the variance between sows not in oestrus to ensure a vague prior estimate. Each time a new observation is done at time t, the distribution for θ_t is updated as described by West and Harrison (1997, pp. 56-57) so that based on the new observation combined with the previous mean m_{t-1} and variance C_{t-1}, a new mean, m_t, and a new variance, C_t, are calculated by the updating equations.

The prior probabilities $(\pi_N, \pi_{OU}, \pi_{LS})$ and a prior point estimate, S_0, of the variance V, together with the time threshold (Δ) are optimised by maximising the sensitivity and specificity per 24 hours. The sensitivity and specificity are defined as:

$$sensitivity = \frac{True\,Positive}{True\,Positive + False\,Negative}, \tag{4}$$

$$specificity = \frac{True\,Negative}{True\,Negative + False\,Positve}. \quad (5)$$

The optimisation was based on the learning data. Each sow is defined as in oestrus from 72 hours before oestrus onset until 48 hours after. This rather prolonged period where the sow is defined as in oestrus is to ensure that the model is "rewarded" for an early reaction on oestrus; however, it entails that the apparent sensitivity becomes low. A true positive is when at least one alarm occurs during 24 hours and the sow is defined as in oestrus; a true negative is when no alarms occur during 24 hours and the sow is defined as not in oestrus. This means that the sensitivity and specificity are measured per 24 hours similarly to the calculations performed by Korthals (1999). The indicator of oestrus, $P(M_{OE})$, is monitored with a threshold value of 0.8. In other words, if the probability exceeds the level of 0.8, an alarm is generated.

The principle applied in the optimisation is to maximise the sensitivity per 24 hours given that the specificity assumes the critical value of 99.5 %. In practice S_0 was varied in the range [1,1.4] by steps of 0.2. The value of π_N was varied in the range [0.25,0.55] by steps of 0.10. π_{OU} was calculated as a proportion of the distance between 1 and π_N, which was varied in the range [0.25,0.55] by steps of 0.10. π_{LS} was calculated as $\pi_{LS} = 1 - \pi_N - \pi_{OU}$. Δ was varied in the range [6,30] by steps of 6. The result of the optimisation is reported in Table 1.

[Table 1 about here.]

The low sensitivity of 50 % does not indicate that only 50 % of the sows are detected, but that 50 % of the days defined as oestrus are detected. In fact, only 1 of the 23 sows entering oestrus is not detected – indicating a

sensitivity of 96 % per sow.

4. Detecting Oestrus via the Frequency of Visits to a Boar

In Section 2.3.2, it is argued that the model describing frequency should be adaptive and capable of handling Poisson distributed data, which suggests a dynamic generalised linear model (West and Harrison, 1997). Furthermore, a dynamic generalised linear model permits a number of levels, which meets the need of modelling diurnal patterns. The response variable, Y_t, is the number of visits per 6 hours.

4.1. Model Design

The dynamic generalised linear model handles all distributions belonging to the exponential family (i.e. binomial, normal, Poisson and gamma distributions). When the data are Poisson distributed, the observation equation can be formulated as Eq. (6). Eq. (6) defines the probability function, and the impact of the underlying parameter vector θ_t on the natural parameter η_t (Kristensen et al., 2009).

$$\begin{aligned} (Y_t|\eta_t, V_t) &\sim \mathcal{P}(\lambda) = \mathcal{P}(e^{\eta_t}), && (6) \\ where\, \eta_t &= F_t'\theta_t, \\ \theta_t &= G\theta_{t-1} + w_t,\ where\ w_t \sim (0, W). && (7) \end{aligned}$$

The system variance, W, is assumed known and constant. The regression vector, F_t, describes the diurnal pattern by assuming different values for the two diurnal periods defined in Section 2.3.2 as high and low activity periods.

$$F^{high} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ F^{low} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},\ G = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (8)$$

This dynamic generalised linear model can estimate a 1-step forecast of the mean, allowing for a comparison with the actual observed number of visits. The chosen indicator of oestrus (9) is the relative deviation from the expected frequency. The indicator is monitored with a threshold value.

$$oestrus\ indicator = \frac{observed\ freq - forecasted\ freq}{forecasted\ freq + 1}. \tag{9}$$

Fig. 7 exemplifies how oestrus is detected based on frequency. Fig. 7 (a) depicts the observed and forecasted frequency, whereas Fig. 7 (b) shows the development of the oestrus indicator.

[Figure 7 about here.]

4.2. Estimation of Model Parameters

The initial values of the underlying mean vector (m_0) and the sampling variance matrix (C_0), together with the value of system variance matrix (W) were optimised by minimising the squared forecast errors of the learning data during normal conditions (i.e. until 72 hours before oestrus). The values of m_0 were varied in the range of [0.1,1.6] by steps of 0.5 and the diagonal values of C_0 were varied in the range [0.01,0.81] by steps of 0.4. The diagonal values of the system variance matrix were varied in the range of [0.001,0.041] by steps of 0.02. The optimised values are as follows:

$$m_0 = \begin{pmatrix} 1.1 \\ 0.1 \end{pmatrix}, \ C_0 = \begin{pmatrix} 0.01 & 0 \\ 0 & 0.01 \end{pmatrix}, \tag{10}$$

$$W = \begin{pmatrix} 0.041 & 0 \\ 0 & 0.021 \end{pmatrix}. \tag{11}$$

Note that C_0 assumed the lowest possible value in the optimisation, indicating that the optimal value could have been lower. However, this is presumably of less importance as a large value of C_0 makes the prior, m_0, more vague, and thus makes the model more adaptive to the individual sow.

The threshold value can assume any arbitrary positive value. However, the value affects the sensitivity and specificity of the detection. The threshold value was chosen by maximising the sensitivity given that the specificity assumed the critical value of 99.5 %. A threshold of 3.3 yielded a sensitivity of 35.7 % and a specificity of 99.6 %, measured per 24 hours.

Despite a low sensitivity, 22 of the 23 sows in oestrus were detected between 48 hours after and 72 hours before onset of oestrus.

5. Alarm System Based on Frequency and Duration

To create an expression that includes both the knowledge about the duration and the frequency of the visits to a boar, a probabilistic approach based on Bayes Theorem seems straightforward. Because the duration yields a probability of oestrus, this can be utilised as a prior (unconditional probability - $P(oestrus)$). The test result from the frequency model can then, via the sensitivity and specificity of the test, affect this probability. The combined probability of oestrus given the result from the frequency model is positive, $P(oestrus|+)$, or negative, $P(oestrus|-)$, are calculated as:

$$P(oestrus|+) = \frac{P(+|oestrus) \cdot P(oestrus)}{P(+)}, \tag{12}$$

$$P(oestrus|-) = \frac{P(-|oestrus) \cdot P(oestrus)}{P(-)}. \tag{13}$$

The two probabilities, $P(+)$ and $P(-)$, simply act as normalising constants that express the probability of a random positive or negative outcome of the frequency test. These probabilities are dependant on the prior probability of oestrus. $P(+|oestrus)$ is the probability of a positive test outcome from the frequency method if the sow is in oestrus (sensitivity), whereas $P(-|oestrus)$ is 1 - sensitivity of the frequency method.

The combined probability is calculated every hour the sow has visited the boar. If there has been an alarm regarding frequency within 24 hours, Eq. (12) is applied. If no alarms regarding frequency have occurred within 24 hours, Eq. (13) is applied. The combined probability of oestrus is monitored with a threshold value of 0.95.

6. Evaluation Methods

In order to evaluate the efficiency of the models, both overall and in comparison, some evaluation methods must be supplied. The sensitivity and specificity are the classical evaluation methods and are calculated according to Eq. (4-5). The error rate expresses the proportion of alarms being false and is, in contrast to sensitivity and specificity, dependant on the number of positives and negatives. It is calculated as:

$$error\,rate = \frac{False\,Positive}{False\,Positive + True\,Positive}. \tag{14}$$

In order to have another instrument for comparison between the models, the likelihood ratio (LR) is calculated for a positive and negative test result (15-16), according to Woodward (1999).

$$LR^{+} = \frac{sensitivity}{1 - specificity}, \tag{15}$$

$$LR^{-} = \frac{1 - sensitivity}{specificity}. \tag{16}$$

If LR^{+} equals 1, the reliability of positive test is similar to tossing a coin, whereas if $LR^{+} = 2$, the reliability of a positive test is twice as big as tossing a coin. Furthermore, if LR^{-} equals 1, the reliability of a negative test result is similar to tossing a coin, whereas if $LR^{-} = 0.5$, the reliability of a negative test is twice as big as tossing a coin. These expressions do not weigh according to prevalence of positives and negatives. However, this does not affect model comparison as the prevalence of positives and negatives is the same for the three detection methods.

The three detection methods are evaluated according to blocks of 24 and 72 hours. A true positive is when at least one alarm occurs during the block and the sow is defined as in oestrus. The block of 72 hours is designed with date of oestrus as point of reference, in order to obtain only one true positive per oestrus case. In this manner, the sensitivity measured according to the block of 72 hours expresses the sensitivity per sow.

7. Results

The three detection methods (duration, frequency and combined) were tested on the test data, which included 318,267 sow days, where 331 days are defined as in oestrus (from two days before first day of service until the first day of service in the test data). The model detecting oestrus based on

duration yields an alarm whenever the posterior probability exceeds 0.8. The model describing frequency yields an alarm whenever the oestrus indicator (9) exceeds the value of 3.3. The combined probability of oestrus yields an alarm whenever the probability exceeds 0.95.

Table 2 reports the performance of the three methods in blocks of 24 hours, whereas Table 3 reports the results in blocks of 72 hours.

[Table 2 about here.]

[Table 3 about here.]

Results indicate that the model describing duration of visits has both a greater sensitivity and specificity than the model describing frequency of visits. This applies to both the block of 24 hours and the block of 72 hours. The sensitivity of the model describing duration is 55.6 %, measured in blocks of 24 hours, and 87.4 % measured in blocks of 72 hours. The specificity is 99.4 %, measured in blocks of 24 hours.

The result of the combined probability of oestrus indicates a more satisfactory sensitivity than both the models describing duration and frequency alone. However, the specificity of the combined model is lower than the model describing duration alone. This is reflected in the values of LR^+, where a positive test result is more reliable for the model describing duration than the combined model. In contrast, LR^- reveals that a negative test result is more reliable for the combined model than for the model describing duration alone.

8. Discussion

A method for detecting oestrus by means of monitoring visits to a boar is implemented. The duration of visits is modelled with a multiprocess dynamic linear model – yielding a probability of oestrus at a given observation time. The frequency is modelled with a dynamic generalised linear model, where the relative deviation is monitored via a threshold. The results of modelling the duration and frequency are combined by means of Bayes Theorem.

The results of the three models clearly show that the model based on frequency is less efficient in detecting oestrus than the two other models. It is difficult to elucidate which of the two other models is the most efficient. The model describing duration alone yields a more reliable test result (measured by LR^+) when the result is positive, due to its greater specificity. However, this model has a lower sensitivity, which means the reliability of a negative test result is lower (LR^-). Because of the small proportion of sow days defined as oestrus, and because the model describing duration alone is more simple and has greater specificity, this model is presumably preferable.

Results indicate that all three models applied are superior regarding both specificity and response time, but less satisfactory regarding sensitivity, compared to previously published attempts. Bressers et al. (1995) achieved a sensitivity of 96 % measured per sow, but only a specificity of 93 % per sow day. This was, however, only tested on 74 sows. Korthals (1999) achieved a sensitivity of 76 % and a specificity of 80 % measured per sow day. If the relation between the number of days defined as positive and negative was the same as in the present study, the error rate would be 99.6 % – i.e. only one out of 250 alarms would be correct.

Because of the conflicting relation between sensitivity and specificity, it is difficult to determine whether the models applied in this study are preferred over previously published studies. Previous studies were performed in a mating/gestation section, which means that the prevalence of positive sow days was significantly higher than in the present experiment, implying that the specificity was of less importance in these studies. However, in the present study, specificity is of utmost importance, as the prevalence of positive sow days is very low. Since the presented method achieves much larger specificities than previously published methods, it is concluded that the presented detection method is better, when applied in a gestation section.

The issue of how to determine an overall efficiency when comparing the models - i.e. sensitivity versus specificity could be resolved by applying cost of not detecting an oestrus case in time versus the cost of a false alarm. The cost of not detecting an oestrus case is made up of feeding costs and cost of space in the stable. A false alarm will cost the time of the worker to find the sow in the group and test whether it is in oestrus. However, except the feeding costs, these economic values are difficult to determine and are dependant on the individual farm. Furthermore, the cost of a false alarm is presumably non-linear, because at some point the cost exceeds the cost of simply monitoring all sows on a daily basis. Based on this, it seems reasonable that the optimisation was performed on the sensitivity and specificity without considering cost.

The unsatisfactory outcome of combining the two models is somewhat surprising, as previously published studies show that both frequency and

duration are important when monitoring sows' visits to a boar (Bressers et al., 1991; Korthals, 1999). However, the findings are possibly linked to the fact that the model describing duration includes the time distance between the last two observation hours, which is very closely linked to frequency. Furthermore, by accumulating duration per hour, some of the frequency is included in the response variable. Thus, some information about frequency is already embodied in the duration model, implying that the model describing frequency contributes with less new information. A different approach could be to extend the method applied to monitor the duration of visits. The idea would be to utilise both the information about frequency and duration in one model. For instance, the time distance between the last two observations could influence the prior probability of a level shift.

The definition of a positive sow day in the test data (from two days before service until the day of service) and a negative sow day is questionable. In a few cases, the data indicate an oestrus period exactly three weeks prior to the time of service, which indicates that the sow was in oestrus, but was undetected by the staff of the farm. Alarms during this period are defined as false positives, even though the sow might actually be in oestrus. The same problem occurs if a sow is still in oestrus one day after the first day of service. Definition of positives and negatives should be taken into consideration, since the apparent efficiency of the detection model can appear lower. However, with the data available in the present study this problem is difficult to overcome.

It seems that the greatest challenge in detecting oestrus in the gestation

section is to reduce the error rate. For the duration model, the error rate is 92 % despite the very large specificity of 99.4 %. To achieve an error rate of 50 % measured per sow day for the model describing duration, the specificity should be 99.94 % if the sensitivity is unaltered. The challenge of improving the specificity to such extent is presumably difficult if the detection is purely based upon visits to a boar. A way of obtaining greater specificity and sensitivity is to utilise all sources of information about the sow. There are several information sources worth considering.

The order in which the sows enter an electronic sow feeder (eating rank) has been applied for oestrus detection with varying success (Bressers et al., 1993; Søllested, 2001; Cornou et al., 2008). Bressers et al. (1993) and Søllested (2001) concluded that the eating rank was not sufficiently stable to be used as indicator of oestrus. However, Cornou et al. (2008) were able to achieve sensitivities between 59 % and 75 % measured per sow and specificities between 81 % and 95 % measured per sow day. Other automatic methods could include body activity measurements as mentioned by Cornou (2006). This is, however, less straightforward on the short run as the required sensor technology for this purpose is not yet available on commercial farms.

Relevant information sources could also be non-automated. An example could be the pregnancy test, which often is performed after day 21 after service. The efficiency of the pregnancy test is rather well documented, and although it only indirectly affects the probability of oestrus it seems straightforward to include this information, as a truly pregnant sow is unlikely to enter oestrus.

Knowledge of the farrowing rate on the individual farm at the specific

time also seems like a vital information, as well as the proportion of sows entering oestrus around day 21.

The differing character of these information sources makes combination of them challenging. However, a Bayesian Network (see Jensen and Nielsen (2007) for an introduction) could potentially meet these challenges.

9. Conclusion

The combined model presented is not unambiguously more efficient than the duration model alone, which probably is due to the duration model containing some information about frequency via the time distance between observations. Thus, the model describing duration alone yields the most satisfactory specificity – 99.4 % per sow day. This is considerably greater than previously published studies that achieved specificities of 80 and 93 %. Furthermore, the model detects 87.4 % of the sows entering oestrus, which is lower than previously published studies. In other words, a conclusion regarding whether the efficiency has been improved is difficult to present. However, if the purpose is to detect oestrus in a gestation section, the presented method is more suitable than previously published methods. The response time of all three models is better than previous attempts (1 or 6 hours as opposed to 1 day).

Even though the specificity is considerably greater than earlier attempts, the proportion of false alarms on a day-to-day basis is still too high (91.0 %), which is due to the very large proportion of the sow days defined as non-oestrus. If only 50 % of the alarms are permitted to be false, the specificity should be 99.94 % per sow day, given that the sensitivity is unaltered.

In order to improve the specificity of the detection model, it is suggested to combine the detection method in the present study with other information sources regarding oestrus. A Bayesian Network is an option for combining such information sources.

Acknowledgements

The authors gratefully acknowledge the employees of the production herd, Askelygaard, and Danish Agriculture and Food Council, Pig Research Centre for supplying the data.

References

Backus, G. B. C., Vermeer, H. M., Roelofs, P. F. M. M., Vesseur, P. C., Adams, J. H. A. N., Binnendijk, G. P., Smeets, J. J. J., Peet-Schwering, C. M. C., van der Wilt, F. J., 1997. Comparison of four housing systems for non-lactating sows. Proefverslag - Proefstation voor de Varkenshouderij.

Bressers, H. P. M., te Brake, J. H. A., Engel, B., Noordhuizen, J., 1993. Feeding order of sows at an individual electronic feed station in a dynamic group-housing system. Appl. Anim. Behav. Sci. 36, 123–134.

Bressers, H. P. M., te Brake, J. H. A., Noordhuizen, J. P. T. M., 1991. Oestrus detection in group-housed sows by analysis of data on visits to the boar. Appl. Anim. Behav. Sci. 31, 183–193.

Bressers, H. P. M., te Brake, J. H. A., Noordhuizen, J. P. T. M., 1995. Automated oestrus detection in group housed sows by recording visits to the boar. Livest. Prod. Sci. 41, 183–191.

Buré, R. G., Houwers, H. W. J., 1989. Automation in group housing of sows. In: Agricultural engineering: proceedings of the 11.International Congress on Agricultural Engineering. pp. 827–830, 1 - Land and Water Use.

Cornou, C., Heiskanen, T., 2007. Automated oestrus detection method for group housed sows using acceleration measurements. In: Precision Livestock Farming '07. pp. 211–217.

Cornou, C., Lundbye-Christensen, S., 2008. Classifying sows´ activity types from acceleration patterns. an application of the multi-process kalman filter. Appl. Anim. Behav. Sci. 111, 262–273.

Cornou, C., Vinther, J., Kristensen, A. R., 2008. Automatic detection of oestrus and health disorders using data from electronic sow feeders. Livest. Sci. 118, 262–271.

Cornou, C., 2006. Automated oestrus detection methods in group housed sows: Review of the current methods and perspectives for development. Livest. Sci. 105, 1–11.

Houwers, H. W. J., 1988. Localty registration as a way of heat detection in an integrated group housing system for sows. In: International Congress of Applied Ethology in Farm Animals. pp. 44–50.

Jensen, F. V., Nielsen, T. D., 2007. Bayesian networks and decision graphs, 2nd Edition. Springer, New York, USA.

Korthals, R. L., 1999. The effectiveness of using electronic identification for the identification of estrus in swine. In: 1999 ASAE/CSAE-CGCR Annual Meeting.

Kristensen, A. R., Jørgensen, E., Toft, N., 2009. Herd Management Science II. Advanced topics. SL books, Copenhagen, Denmark.

Madsen, T. N., Andersen, S., Kristensen, A. R., 2005. Modelling the drinking patterns of young pigs using a state space model. Comput. Electron. Agric. 48, 39–62.

Rasmussen, J., Duus, L. K., 2003. Erfaringer med tidsforbrug i drægtighedsstalde. Tech. rep., Erfaring nr. 0301. Dansk Svineproduktion.

R Development Core Team, 2009. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.
URL `http://www.R-project.org`

Søllested, T., 2001. Automatic oestrus detection by modelling eating behaviour of group housed sows in electronic sow feeding systems. Master's thesis, Department of Animal Science and Animal Health, Royal Veterinary and Agricultural University. Copenhagen.

Thysen, I., 1993. Monitoring bulk tank somatic cell counts by a multi-process kalman filter. Anim. Sci. 43, 58–64.

West, M., Harrison, J., 1997. Bayesian Forecasting and Dynamic Models, 2nd Edition. Springer, New York, USA.

Willemse, A. H., Boender, J., 1966. A quantitative and qualitative analysis of oestrus in gilts. Tijdschr Diergeneesk 91, 349–363.

Woodward, M., 1999. Epidemiology: study design and data analysis. CRC Press, Florida, USA.

List of Figures

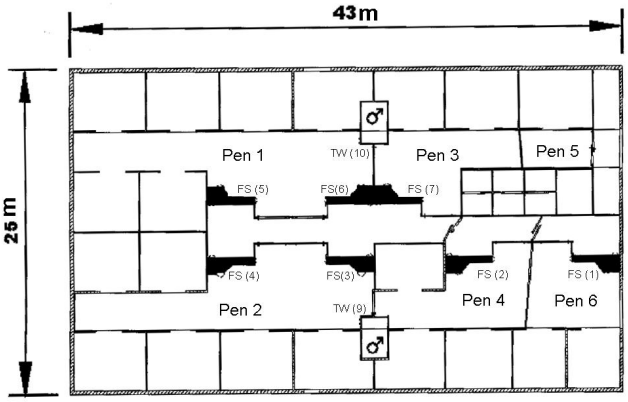

Figure 1: Layout of the gestation section. Only data from pens 1 and 2 were used. Note the location of the feeding stations (FS) and ticket windows (TW).