A Model for Monitoring the Condition of Young Pigs by their Drinking Behaviour

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

Madsen, T.N. & A.R. Kristensen. 2005. A Model for Monitoring the Condition of Young Pigs by their Drinking Behaviour. *Computers and Electronics in Agriculture 48*, 138-154.

DOI: 10.1016/j.compag.2005.02.014

A Model for Monitoring the Condition of Young Pigs by their Drinking Behaviour

Thomas Nejsum Madsen [a,b,*], Anders Ringgaard Kristensen [b]

[a]*Department of Large Animal Sciences, The Royal Veterinary and Agricultural University, Grønnegårdsvej 2, DK-1870 Frederiksberg C, Copenhagen, Denmark*

[b]*The National Committee for Pig Production, Danish Bacon & Meat Council, Axeltorv 3, DK-1609 Copenhagen V, Denmark*

Abstract

A uniform water or feed consumption pattern is required for assessment of changes in pig health and well-being. Water consumption rate was measured and monitored continuously in pigs from 4 to 11 weeks of age. The study comprised three herds and 18 batches of pigs each with 400-900 animals. The management system was "all-in all-out". Water consumption was measured electronically and data were transferred to a computer for time series analysis. Management interferences such as change of diet and treatment of pigs were recorded daily in a log book. The results indicated that water consumption was associated with a distinct circadian rhythm. Water consumption rate peaked between 4 and 6 pm and was lowest between 3 and 5 am regardless of herd and housing system. The circadian pattern persisted throughout the growing period while total water consumption rate increased.

The pigs showed a very stable diurnal drinking pattern as long as they were healthy whereas the pattern often changed when the pigs were affected by a disease.

A method using a state-space model in conjunction with a Cusum control chart is presented as a tool for on-line monitoring of young pigs, based on the water consumption. By an example it is shown that an outbreak of a disease (diarrhea) can be detected by the monitoring method approximately one day before physical signs are seen on the pigs.

Key words: Dynamic linear model, cusum control chart, drinking behaviour, diurnal pattern

---

* Corresponding author. Present address: TNM Consult Aps, Gydevang 39-41, DK-3950 Allerød, Denmark

Email addresses: `tnm@tnmconsult.dk` (Thomas Nejsum Madsen), `ark@dina.kvl.dk` (Anders Ringgaard Kristensen).

1 Introduction

Management decisions while taking care of pigs are frequently based on subjective judgments. Normally, such decisions derive from a combination of information sources including visual observations of pigs (e.g. aggression) and pens (e.g. signs of diarrhea), or other senses (e.g. temperature) as well as results from monthly or quarterly performance records.

Outbreak of a disease such as diarrhea often spreads fast within a group of pigs, and if it is not detected and treated immediately the outcome might be losses in terms of reduced growth rate and increased mortality. Similarly, poor feed quality might result in lower growth rate and poor feed conversion.

Because of the general trend towards larger herds and more animals managed per person there is little time available for observing individual animals in weaning and finishing units. Formerly, when daily caretaking was associated with manual feeding and mucking out more time was spent among the animals, which increased the chance of detecting outbreak of disease or other problems.

Modern pig housing includes fully slatted flooring, automatic climate control, and automatic feeding and watering, which allow the caretaker to manage a large number of animals. Theoretically, labour saving equipment should increase time available for inspection and supervision of animals. However, in commercial practice improved technology has meant that time spent on manual labour has been converted to more animals managed per person. According to data from the Danish Agricultural Advisory Service, the average labour input being only 10-12 min per finishing pig produced. Therefore, it is important that time available for inspection is devoted to those animals that are at risk of being infected with disease or exposed to stressors. The problem is to determine the risk level or health status of various groups of pigs. Of course, the manager might have some prior knowledge, i.e. for instance, that certain disease outbreaks appear at a given stage of the production period. However, it must be assumed that some kind of guidance might be helpful as to where to concentrate management efforts. Thus, it has been suggested that automatic monitoring systems and use of time series analysis might be promising management tools (Frost et al., 1997; Bird and Crabtree, 2000; Bird et al., 2001). Yet, surprisingly few attempts have been made in terms of using such methods in modern pig production.

There is general agreement that animal well-being might be measured indirectly using indicators such as animal behaviour (Smidt, 1983). Changes in eating and drinking patterns are usually the first visual signs that pigs are experiencing environmental stress. Eating patterns in pigs have been reported in several studies (Slader and Gregory, 1988; Nienaber et al., 1991; Xin and DeShazer, 1992; Young and Lawrence, 1994; Hyun et al., 1997). A common finding was that the eating pattern is characterized by a distinct diurnal rhythm. Bigelow and Houpt (1988) have shown that pigs' drinking behaviour is closely correlated to their feed consumption which leads to the hypothesis that changes in pig health that affect feed intake also impinge on water intake and thus drinking behaviour. This leads to the idea of using statistical quality control for monitoring the pigs' water consumption in order to detect when they change their behaviour.

Statistical quality control methods are quite commonly used in the manufacturing industry (see e.g. Montgomery (1996)), but more rarely in animal husbandry. However, a few examples exist, e.g. monitoring daily milk yields in dairy cows (Van Bebber et al., 1999), detection of oestrus and disease in dairy cattle based on time series analysis (de Mol et al., 1999); detection of changes in feed consumption in broilers (Roush et al., 1992) and detection of changes in daily milk production in cows (Thysen, 1992). Common for all these monitoring systems are that they: 1) incorporate an automatic method for collection of production traits, and 2) include an adequate model for analysis of the collected data.

Modern computer technology has extended the possibilities of real-time monitoring at the farm level. Process computers, known from factory automation, used together with electronic water flow-meters can easily be set up at farm level to provide real-time data of water consumption on a PC. The real challenge lies in processing the recorded data in order to achieve as much information as possible. Monitoring production traits is often difficult because of random as well as more systematic variation in responses, which may complicate the interpretation of data.

The basic problem is to detect change-points in a sequence of discrete sums of water consumption. There is much literature on change-point problems, see e.g. Christensen and Rudemo (1996) and Csörgö and Horváth (1988) for discussion of different methods. In our case the change-point problem is complicated by the dynamic aspects of data. The time series exhibit varying diurnal patterns as well as unpredictable growth rates in the overall level and there is very little, if any, prior information at the beginning of a time series.

The objective of this paper is to develop a method for monitoring the condition of young pigs by measuring their water intake at a real-time basis. The method combines Bayesian modelling with a traditional statistical quality control model, namely the Cusum control chart. A similar approach is seen in Iwersen (1997). The method is intended to be used as part of a computer-based monitoring system [1] for pig farms.

[1] The Farm*Watch*™ system.

Table 1
Production conditions were almost identical on farm A and farm B, with relatively small pens and slatted floors, whereas farm C had much larger pens with deep litter. There were twice as many pigs per drinking bowl on farm C as compared to farms A and B.

	Farm A	Farm B	Farm C
Pens per section	12	24	4
Pigs per pen	35	30	250
Number of sections in study	2	1	1
Flooring	Partly slatted	Partly slatted	Deep litter
Feeding system	Tube feeder	Tube feeder	Tube feeder
Feeding regimen	Ad lib	Ad lib	Ad lib
Pigs per drinking bowl	17	15	30

2 Materials

Water consumption was measured in 18 batches of 4-wk- to 11-wk-old pigs on three commercial farms. One of the farms participated with 2 sections whereas the others participated with one section each. Each batch of pigs originated from the same weaning. Pigs were housed in rooms comprising 400-900 animals, which were managed all in all out. On all three farms, pigs were fed a standard cereal-soy diet according to Danish feeding standards, and in all cased, water was provided by automatic drinking bowls. Production conditions are shown in Table 1.

Electronic water flow-meters were installed in four weaning sections for measurements of water disappearance, as illustrated in Figure 1.

Water consumption was recorded at 2 min intervals as whole litres consumed. All data were logged every two minutes and stored in a local process computer. Every 24 hours data were transmitted to a central database by means of a modem link. In addition, the farm staff recorded any action such as treatment of sick animals, change of feed, removal of dead animals, etc., in a log book.

Electronic recording and log book data were arranged batchwise. A single data set contained about 36000 observations and 50 log book pages. Due to technical problems only 12 time series out of 18 batches were complete.

Water consumption in a batch of 300-900 pigs can be regarded as a continuous process, but for practical reasons the flow has been measured in discrete intervals.

An example of the 2 minute sums of water consumption for a batch of 350 pigs 29 days after weaning is shown in Figure 2. The figure reflects the fact that the water consumption is measured as the number of whole litres consumed in a period of

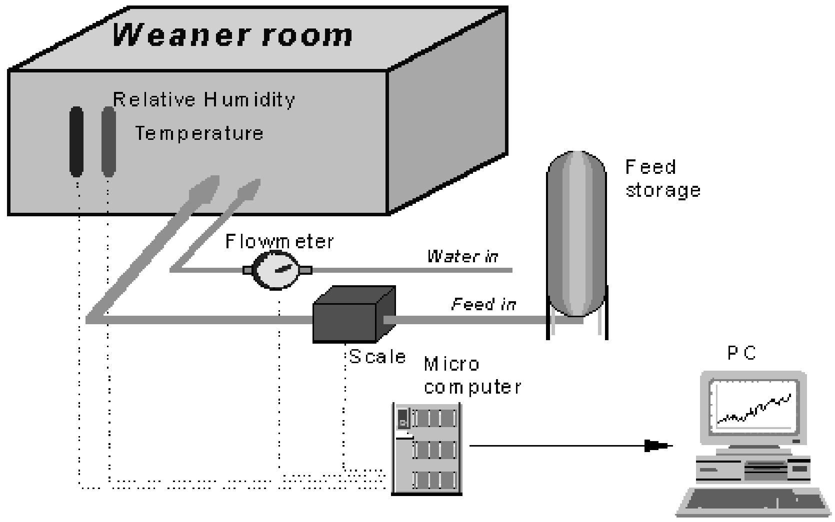


Fig. 1. Water and feed disappearance, temperature and relative humidity were monitored electronically. Data were transferred to a computer database for processing. In this study only data on water consumption was used.

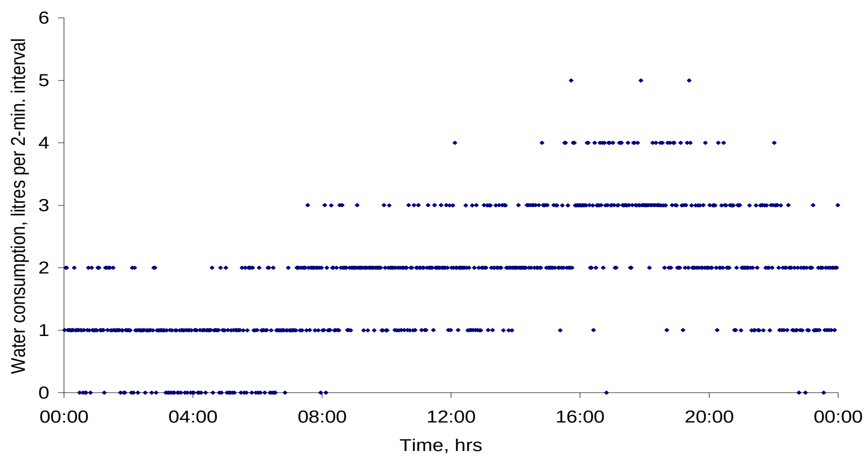
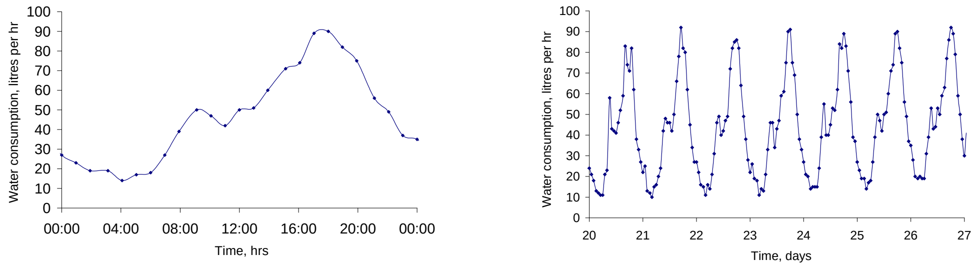


Fig. 2. An example of 2-minute observations of the water consumption for a group of 350 pigs, 29 days after weaning.

2 minutes. The density of the points in the plot indicates a non-constant drinking behaviour during the day.

The primary aim of the data recording was to investigate whether the pigs' drinking pattern can be used as an indicator of their health condition. The first requirement is that data somehow repeat themselves, that is, we should be able to formulate a



(a) 1 hour sums, 1 day

(b) 1 hour sums, 7 days

Fig. 3. The 2-minute observations, shown in Figure 2, are aggregated to 60-minutes sums.

well-founded expectation for each observation. From Figure 2 it is obvious that the 2 minute sums contain a lot of random noise. To diminish the effect of the random noise, and to make data easier to handle, the 2 minute registrations have been aggregated. Aggregation implies loss of information so it is desirable not to aggregate data to much. An evaluation of 20 minutes, 1 h and 4 h summation intervals revealed that a 1 h interval provided the best compromise between being able to handle data in a model and not losing too much information. The 1-hour sums of water consumption show a much more stable pattern (Figure 3). The drinking patterns of the young pigs were almost identical in all four sections (and thus also all three farms), despite the fact that there were considerable differences in the housing systems as shown in Table 1. The recorded data can be characterized by the following properties (Madsen et al., 2005):

- The level of water consumption increases as the pigs grow.
- The daily drinking pattern is stable in a normal situation.
- Data contain some random noise because of biological variation and measurement errors.

A monitoring method designed to fit the characteristics of the water consumption data will be presented in the following section.

3 Method

In the 1920s, Walter Shewhart developed a number of business production analysis techniques, which were designed to detect changes in the quality of the output from continuous production processes (see Montgomery, 1996, Chapter 4). Statistical Process Control (SPC), or quality control, is widely used in many industries to facilitate objective evaluation of business operations and production processes. SPC is used to monitor the level of a production trait, and to give a notification when the level changes beyond some predefined limit. The basic control charts are designed to monitor a process that is expected to be constant, although it allows for some random fluctuation. Monitoring the water intake of growing pigs is somehow more complicated since the expected level of the process is not at all constant.

Consider the water intake as a stochastic observable process that we want to monitor. We do have some idea of how the observable process will evolve as the water intake is basically an effect of the latent physiological processes in the pigs. As described in Section 2, the dynamic nature of the recorded data on water consumption does not conform to the "constant process" assumptions on which the control charts are based. Therefore the time series of water consumption needs to be transformed somehow to fit into the SPC framework. Another problem is that the observed time series contains some random noise, as described in Section 2.

Data are modeled by means of a Dynamic Linear Model (DLM), which is well suited to model the dynamic/cyclic evolution in data, described in Madsen et al. (2005). The model is defined by superposition of a linear growth model and a cyclic model with period 24 (24 hourly sums), where the cyclic model itself is also constructed by several sub models. By using this principle, it is possible to split the fitted model up into a level component and a cyclic component, as it will be shown in Section 4.3.

Usually, the DLM is used as a tool for making forecasts, based on prior knowledge including former observations. In this framework, the DLM is used to make a prediction of the water consumption one step ahead in time as described in detail by Madsen et al. (2005). The difference between the one step forecast at time $t - 1$ and the observation Y_t is then used as a measure of the deviation from the "normal" level of water consumption. The deviation or forecast error can be considered as an independent random error term with zero mean as long as the process model is valid. If, on the other hand, the pigs change their drinking behaviour, data will no longer conform to the model predictions, and the numerical value of the forecast errors will increase. For details about how to calculate the forecast, reference is made to West and Harrison (1997, pp 103-104).

For practical purposes, one has to distinguish between deviations from the normal drinking behaviour caused by a change in the growth rate of the pigs, and deviations caused by a disease that implies increased/decreased water consumption. The latter is illustrated in figure 4, which shows the sum of standardized forecast errors from the DLM, based on a batch of weaners having had a serious outbreak of diarrhea within the first 20 days after weaning. The outbreak was detected on day 11 when all the pigs were treated with antibiotics for a period of 5 days. It clearly appears from the figure that the sum of forecast errors increases radically 1-2 days before the disease is detected, indicating a systematic deviation between data and model. This pattern has been found in almost all of the time series where outbreak of diarrhea has been reported.

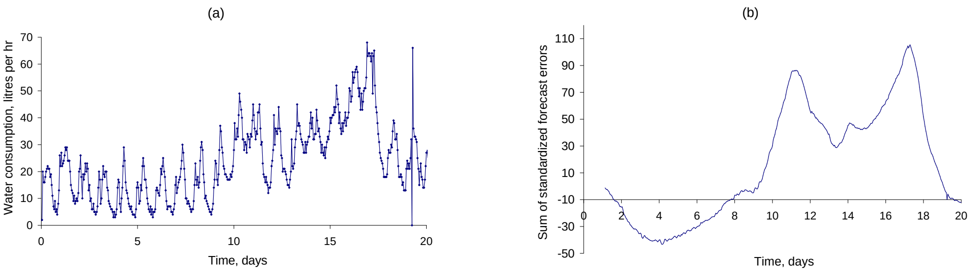


Fig. 4. (a) A sequence of one-hour sums of water consumption for a group of weaners. Outbreak of diarrhea is detected by the caretaker at the end of day 11, but increasing water intake was evident already on day 9. (b) Sum of standardized forecast errors from the DLM.

The monitoring concept described below is based on a DLM used in conjunction with a cumulative-sum or Cusum control chart. The idea is to make a forecast at each time t of the level of the next observation at time $t+1$, based on all the former observations $Y_0, \ldots, Y_t$. Then the Cusum control chart is applied to detect when the time series of forecast errors deviates from the zero level.

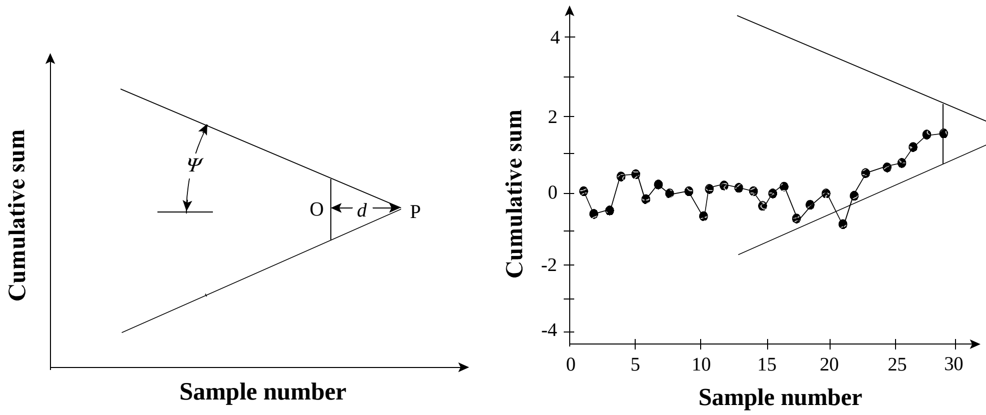
3.1 Cusum charts

The cumulative-sum (or Cusum) control chart directly incorporates all the information in a sequence of sample values. The Cusum can be used to detect when a process deviates from a given target value. If the forecast f_t from the DLM is the target for a process mean, the cumulative-sum control chart is formed by plotting the quantity

$$S_i = \sum_{t=1}^{i} (Y_t - f_t) \tag{1}$$

against the sample number i (Montgomery, 1996).

If the process remains in control, the cumulative sum S_i should fluctuate stochastically around zero. However, if the underlying mean of the process changes, either upwards or downwards, it will affect the cumulative sum in a positive or negative direction. Therefore, if a trend develops in S_i it should be considered evidence that the process mean has shifted. The key issue of the Cusum technique is to detect when the sum S_i starts to drift. Two representations of the Cusum have been considered, the tabular Cusum and the V-mask form of the Cusum. The tabular Cusum (Montgomery, 1996) is calculated as separate upper and lower one-sided Cusums. The tabular upper Cusum works by accumulating deviations from zero that are above the target, and the lower Cusum accumulates the deviations that are below



(a) (b)

Fig. 5. The cumulative-sum control chart. (a) The V mask and scaling. (b) The cumulative–sum control chart in operation.

the target. When the sum of accumulated deviations exceeds a given threshold the process is said to be out of control.

The V-mask version of the Cusum (see e.g. Barnard (1959)) is based on the principles shown in Figure 5. The decision procedure consists of placing the V-mask on the cumulative-sum control chart with the point O on the last value of S_i. If all the previous cumulative sums, $S_1, S_2, \ldots, S_{i-1}$ lie within the two arms of the V-mask, the process is considered to be in control. But if the sum S_i at any time lies outside the arms of the V-mask the process is considered to be out of control.

Both methods have been considered and despite the fact that Montgomery (1996) advises not to use the V-mask Cusum for quality engineering use, the method has been chosen for our monitoring purpose. For the practitioner it is important to be able to distinguish between a drift in the sum of forecast errors caused simply by the pigs growing, and a shift caused by a sudden problem such as outbreak of a disease. This can, to some extent, be seen as long vs. short term control. In Section 3.1.1 it is mentioned how one can distinguish between long and short term with the V-mask version of the Cusum.

See e.g. Page (1954), Johnson and Leone (1974) and Lucas (1976) for a further description of the Cusum control charts.

3.1.1 Design of the V-mask

The V-mask can be defined by the lead distance d, the distance between the origin and the vertex (i.e. the intersection of the two "arms" of the mask), as shown in Figure 5 (a) and the angle ψ defined as half of the angle formed by the V-mask arms. The setting of the lead distance can be interpreted as the model's sensitivity to short term changes while the angle of the arms determines how much (long term) drift can be accepted.

This way of specifying the Cusum is used by, e.g. Johnson and Leone (1974) and Montgomery (1996). The latter suggests choosing

$$d = (\frac{2}{\delta^2}) \ln(\frac{1-\beta}{\alpha}) \tag{2}$$

and

$$\psi = \tan^{-1}(\frac{\Delta}{2A}) \tag{3}$$

where α is the probability of incorrectly concluding that a shift has occurred (a false alarm), β is the advertised probability of failing to detect a shift, and Δ is the shift that it is desired to detect. The term $\delta = \frac{\Delta}{\sigma}$ is the magnitude of the standard deviation shift that we wish to detect where σ is deviation of the process, in this case, the forecast error produced by the DLM . Montgomery (1996) recommends that A lie between σ and 2σ.

Since the model forecast errors are used as an indicator of model collapse and the control facility, the Cusum, is designed to monitor on a fixed scale, it is important that a forecast error of a given numerical size indicates the same degree of instability in the health condition of the pigs. In weaner production, the average daily water consumption per pig increases from approx. 1 litre to 4 litres per day during the 50 days' production period, implying that the average level of the recorded water consumption data increases by a factor of 4. A logical consequence is that a deviation in water consumption at the beginning and at the end of the period, which is identical on a relative scale, will deviate numerically by a factor of 4. Another problem is that there are large differences in the number of animals contributing to the recorded sum of water consumption differs a lot among different housing systems.

Since the variance of the forecast error is expected to vary considerably, the design of the V-mask is complicated. To overcome this problem it seems obvious to transform the forecast errors to a relative scale, which can be done simply by using standardized instead of absolute errors. The standardized forecast error has expectation 0 and, of course, variance 1, which simplifies the scale factor A to lie in the interval $1-2$. The specific values of the parameters Δ, α and β depend on how much the accumulated forecast error should deviate from zero before one can say that there is a problem with the pigs' health condition. This issue is considered in section 4.

4 Illustration

In the preceding section the focus has been on the theoretical structure of the model. In this section, the model's ability to give warnings in critical situations is illustrated by a data set from one of the farms that joined the experiment. The data set was chosen for the following reasons

- The time series is complete, i.e. there are no missing values.
- It contains a sequence where outbreak of diarrhea has been detected.
- It contains a sequence where the pigs have been stressed.
- It contains a sequence where the pigs have been without feed.

Data were recorded from a batch containing 405 piglets, starting on the day of weaning (the average age was 28 days). The data recording continued until the piglets left the production unit after 50 days.

4.1 Setup of the V-mask

As described in Section 3.1.1, the sensitivity of the Cusum depends on the size of the shift in forecast error that it is desired to detect (deviation from zero), the advertised probability of incorrectly concluding that a shift has occurred and the probability of failing to detect a shift. The more sensitive we want the model to be, the higher is the risk of getting a false alarm. We decided to accept 1 % probability for false alarms and 1 % probability for failing to detect a shift. As mentioned in Section 3.1.1 the scale factor A is expected to lie in the interval $1-2$. The choice of A has an influence on the slope of the arms in the V-mask and thereby on the sensitivity of the Cusum (higher values of A narrow the gap of the V-mask). The initial analysis was carried out with the following setup:

- The probability of incorrectly concluding that a shift has occurred, $\alpha = 0.01$.
- The probability of failing to detect a shift, $\beta = 0.01$.
- The shift that it is desired to detect, $\Delta = 1.5$.
- The scale factor $A = 1$.

With $\delta = \frac{\Delta}{\sigma} = \frac{1.5}{1} = 1.5$, the lead distance is

$$d = (\frac{2}{\delta^2})\ln(\frac{1-\beta}{\alpha}) = (\frac{2}{1.5^2})\ln(\frac{1-0.01}{0.01}) = 4.08 \qquad (4)$$

and

$$\psi = \tan^{-1}(\frac{\Delta}{2A}) = \psi = \tan^{-1}(\frac{1.5}{2(1)}) = 36.9^\circ \qquad (5)$$

These values are rounded to $d = 4$ and $\psi = 37^\circ$.

Some aspects of the DLM have to be considered. In Madsen et al. (2005) the optimization of model settings was performed on the basis of one particular data set which was chosen as representative for the recorded time series. It was discussed how many harmonic components should contribute to the cyclic part of the model and it was found that a model with 3 harmonics performed best. The optimal settings of the discount factors were $\delta_T = 0.98$ and $\delta_S = 0.97$ (the discount factors express the rate of decay of information).

To review the robustness of these optimal settings, the optimization procedures described in Madsen et al. (2005) have been applied to the 12 complete time series described in Section 2. Almost all the time series contained sequences with abnormal data due to different causes, for instance data set no 9 has a 20-hour sequence with observations 4 times as high as usual because of a leak in the water pipe. Most of the other irregularities were caused by changing conditions for the pigs, e.g. gastro-intestinal disorders, blockage of feeders, movement of pigs or tail-biting (such events were recorded by the staff in the log book).

Table 2 shows the results of MSE (mean square error) calculations using the DLM's with linear growth and a cyclic sub model containing from 3 to 9 harmonic components (MSE-calculations were made with models containing from 1 to 12 harmonic components). The different models always contain harmonics with the highest possible period, so for instance, in the DLM containing 3 harmonic components, the harmonics have periods 24, 12, and 8 respectively. It appears that for 7 of the time series the DLM with 4 harmonic components is the one that provides the lowest value of MSE. It should be noted that the 4 time series needing more than 4 harmonic components to describe the diurnal drinking pattern are all recorded during the winter, whereas the other 8 series are recorded in the summer. This could indicate that the characteristics of pigs' drinking behaviour might depend on the time of year, although further research is needed before any conclusions can be drawn on that issue. For the rest of the analysis we will use the model containing 4 harmonic components.

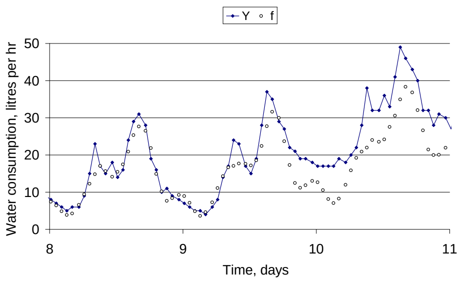
The optimal settings of the two discount factors were determined by minimizing MSE for each data set and each model in the same way as described in Madsen et al. (2005). The discount factor for the trend sub model, δ_T, varies in the interval from 0.88 to 0.98, whereas the discount factor for the cyclic sub model, δ_S, varies from 0.94 to 0.98. These levels are in good agreement with Madsen et al. (2005), where it is shown by an example that the predictive performance of the DLM is more sensitive to changes in δ_S than it is to changes in δ_T. The data sets with low estimates of δ_T and δ_S are typically the ones with many irregular sequences, which appears from the relatively high MSE-values. Low values for the discount factors increase the system variance, W, which implies that the model becomes more adaptive to

Table 2
Mean square error, MSE, values, from DLMs with 3-9 harmonic components, calculated for the 12 time series. For 7 of the series, the model with 4 harmonics provides the lowest value of MSE. High level of MSE values is caused by unpredictable time series. The extreme values from data set number 9 are due to a 20-hour sequence with observations that were 4 times higher than normal because of a leak in the water pipe.

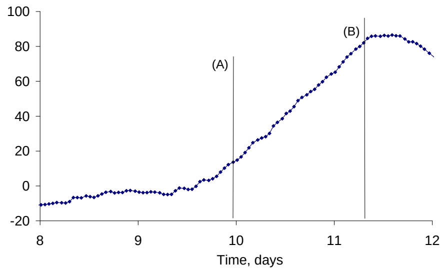
Data set	Number of harmonics							Optimal discount factors	
	3	4	5	6	7	8	9	δ_T	δ_S
	....................MSE....................								
1	127	100	97	**95**	99	102	106	0.91	0.98
2	142	131	105	98	97	**95**	98	0.96	0.97
3	118	**106**	109	114	118	124	135	0.88	0.96
4	115	72	66	**47**	48	51	59	0.98	0.97
5	**40**	40	41	41	42	43	44	0.98	0.97
6	135	**130**	134	141	148	155	164	0.89	0.94
7	789	742	**703**	715	720	734	747	0.93	0.95
8	581	**569**	570	582	603	622	646	0.93	0.95
9	2280	**2228**	2333	2440	2549	2689	2843	0.88	0.94
10	447	**445**	446	464	483	508	534	0.89	0.94
11	406	**358**	369	387	406	425	444	0.88	0.95
12	547	**543**	547	561	579	603	628	0.92	0.95

abrupt changes in data. Especially low values of δ_T make the model fit sudden level changes. But since we want the model to reflect level changes in data by forecast errors, the response obtained with low values of δ_T is not desirable. Therefore the settings $\delta_T = 0.98$ and $\delta_S = 0.97$, found on the regular data set (number 5 in Table 2), are recommended for practical use.

In Madsen et al. (2005), the inclusion of a quadratic growth term in the DLM was discussed. The quadratic model was shown to be much more adaptable to changes in the level of water consumption as compared to the model with only a linear growth term. It seems obvious that the quadratic growth model is the optimal choice if one wants to use the model for accurate step ahead predictions, no matter how the drinking pattern evolves. But when the DLM is used in the context of a monitoring system for detection of deviations from the normal behaviour, it should not adapt to sudden changes in growth rate, i.e. it is the deviation between data and model predictions that is of interest. The problem is illustrated in the following example: A sequence of water consumption data with changing pattern on day 9 (caused by

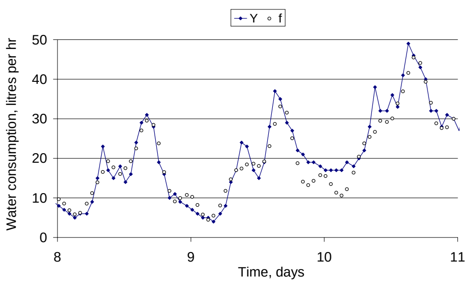


(a) Forecast and observations

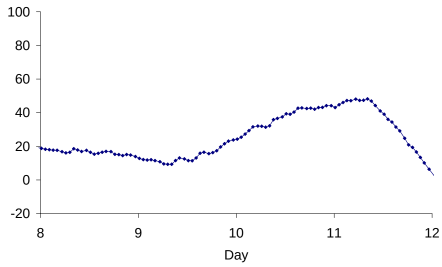


(b) Sum of forecast errors

Fig. 6. Observations (connected with a solid line) and one-hour forecasts from the DLM with only linear growth between time steps (model with 12 harmonics). The model is rather slow in adapting to the new pattern, which was caused by an outbreak of diarrhea resulting in irregular drinking behaviour is well reflected in the sum of forecast errors.



(a) Forecast and observations



(b) Sum of forecast errors

Fig. 7. Observations (connected with a solid line) and one-hour forecasts from the DLM with quadratic growth between time steps (model with 12 harmonics). The model adapts quickly to the new pattern, and the increased water intake is not reflected as clearly as with the linear growth model.

outbreak of diarrhea), is evaluated by the model with only linear growth (Figure 6) and by the model containing a quadratic growth term (Figure 7). Both models contain all 12 harmonics.

The linear model is rather slow in adapting to the new pattern. The effect of the slow response is that the irregular drinking behaviour is well reflected in the sum of forecast errors. The quadratic model, on the other hand, adapts quickly to the new pattern, and the increased water intake is not reflected as clearly as with the linear growth model, therefore, the linear growth model is preferred since the purpose of monitoring is to detect changes in the drinking pattern.

With the setup described above, the Cusum has been applied to the series of forecast errors from the DLM. Each time an out-of-control alarm is issued, the sum is reset to zero.

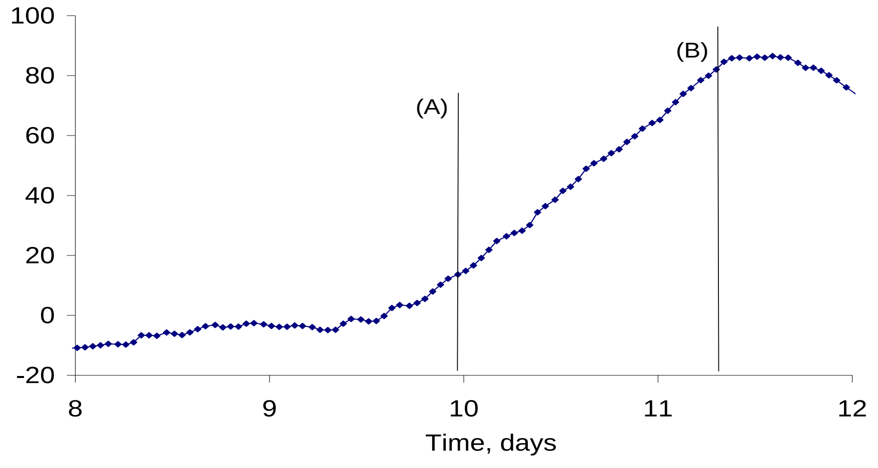


Fig. 8. Sum of forecast errors. An alarm is issued because of deviating drinking pattern at point (A). Outbreak of diarrhea is detected and treated at point (B). Under the circumstances, the sum would have been reset to zero at point (A), but for illustration of the course it been omitted.

Figure 8 shows a sequence of forecast errors from a batch of piglets with an outbreak of diarrhea. When the V-mask is applied to the sum of forecast errors, it reaches the threshold where the process is considered out of control at the point that corresponds to 11 pm on day 9 after weaning (the resetting to zero is not shown in the figure). The disease was detected by the caretaker at 8 am on day 11 (the caretaker did not have access to the monitoring system). In this case, the monitoring method provided information of the disease 33 hours before it was actually detected. Of course, since there was no visual inspection of the pigs between 4 o'clock in the afternoon and 8 o'clock the next morning, visual symptoms of the disease could have been present somewhere in that time interval. But evidently, symptoms were not present at 4 pm on day 10 and it can thereby be concluded that the monitoring system, in this case, reacted at least 17 hours before visual symptoms occurred.

Another example (Figure 9) shows the level component from the DLM for a 48-day production period of piglets. The vertical bars indicate the alarms issued by the monitoring system. The alarms on day 4 and day 9 are caused by an outbreak of diarrhea. The huge deviations in the level of water consumption around day 39

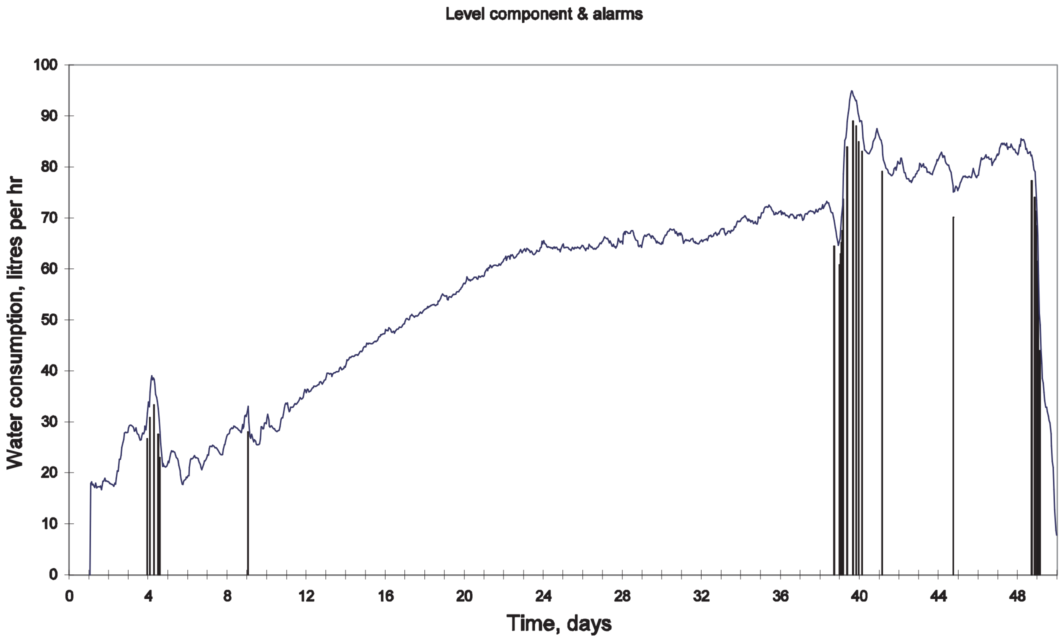


Fig. 9. The alarms from the Cusum are shown as bars together with the level component from the DLM. The alarms on day 4 and day 9 are caused by an outbreak of diarrhea. The huge deviations in the level of water consumption around day 39 and day 40 are due to a 12-hour stop in the automatic feeders, whereas there is no explanation of the alarm on day 41. On day 44, formic acid was added to the drinking water, and finally on day 48, the pigs were taken out of the compartment.

and day 40 are due to a 12-hour stop in the automatic feeders, whereas there is no explanation of the alarm on day 41 (unless it is a delayed consequence of the feed stop). On day 44, formic acid was added to the drinking water, and finally on day 48, the pigs were taken out of the compartment.

5 Discussion

Although this study is only based on 12 time series, it strongly indicates that monitoring young pigs' drinking behaviour can provide useful information for managing the production as suggested by Bird and Crabtree (2000). The method for monitoring the drinking behaviour of young pigs, based on a combination of a dynamic linear model and a Cusum control chart, has proved to be a useful tool in detection of diseases and other production related problems that affect the pigs' water intake.

The approach has been to pose a model that fits the normal pattern of growing pigs and then to detect when data depart from the model. The method gives simple information: alarm/no-alarm at each time step, combined with the information on whether the alarm is caused by an increase or a decrease in the observed time series. An alarm does not give information on which problem caused the deviating drinking behavior, but in combination with the caretaker's knowledge of the problems that are common at a given stage of the pigs' life cycle, he should often be able to point out the possible causes. If, for instance, it is usual in a given herd to see outbreaks of diarrhea occurring 3-7 days after weaning and there is an alarm for increasing water consumption on day 5, it is likely that treatment should be considered.

Other approaches for detection of changes in the drinking behaviour could have been chosen. Thysen (1992) modeled the level of somatic cell count in milk by using a multi-state dynamic linear model. The model had three possible states: normal level, an outlier, or a change of level. For each time step, the probability was calculated for the three models and systematic changes in level were indicated by high probability on the level shift model. The multi-state approach is not straightforward in our case since the deviation from the normal model cannot be described only by one single (level shift) model. Deviations can be caused by changes in the diurnal pattern with a steady overall level or by changes in the level with steady diurnal pattern. Sometimes changes are seen as increased water consumption at night but unchanged consumption during the day. As a consequence, a multi-state model would have to contain a number of competing submodels and they would all have to be quantified in terms of their specific way of deviating from the normal model. The multi-state model might, on the other hand, give more detailed information on the <u>kind</u> of changes in the pigs' behaviour when there is a problem.

Indoor temperature was recorded as supplementary registrations for all the time series, to estimate the effect of high temperature on water consumption. But the temperature never exceeded 28 C within the period of data recording, and below that temperature there was no effect on the drinking behaviour. However, it is likely that temperatures above 28-30 C will have some effect on the drinking pattern, and if that is the case, the effect should be incorporated as a regression effect in the dynamic linear model.

The specification of prior distributions is necessary to initialize the model. If there is no information available except the time series itself, the model can be initialized by means of <u>reference analysis</u>. Reference analysis uses the first observations of the series in question to estimate the parameters. For practical purposes the method uses an observation for each of the model parameters (including V) to obtain a fully specified joint posterior distribution of the parameters. During the reference analysis, it is assumed that the system variance is zero (i.e. $\mathbf{W}_t = \mathbf{0}$), non-zero matrices would allow for change, which cannot be estimated since any changes in a parameter is impossible to detect before an estimate of the parameter exists. A detailed description of the estimation method can be found in West and Harrison (1997) pp 128-136. The use of dynamic linear models with "batch-specific" parameters, initialized by reference analysis, makes it possible to formulate a general model that applies for all batches even though there are differences in the daily drinking pattern. The only task is to determine the number of harmonic components that should be included in the model and the setting of the discount factors. A model with 12 harmonics contains enough parameters to describe any diurnal pattern, but if the structure of a specific drinking pattern can be described with less harmonics, the spare ones only provide unnecessary noise in the model. It has been shown that the optimal model for 8 of the 12 time series contained 4 or less harmonics, but general recommendations on this issue should be based on more data than were available for this study.

It should be noted that the data sets contain several sequences with deviating drinking behaviour for which there are no explanation, i.e. nothing is reported in the log book. This should not, however, be regarded as poor performance by the model, but rather as an indication of the fact that the pigs have been affected by a kind of stressor that has not been observed by the caretaker. The exact number of such alarms without explanation will depend on the settings of the V-mask. In other words, if the manager thinks that the system issues too many alarms, the settings may be adjusted.

The described model for monitoring the behaviour of pigs produced in "all-in-all-out" operations seems to have great potential for practical implementation. A warning system monitoring each batch at farm level would improve the caretaker's opportunity to concentrate the efforts on those housing units where the pigs show deviating behaviour. The possibility of detecting outbreak of diseases like diarrhea before any visual signs are seen gives potential for reduced use of medicine as well as improved productivity.

6 Conclusion

The present study indicates that pigs' water consumption is characterized by a distinct circadian pattern, which is a prerequisite for utilizing drinking behaviour for detection of production and health in growing pigs.

The system for monitoring the drinking behaviour of young pigs based on a combination of a dynamic linear model and a Cusum control chart has proven to be a useful tool in modelling water consumption rate in pigs including forecasts of altered water intake. The method utilizes the fact that pigs' diurnal drinking pattern is stable as long as they are healthy while they often change their drinking behaviour when they are affected by a disease or a stressor. Information of this kind might alert the caretaker that intervention is needed in a specific pig building for prevention of diseases and stressors. No systematic investigation has been conducted to identify what diseases lead to changes in the pattern of water consumption, but an example has shown that an outbreak of a diarrhea can be detected by the method approximately one day before physical signs are seen on the pigs.

References

Barnard, G., 1959. Control charts and stochastic processes. Journal of the Royal Statistical Society 21.

Bigelow, J. A., Houpt, T. R., 1988. Feeding and drinking patterns in young pigs. Physiology & Behaviour 43 (1), 99–109.

Bird, N., Crabtree, H., Schofield, C. P., 2001. Engineering technologies for pig production. No. 28 in British Society of Animal Science Occational Publication. BSAS, Edinburgh, pp. 105–112.

Bird, N., Crabtree, H. G., 2000. Real-time production performance monitoring. The Pig Journal 45, 55–68.

Christensen, J., Rudemo, M., 1996. Multiple change-point analysis of disease incidence rates. Preventive Veterinary Medicine 99, 53–76.

Csörgö, M., Horváth, L., 1988. Nonparametric methods for changepoint problems. Handbook of Statistics 7, 403–425.

de Mol, R., Keen, A., Kroeze, G., 1999. Description of a detection model for oestrus and diseases in dairy cattle based on time series analysis with a kalman filter. Computers and Electronics in Agriculture 22, 171–185.

Frost, A., Schofield, C., Beaulah, S., Mottram, T., Lines, J., Wathes, C., 1997. A revview of livestock monitoring and the need for integrated systems. Computers and Electronics in Agriculture 17, 139–159.

Hyun, Y., Ellis, M., McKeith, F. K., Wilson, E. R., 1997. Feed intake pattern of group-housed growing-finishing pigs monitored using a computerized feed intake recording system. Journal of Animal Science 75, 1443–1451.

Iwersen, J., 1997. Statistical process control for autocorrelated processes: A case-study. Frontiers in Statistical Quality Control 5, 138 – 155.

Johnson, N., Leone, F., 1974. Statistics and Experimental Design, Second Edition, Volume 1. John Wiley & Sons, Inc.

Lucas, J., 1976. The design and use of cumulative sumquality control schemes. Journal of Quality Technology 8.

Madsen, T. N., Andersen, S., Kristensen, A. R., 2005. Modelling the drinking patterns of young pigs using a state space model. Computers and Electronics in Agriculture, Submitted.

Montgomery, D., 1996. Introduction to Statistical Quality Control, 3rd Edition. John Wiley & Sons.

Nienaber, J. A., McDonald, T. P., Hahn, G. L., Chen, Y. R., 1991. Group feeding behavior of swine. Transactions of the ASAE 34 (1), 289–294.

Page, E., 1954. Continuous inspection schemes. Biometrika 41, 100–114.

Roush, W. B., Tomiyama, K., Garnaoui, K. H., D'Alfonso, T. H., Cravener, T. L., 1992. Kalman filter and an example in poulty production responses. Computers and Electronics in Agriculture 6, 347–356.

Slader, R. W., Gregory, A. M. S., 1988. An automatic feeding and weighing system for ad libitum fed pigs. Computers and Electronics in Agriculture 3, 157–170.

Smidt, D., 1983. Advances and problems of using integrated systems of indicators as compared to single traits. In: D. Smidt. (Ed.) Indicators Relevant to Farm

Animal Welfare. Martinus Nijhoff Publishers, Boston, Mass.
Thysen, I., 1992. Detection of changes in daily milk production by a multi-state kalman filter. Dina Rapport No. 3.
Van Bebber, J., Reinsch, N., Junge, W., Kalm, E., 1999. Monitoring daily milk yields with a recursive test day repeatability model (kalman filter). Journal of Dairy Science 82 (11).
West, M., Harrison, J., 1997. Bayesian Forecasting and Dynamic Models., 2nd Edition. Springer.
Xin, H., DeShazer, J. A., January-February 1992. Feeding patterns of growing pigs at warm constant and cyclic temperatures. Transactions of the ASAE 35 (1), 319–323.
Young, R. J., Lawrence, A. B., 1994. Feeding behavior of pigs in groups monitored by a computerized feeding system. Animal Production 58, 145–152.