Modelling the Drinking Patterns of Young Pigs Using a State Space Model

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

Madsen, T.N., S. Andersen & A.R. Kristensen. 2005. Modelling the Drinking Patterns of Young Pigs Using a State Space Model. *Computers and Electronics in Agriculture 48*, 39-62.

DOI: 10.1016/j.compag.2005.01.001

Modelling the Drinking Patterns of Young Pigs Using a State Space Model

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

[a]*Department of Large Animal Sciences, 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

In normal situations pigs show a stable diurnal drinking pattern. Based on experimental data a dynamic model is developed for prediction of the drinking behavior of growing pigs. A state space model with cyclic components is proposed for modelling the diurnal drinking pattern, measured as hourly sums. Determination of variance structure by use of discount factors is suggested. Model performance is investigated by error analysis. The final model contains three cyclic components.

Key words: Dynamic linear model, drinking behaviour, diurnal pattern

1 Introduction

Studies carried out by The National Committee for Pig Production, Danish Bacon & Meat Council (Kai et al., 1999), Bird and Crabtree (2000) and Bird et al. (2001) have pointed out the potential of monitoring the drinking behaviour of growing pigs. In normal situations pigs show a stable diurnal drinking pattern, whereas outbreak of diseases, changes in the quality of feed or ventilation problems often make the pigs' drinking behaviour deviate from the normal pattern. Therefore, real time

* Corresponding author. Present address: TNM Consult Aps, Gydevang 39-41, DK-3950 Allerød, Denmark
Email addresses: `tnm@tnmconsult.dk` (Thomas Nejsum Madsen), `sqa@novonordisk.com` (Søren Andersen), `ark@dina.kvl.dk` (Anders Ringgaard Kristensen).

monitoring of growing pigs' water consumption seems to be a possible way of improving management (Bird and Crabtree, 2000). In order to be able to detect changes in drinking behaviour, it is crucial to have a well founded model to predict the expected behaviour.

The aim of this study is to develop a dynamic model for prediction of the drinking behaviour of growing pigs. The model is part of a computer based monitoring system for all-in all-out management of weaner and finishing accommodations described in Madsen and Kristensen (2005) and Madsen (2001). Water consumption is, of course, a continuous process, but for monitoring purposes the consumption has to be measured in discreet intervals, that is, the amount of water consumed within a given time period. In Madsen and Kristensen (2005), different time intervals are considered and it is concluded that one-hour sums is the the preferable choice when modelling the observed water consumption.

Only very few studies (Bird and Crabtree, 2000; Bird et al., 2001) into pigs' drinking pattern have been found in the literature whereas their eating behaviour seems well described. In some of the literature the eating pattern of group-housed growing-finishing pigs is calculated as an average over the whole test period. In Hyun et al. (1997) the diurnal pattern is estimated as an average over 10 weeks, and Slader and Gregory (1988) use an average over 53 days. Nienaber et al. (1990) made the analysis more flexible by dividing the test period into 5 intervals. For each interval the feeding pattern is estimated as averages of 7-day periods. Common to all of the studies mentioned above is the lack of dynamics. The daily pattern is estimated as an average over a number of days. The disadvantage of such a method is that it does not reflect a change in eating pattern over the period.

For the one-hour ahead predictions of water consumption a model with a more dynamic nature is required, allowing for development in the diurnal pattern as well as in the overall level of water consumption, as the pigs grow. The work described in this paper is based on a Bayesian way of modelling and forecasting and includes Kalman filter techniques. Data are modeled by use of a Dynamic Linear Model (DLM) which is dynamic in nature because it explicitly allows for parameter values to change as time passes, and new observations add to the information. More attention is paid to recent information than to past information, so as time passes information looses its value.

In the literature, there is a variety of studies describing monitoring systems based on Bayesian methods. In human medicine examples are given by Lundbye-Christensen (1991) and Lundbye-Christensen and Christensen (1995). Within the agricultural area, examples include estimation of the lactation curve of a dairy cow (Goodall and Sprevak, 1985), detection of changes in feed consumption in broilers (Roush et al., 1992), an expert system for interpreting control charts (Cook et al., 1992), detection of changes in daily milk production (Thysen, 1992), monitoring bulk somatic cell count (Thysen, 1993), and a model for detection of oestrus and diseases in dairy cattle (de Mol et al., 1999).

Monitoring one-hour sums of growing pigs' water consumption complicates the model structure somewhat compared with the studies described above, i.e. the model must not only include the linear or quadratic growth in the overall level, as the pigs grow, but it also has to contain a cyclic effect to describe the diurnal pattern. Cyclical or periodic models are widely used in commercial business, e.g. in modelling the annual deviations in the demand for oil and gas. Examples of such models are given in Pole et al. (1994) and in West and Harrison (1997).

In this paper modelling the pigs' drinking behaviour is intended as a part of a monitoring system for farm management, but the principles stated might also be useful in general research on animal behaviour.

2 Materials

In connection with the development of a new computer based monitoring system [1] The National Committee for Pig Production, Danish Bacon and Meat Council has installed water flow-meters and micro computers in 4 production units for young pigs, located on 3 different farms in Denmark. The production units had a capacity of 400-900 pigs each, and the production was performed in batches with all-in all-out operations. On all 3 farms there was an automatic dry feeding system with wet/dry feeders (tube feeders).

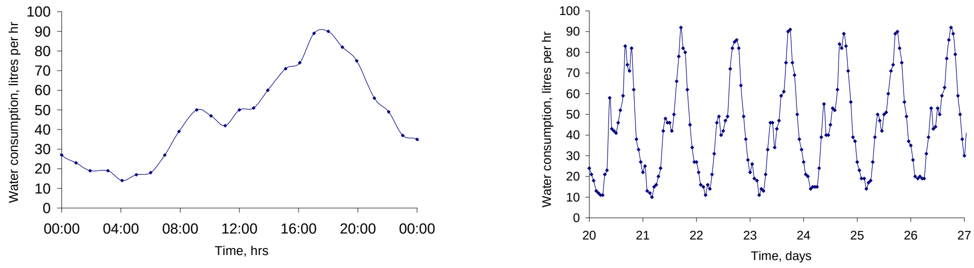
Data from 18 batches have been collected, and from this sample one representative time series has been chosen for illustration. In the following we will refer to this time series as <u>the reference data set</u>. The data set is chosen for the following reasons:

- The time series is complete, i.e. there are no missing values.
- The number of pigs contributing to the recorded sum of water consumption is almost constant all through the time series.
- There have not been any serious problems with diseases or with the feed.

Data were recorded during the months of February and March from a batch containing 405 piglets, starting on the day of weaning (the average age was 28 days), and the recording continued until the piglets left the production unit after 45 days. Pigs were fed ad lib with dry feed through the entire period. Water disappearance was recorded by an electronic water flow meter (Skov Dol92 [2]) connected to a process computer (Skov Dol34). The accuracy of the flow meter was specified by the manufacturer to be 5%. It issues 4 impulses per litre corresponding to a water flow

[1] The Farm*Watch*™ system.

[2] Manufactured by Skov A/S, Hedelund 4, Glyngøre, DK-7870 Roslev, Denmark



(a) 1-hour sums, 1 day (b) 1-hour sums, 7 days

Fig. 1. One-hour sums of water consumption recorded from a group of 400 piglets.

of 250 ml per impulse. Data on water consumption were stored in the computer as the number of whole litres (i.e. the number of impulses divided by 4) consumed within a 2 minute period. The 2 minute sums were then grouped into 1 hour sums (1056 observations). All calculations and illustrations in this paper are based on the reference data set which, besides the water consumption data, contains a log book in which the caretaker has reported actions such as treatment of sick animals, change of feed, removal of dead animals, etc.

The 1 hour sums of water consumption (Figure 1) show a very stable pattern (for a longer period, reference is also made to Figure 6), and can be characterized by the following properties:

- The level of water consumption increases as the pigs grow. This is, in particular, clear from Figure 6.
- The drinking pattern is rather stable from day to day in a normal situation as shown in Figure 1.
- The daily drinking pattern evolves as the pigs grow. If the pattern is followed over a longer period than shown in Figure 1 it becomes clear that, for instance, the amplitude of the cycles become bigger over time as is seen in Figure 6(a).
- Data contain some random noise because of biological variation and measurement errors.

A further description of data can be found in Madsen and Kristensen (2005), and a plot of the entire time series is shown in Madsen (2001).

In the following section a model that was designed to fit the characteristics of the water consumption data is presented.

3 Model

Consider the water intake as a stochastic observable process which we want to monitor. We do have some idea of how the observable process will evolve as the water intake basically is an effect of the latent physiological processes in the pigs.

Data are modeled by means of a Dynamic Linear Model (DLM) which is well suited to model the dynamic/cyclic evolution in data, see, e.g. Pole et al. (1994). Usually, the DLM is used as a tool for making forecasts, based on prior knowledge including former observations. The following model description is mainly based on West and Harrison (1997) and Pole et al. (1994).

3.1 A univariate dynamic linear model

The data recording is performed on a section containing a given number of pigs. The observation, Y_t, is the amount of water consumed by the pigs within the last hour at time t.

The information obtained by observing $Y_1, \ldots, Y_t$ and the initial information present at time $t = 0$ is defined as the information set D_t. The information set D_t is given by:

$$D_t = D_{t-1} \cup Y_t \tag{1}$$

D_0 being a set of information available at time $t = 0$.

Let $\boldsymbol{\theta}_t$ be a vector of parameters at time t, let $\mathbf{F}$ be a matrix of coefficients (the design matrix) and let v_t be a random observation error. The observation equation of a general dynamic linear model is then defined for each time t by:

$$Y_t = \mathbf{F}'\boldsymbol{\theta}_t + v_t \qquad v_t \sim N(0, V). \tag{2}$$

The system equation of a general linear model describes how the parameter vector may change over time:

$$\boldsymbol{\theta}_t = \mathbf{G}\boldsymbol{\theta}_{t-1} + \mathbf{w}_t \qquad \mathbf{w}_t \sim N(0, \mathbf{W_t}), \tag{3}$$

where $\mathbf{G}$ is a matrix of coefficients (the system matrix) and $\mathbf{w}_t$ is a random system evolution error. In order to initialize the model, the distribution

$$(\boldsymbol{\theta}_0 \mid D_0) \sim N(\mathbf{m}_0, \mathbf{C}_0), \tag{4}$$

of the parameter vector at time $t = 0$ (before any observations are made) must be specified for some prior moments $\mathbf{m}_0$ and $\mathbf{C}_0$. The observational and evolution error sequences v_t and $\mathbf{w}_t$ are assumed to be internally and mutually independent, and are independent of $(\boldsymbol{\theta}_0 \mid D_0)$.

The sequence $\boldsymbol{\theta}_t$ is called a latent process. The mean response or the expected value of Y_t given $\boldsymbol{\theta}_t$ is $\mathbf{E}[Y_t \mid \boldsymbol{\theta}_t] = \mathbf{F}'\boldsymbol{\theta}_t$. The mean response can be interpreted as the underlying true level of water consumption and it differs from the observed value Y_t by the observational error term v_t. Contributions to the error term could be biological variation or measurement errors. Water waste is an example of such variation, but with modern drinking bowls, the waste is almost non-existing. The conditional expectation of $\boldsymbol{\theta}_t$ is $\mathbf{E}[\boldsymbol{\theta}_t \mid \boldsymbol{\theta}_{t-1}] = \mathbf{G}\boldsymbol{\theta}_{t-1}$. The evolution in time of $\boldsymbol{\theta}_t$ is a one-step Markov process where $\boldsymbol{\theta}_t$ is independent of D_{t-1}, given $\boldsymbol{\theta}_{t-1}$, the design matrix $\mathbf{G}$ and the evolution variance $\mathbf{W}$. The latter allows the model dynamically to adapt to systematic changes in data, e.g. changes in the diurnal pattern.

3.2 Specifying the design and system matrices

The design and system matrices are defined by superposition of a linear growth model and a cyclic model with a period of 24 (24 hourly sums). The cyclic model itself is also constructed by superposition of several sub models. The principle of superposition simply states that a linear combination of DLM's is a DLM. Hence, the model can be constructed by combining a number of simple models each handling a certain aspect of the data.

The basic idea of a linear growth model is to describe the observation and system equations as:

$$\begin{aligned} Y_t &= \mu_t + v_t \\ \mu_t &= \mu_{t-1} + \beta_t + w_{t,1} \\ \beta_t &= \beta_{t-1} + w_{t,2} \end{aligned} \tag{5}$$

where the incremental growth β_t evolves during the addition of the stochastic element $w_{t,2}$. The level μ_t at time t evolves systematically via the addition of the growth term β_t and the stochastic term $w_{t,1}$.

The cyclic pattern could be described by a set of fixed factors, one for each observation in a cycle. In our case the fixed factor model would require 24 parameters to describe the daily drinking pattern.

An alternative representation of the cyclic pattern employs linear combination of trigonometric functions, called the Fourier form representation. One of the main reasons for using this representation is the possibility of economising on parameters. If data contain some kind of underlying sine/cosine wave form, it will often be possible to reduce the number of cyclic components, and thereby the number of parameters, in the model. Any cyclic pattern with a period of 24 can be described by a complete model with 12 different harmonic waves, including 23 parameters. When monitoring the daily drinking pattern of pigs, some of the harmonics can be left out of the model without reducing the predictive performance, as is shown in section 4.1.

The system matrix describing a harmonic wave is:

$$\mathbf{G} = \begin{pmatrix} \cos(\omega) & \sin(\omega) \\ -\sin(\omega) & \cos(\omega) \end{pmatrix} \tag{6}$$

where $\omega \in]0, \pi[$. Choosing $\omega = 2\pi/24$ yields a wave with a period of 24, and $\omega = 2\pi/12$ results in a 12 period model, etc.

The harmonic having the shortest period (i.e. the period 2) is called the Nyquist harmonic, and it has a particular simple form, because it just oscillates between a maximum and a minimum. The system matrix is therefore very simple:

$$\mathbf{G} = (-1) \tag{7}$$

The full model is constructed as a combination of the linear growth model having two parameters (μ_t and β_t), the 11 cyclic models each having 2 parameters and the Nyquist harmonic having 1 parameter. The result is a 1×25 design matrix

$$\mathbf{F} = \begin{pmatrix} 1 \ 0 \ 1 \ 0 \dots 1 \end{pmatrix}'$$

and a 25×25 system matrix

$$\mathbf{G} = \begin{pmatrix} 1 & 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & \cos(\omega) & \sin(\omega) & & 0 & 0 & 0 \\ 0 & 0 & -\sin(\omega) & \cos(\omega) & & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & 0 \\ 0 & 0 & 0 & 0 & \dots & \cos(11\omega) & \sin(11\omega) & 0 \\ 0 & 0 & 0 & 0 & \dots & -\sin(11\omega) & \cos(11\omega) & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}$$

where $\omega = 2\pi/24$.

Each of the 25 elements of the state vector $\boldsymbol{\theta}_t = (\theta_{t,1}, \ldots, \theta_{t,25})$ has a special meaning. $\theta_{t,1}$ is the underlying, non-cyclic level of the series at time t. $\theta_{t,2}$ is the non-cyclic growth between time $t-1$ and t, whereas $\theta_{t,3}, \theta_{t,4}, \ldots, \theta_{t,25}$ are the Fourier coefficients describing the cyclic pattern.

3.3 The Kalman filter

The dynamic linear model is dynamic in its nature because it explicitly allows for parameter values to change as time passes and new observations add to the information. For each new observation, the parameters are updated according to the Kalman filter introduced by Kalman (1960). The equations update the posterior distribution for the state vector $\boldsymbol{\theta}_t$ when a new observation Y_t is available. The Kalman filter produces a posterior multi-dimensional normal distribution, $N(m_t, C_t)$, of the state vector, based on all prior information, including the new observation. The exact form of the updating equations may be found in West and Harrison (1997).

3.4 Variance components in the Kalman filter

In the general Kalman filter described in section 3.3, the variance terms V and $\mathbf{W}$ (defined in Eqs. (2) and (3)) are assumed constant and known. It is doubtful whether it is reasonable to assume $\mathbf{W}$ constant when the drinking pattern of growing pigs is considered. In the weaner production, the average daily water consumption per pig increases from approximately 1 litre to 4 litres per day during the 45 days' production period, implying that the average level of the recorded water consumption data increases by a factor 4 . This four-fold increase in water consumption level over time has, as a logical consequence, a four-fold numerical change in deviation in water consumption, even though the actual deviation at the beginning and end of the period may be the same on a relative scale.

The observational variance V depends, among other things, on the amount of water wasted when the pigs are drinking. The waste of water is conditioned by the design of the drinking bowls. The observational variance is also affected by the accuracy of the water flow meters. These facts suggest that the level of V is rather stable from one batch to the next within the same farm, but there could be huge differences between farms.

The data on water consumption give the following problems concerning the variance components:

- The system variance, $\mathbf{W}$, is probably changing as the pigs grow up.
- The observational variance, V, should be rather constant within the same time series.

These problems lead to a more informal handling of the variance data, as described in the following section.

3.5 Unknown variance

It seems reasonable to assume that the observational variance, V, is constant during the growth period, but varying between batches. When V is unknown an estimation procedure is necessary. The idea is that V, assuming it is constant, is re-estimated for each new observation.

The main consequences of unknown, constant variance are that the updated parameter vector $\boldsymbol{\theta}_t$ is then distributed according to a Student T distribution that converges to the standard normal distribution as t increases and that the estimated precision $\phi = V^{-1}$ becomes Gamma distributed as described in details by West and Harrison (1997).

In the system equation (equation 3), $\mathbf{W}_t$ leads to an increase in uncertainty about the state vector between times $t-1$ and t.

The system variance $\mathbf{W}_t$ can be defined as a fixed proportion of $\mathbf{C}_t$ by use of discount factors as an aid to choosing $\mathbf{W}_t$ (West and Harrison, 1997). By definition, a discount factor δ satisfies the condition $0 < \delta \leq 1$. For each time step t the system variance is determined as a given proportion of the model variance $\mathbf{C}_t$

$$\mathbf{W}_t = \frac{1-\delta}{\delta}\mathbf{P}_t \tag{8}$$

where

$$\mathbf{P}_t = \mathbf{G}\mathbf{C}_t\mathbf{G}' = V[\mathbf{G}\theta_{t-1} \mid D_{t-1}]. \tag{9}$$

The initial prior distribution for $\mathbf{P}_t$ is provided by reference analysis, as described in Section 3.7.

3.6 Component discounting

The discount factors are independent of the measurement scale of the observation series, which is a useful feature when modelling the drinking pattern of growing pigs. Another advantage is that, because of the conditional independence structure of a DLM comprising the superposition of several sub-models, each sub-model can have its own discount factor. If, for instance, the cyclic characterization is more durable than the trend, this can be incorporated directly in the evolution of the model. Since the system matrix $\mathbf{G}$ is block diagonal, as described in Section 3.2, it was decided to assume independence between the the blocks so that the system variance $\mathbf{W}_t$ becomes block diagonal as well. Given eq. (8) in section 3.5 the block discounting of the evolution variance is

$$\mathbf{W}_t = \text{block diag}\left[\mathbf{P}_{tT}(\delta_T^{-1} - 1), \mathbf{P}_{tC}(\delta_C^{-1} - 1)\right] \tag{10}$$

where:

- $\mathbf{P}_t = \mathbf{G}\mathbf{C}_{t-1}\mathbf{G}'$
- $\mathbf{P}_{tT}$ is the block of $\mathbf{P}_t$ concerning the trend component.
- $\mathbf{P}_{tC}$ is the block of $\mathbf{P}_t$ concerning the cyclic component.
- $0 < \delta_T < 1$ is the discount factor for the trend component.
- $0 < \delta_C < 1$ is the discount factor for the cyclic component.

By defining discount factors for the trend component and the cyclic component individually it is possible to allow for different beliefs in the value of information received in the updating. The bigger the discount factor, the less information is believed to be received in each new observation.

Discount factors in the range of 0.8 - 1 are usually suggested for practical modelling, see e.g. Pole et al. (1994). It is desirable to choose the discount factors in a way that maximizes the models' ability to make forecasts, in periods where the time series is in a stable mode.

To get an impression of how differing values of the discount factors affect the predictive performance of the model, it was applied to the reference data set with δ_T and δ_C varying in the interval 0.80 to 0.99 by 0.01, resulting in 400 successive runs. Initial specifications of model parameters were made by means of reference analysis, as described in Section 3.7. The mean square error, or MSE defined as $\frac{1}{n}\sum e_t^2$ was used as a measure of model predictive performance for comparison of the alternative models. The surface plot in Figure 2 shows the result of changing the discount factors. The MSE has minimum in $\delta_T = 0.98$ and $\delta_C = 0.97$. The plot indicates that the model is more sensitive to changes in δ_C as compared to δ_T, and it is seen that MSE has reasonably low values when δ_C is in the interval 0.96 to 0.98.

3.7 *Reference analysis*

The specification of prior distributions is necessary to initialize the model. If there is no information available except from the time series itself, the model can be initialized by means of reference analysis.

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

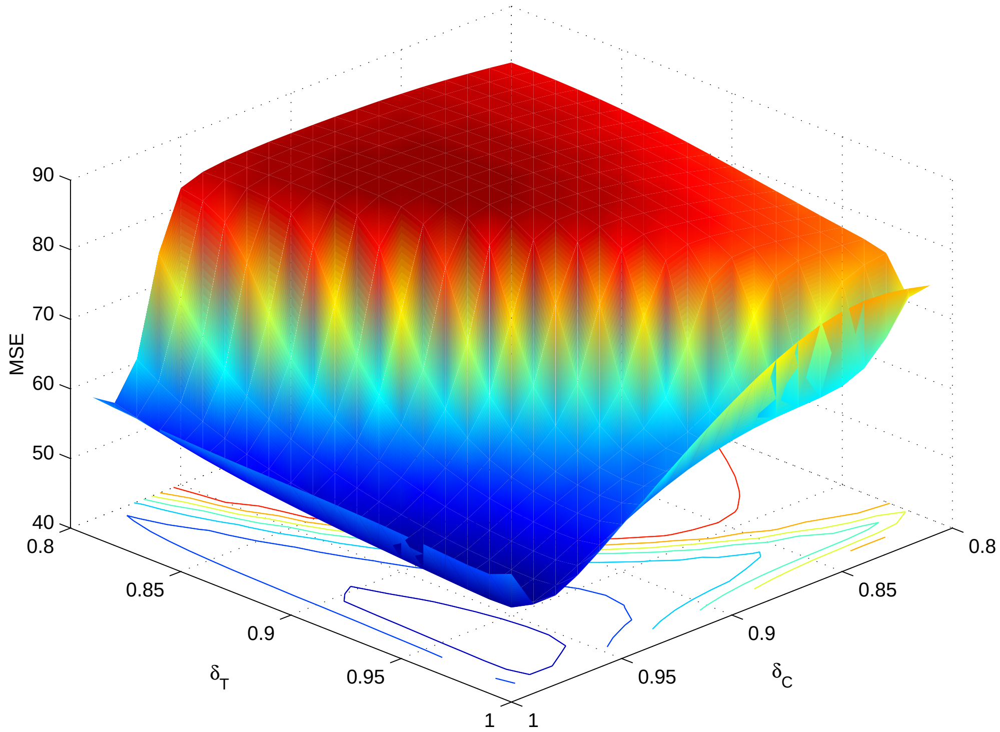


Fig. 2. The total square error, MSE, with minimum in $\delta_T = 0.98$ and $\delta_C = 0.97$.

variance is zero (i.e. $\mathbf{W}_t = \mathbf{0}$), non-zero matrices would allow for changes which cannot be estimated since it is impossible to detect any changes in a parameter 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).

3.8 Missing values

So far, all updating equations have been based on the assumption that the observed time series is complete. Practical experience with the recording of water consumption data has shown that missing values might occur, e.g. due to a power outage or a computer shutdown. In these cases the information set at time t is just $D_t = D_{t-1}$, and since there is no new information the posterior distributions are equal to the prior distributions. The forecast error e_t is, of course, zero.

4 Application of the model

In the preceding section the focus has been on the theoretical structure of the model. In this section, the model's performance is illustrated by the reference data set. The complete time series is illustrated in Figure 6(a). The figure is rather diffuse and does not contain much information, except that it is clear that the amount of water consumed increases as the pigs grow.

4.1 Model reduction

The full model described in Section 3.2 contains 24 parameters describing the cyclic pattern which gives the model a high degree of flexibility allowing it to adapt to any 24-hour pattern.

The advantage of using seasonal instead of fixed factors, described in Section 3.2, in the cyclic part of the model, is the possibility of omitting some of the harmonics and thereby reducing the number of model parameters. It can be argued that when appropriate, a reduced form model produces a better forecasting performance than a full model, because the harmonic components that have very little or no effect degrade forecast performance since their assessment introduces extra variation and correlated forecast errors.

The very stable and systematic pattern in the pigs' drinking behaviour appearing in Figure 1 suggests that the 24-hour cycles can be described with less than 12 harmonics. The importance of the individual harmonics has been evaluated by an informal statistical test. At time t, let the posterior distribution for the coefficient vector $\boldsymbol{\theta}_t$ be $(\boldsymbol{\theta}_t|D_t) \sim N[\mathbf{m}_t, \mathbf{C}_t]$. Let $\boldsymbol{m}'_{tr} = (m_{t,2r+1}, m_{t,3r})$ be the posterior mean of the two parameters describing harmonic r and let $\mathbf{C}_{tr}$ be the corresponding 2×2 posterior variance-covariance matrix of the same two parameters (cf. Section 3.2). Then the significance is determined by calculating the quantity $\mathbf{m}'_{tr}\mathbf{C^{-1}}_{tr}\mathbf{m}_{tr}/2$ (West and Harrison (1997), pp 257) and comparing it to the F-distribution F_{2,n_t} with r being the harmonic number $(1, \ldots, 12)$ and n_t the number of valid observations at time t. Each value of the quantity $\mathbf{m}'_{tr}\mathbf{C^{-1}}_{tr}\mathbf{m}_{tr}/2$ relates to a given component in the full model, i.e. the first value $(r = 1)$ relates to the first harmonic, the 2nd value $(r = 2)$ to the 2nd harmonic, etc.

The full model was applied to the reference data set and F-values were calculated at each time step. Discount factors of $\delta_T = 0.98$ and $\delta_S = 0.97$ were used. Initial information was provided by reference analysis. Values concerning harmonics 1 to 11 were compared to the $F_{2,n'}$, and the 12th harmonic to the $F_{1,n'}$, where n' denotes the number of valid observations minus 24, since the first 24 observations were used for reference analysis. So, for the first 11 harmonics, the probability for observation number 25 was compared to $F_{2,1}$, number 26 to $F_{2,2}$, etc. The results can be interpreted as the significance of each harmonic, given all other harmonics are present.

The importance of the individual harmonics turned out to be very fluctuating as is seen from the example in Figure 3. The figure shows the F-values from harmonic 5 together with the 90 % probability limit. It clearly appears that the observations

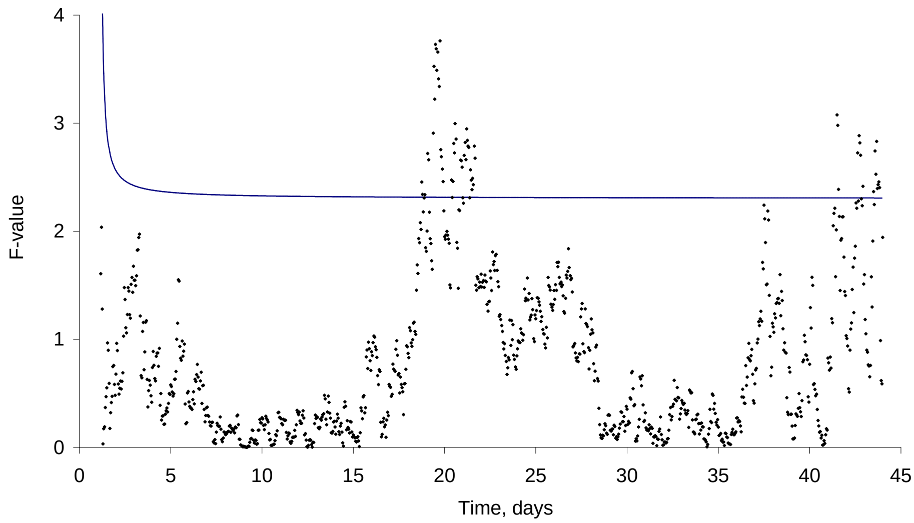


Fig. 3. Test of the significance of harmonic 5 given that all other harmonics are present. F-values together with the 90% probability limit (the solid line) are shown. The harmonic component proves only to be significant at the 90 % level in 59 of the 1070 time steps.

Table 1
The first and second rows show the number (N^*) and the percentage of the 1026 time steps where the individual harmonic components are significant at the 90% probability level. The final row shows the average F-values for each model. The results can be interpreted as: the significance each harmonic given all other harmonics are present. It is clear that the first 3 harmonics are, by far, the most important ones.

	Harmonics											
	1	2	3	4	5	6	7	8	9	10	11	12
N^*	1026	949	971	115	59	27	16	5	13	22	35	0
% Sign.	100	92.5	94.6	11.2	5.8	2.6	1.6	0.5	1.3	2.1	3.4	0
F	418	35.8	22.5	4.0	2.8	3.3	3.0	1.6	1.3	1.2	1.5	0.5

are autocorrelated. Except sequences around day 20 and day 44, the F-values are below the 90 % probability limit. The same tendency is seen for some of the other harmonics as well. The numbers of time steps in which the individual harmonics were significant are listed in Table 1. Harmonics 1, 2 and 3 are significant in nearly all the time steps, clearly indicating dominance in the description of the daily cycles. Harmonic 4 is significant in 115 of the 1070 time steps, corresponding to 11.2 % of the time. Harmonics 5 - 12 were found to be significant in 0 - 5.8 % of the time steps, indicating less importance.

The results shown in Table 1 suggest that some of the harmonic components could

Table 2
Mean square error, MSE, calculated for models containing 12,11,10, . . . , 1 harmonics. The MSE decreases until 3 harmonics remain.

	Harmonics											
	1	2	3	4	5	6	7	8	9	10	11	12
MSE	47.97	46.4	45.21	44.34	43.03	41.98	41.34	40.55	40.02	39.99	55.96	82.7

be removed from the model without reducing the predictive performance. The question is how to evaluate/compare the predictive performance among different (reduced) models. One approach could be to look at the likelihood function, another to use the Akaike information criteria (Akaike, 1974). But for an informal comparison of alternative models the mean square error, or MSE, is determined to suffice as a measure of model performance.

In Section 3.6, the full model (with 12 harmonic components) was optimized (by minimizing MSE) with respect to the two discount factors. The same procedure has been repeated with the reduced models, starting with the full model minus harmonic 12 and then in every step skipping the harmonic component with the highest phase, ending with a model that contains only the trend component and harmonic 1.
The MSE, minimized with respect to the two discount factors, is shown for each model in Table 2. The mean square error decreases as the model is reduced, until there are 3 harmonics left. It could be discussed whether the 4th harmonic should be included, as the model performance is almost equal with 3 and with 4 harmonics. Including the 4th harmonic makes the model more flexible to changes in the the daily drinking pattern, but as the rest of this analysis is based on the reference data set only, we will concentrate on the model with 3 harmonics.

4.2 Model analysis

In this section we will examine how the reduced model fits with the reference data set. The exact specification of the reduced model is

$$\mathbf{F}' = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$

and an 8×8 system matrix

$$\mathbf{G} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \cos(\omega) & \sin(\omega) & 0 & 0 & 0 & 0 \\ 0 & 0 & -\sin(\omega) & \cos(\omega) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cos(2\omega) & \sin(2\omega) & 0 & 0 \\ 0 & 0 & 0 & 0 & -\sin(2\omega) & \cos(2\omega) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cos(3\omega) & \sin(3\omega) \\ 0 & 0 & 0 & 0 & 0 & 0 & -\sin(3\omega) & \cos(3\omega) \end{pmatrix}$$

where $\omega = 2\pi/24$.

Since the DLM is constructed by combining several component DLMs based on the principle of superposition, each of the model components can be identified by model decomposition. So, once the model has been applied in forecasting the water consumption data, the individual components of the DLM can be analyzed. Recall that the 8 elements of the state vector $\boldsymbol{\theta}_t = (\theta_{t,1}, \ldots, \theta_{t,8})$ each has a special meaning. $\theta_{t,1}$ is the underlying, non-cyclic level of the series at time t. $\theta_{t,2}$ is the non-cyclic growth between time $t-1$ and t, whereas $\theta_{t,3}, \theta_{t,4}, \ldots, \theta_{t,8}$ are the Fourier coefficients describing the cyclic pattern. Harmonic 1 describes the overall daily cycle, whereas harmonics 2 and 3 describe minor deviations from the smooth curve during the day.

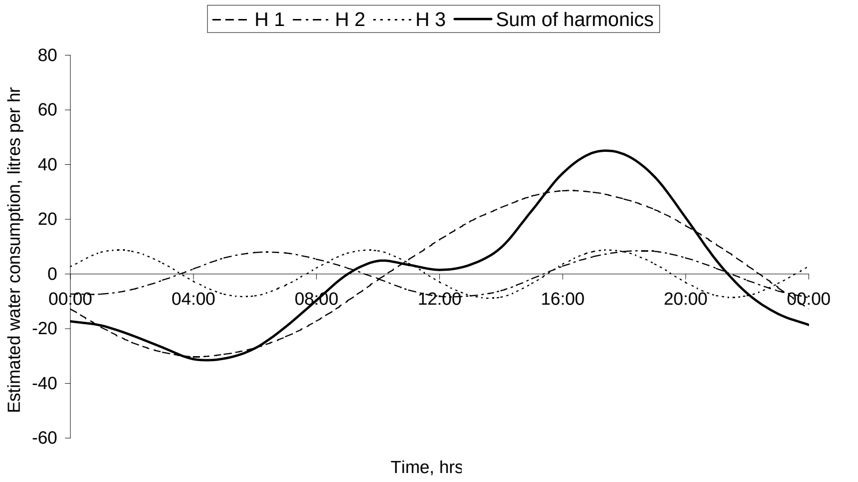
The decomposition of the DLM is made by choosing an appropriate vector $\mathbf{x}$. Let $x_t = \mathbf{x}'\boldsymbol{\theta}_t$ then

$$(x_t|D_t) \sim T_{n_t}\left[\mathbf{x}'\mathbf{m_t}, \mathbf{x}'\mathbf{C_t}\mathbf{x}\right]$$

Choosing $\mathbf{x} = (1,0,0,0,0,0,0,0)$ gives $x_t = \theta_{t,1}$, which is the level component, and $\mathbf{x} = (0,0,1,0,0,0,0,0)$ yields the effect of harmonic 1. Other parameters can be referenced by choosing the appropriate value of $\mathbf{x}$.

Parameter estimates from a 24 hour period (day 22 of the reference data set) are shown in Figure 4. From 4 (a) it clearly appears that harmonic 1 (H1) describes the overall daily pattern, whereas harmonics 2 and 3 (H2 and H3) add minor contributions to the model. It should be noted that the effect of the harmonics sum to zero over a full 24-hour period, as the cyclic part of the model describes deviations from the level component $\theta_{t,1}$. Figure 4 (b) shows the effect of the full model, that is, the sum of the level and the cyclic components. The fitted model (the solid line) and observations from day 22 of the reference data set are shown in Figure 5

(a)



(b)

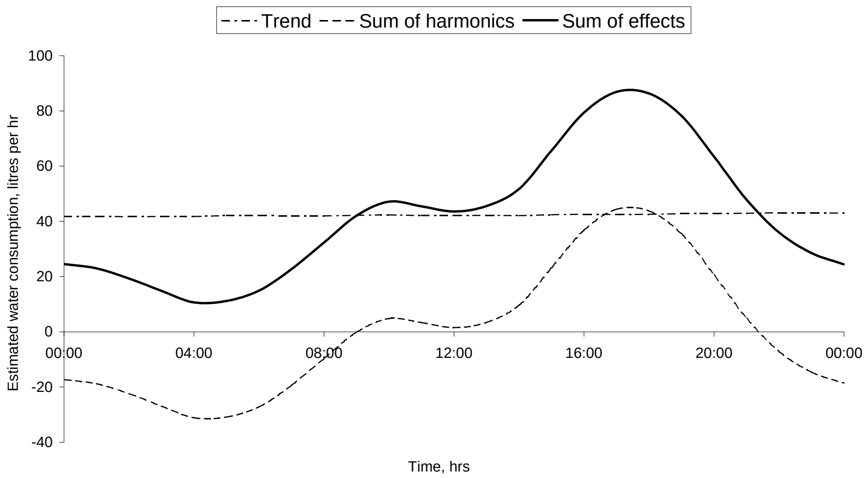


Fig. 4. (a) Estimated model parameters based on observations from the reference data set on day 22. The sum of the 3 harmonic components (H 1, H 2 and H 3) forms the cyclic pattern, indicated by the solid line. (b) The sum of cyclic components and the level component forms the model prediction (cf Figure 5).

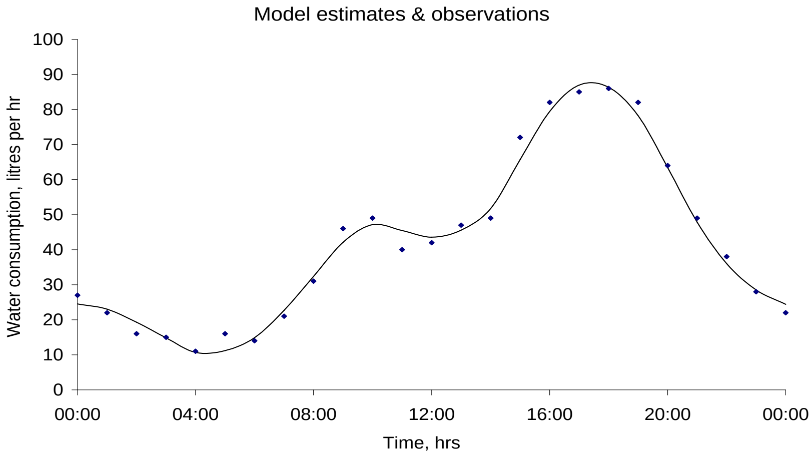


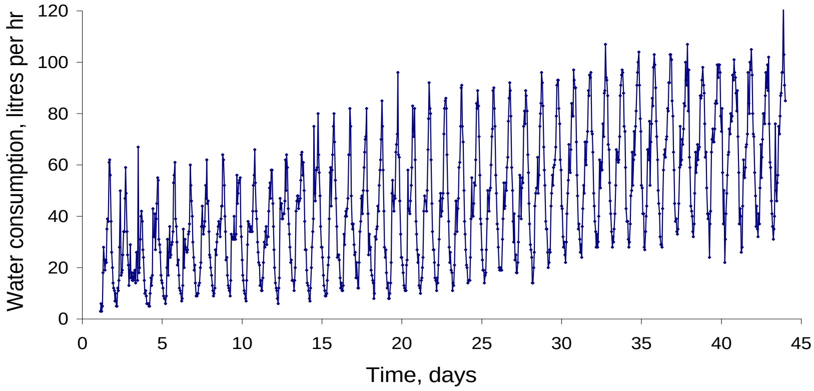
Fig. 5. An example of observed water consumption together with the fitted model, indicated by the solid line (day 22 from the reference data set).

4.3 Visual inspection of model components

A plot of the raw data together with a plot of $\theta_{t,1}$ is shown in Figure 6. The term $\theta_{t,1}$ describes the underlying, deseasonalised level of water consumption which is well suited for monitoring purposes (it should be mentioned that an almost identical curve can be produced by applying a 24-hour moving average to the raw data). A systematic change in the pigs' drinking behaviour, which might be hidden in the raw data, immediately shows up by a visual inspection of $\theta_{t,1}$. For example, Figure 6 (b) shows a clear drop in water consumption on day 4, which is difficult to detect in Figure 6 (a).

The effect of the 3 harmonic components is shown in Figure 7. Harmonic 1 (H1) evolves rather stable with increasing amplitude until somewhere around day 20 when it stabilizes, whereas harmonics 2 and 3 (H2 and H3) appear to be more fluctuating. The effect of harmonic 2 decreases around day 15. Harmonic 1 seems to compensate for that by a short increase, which indicates a brief change in the daily drinking pattern, without any significant change in the underlying level ($\theta_{t,1}$). The contribution from harmonic 3 is decreasing at the end of the time period and the 24-hour cyclic pattern tends to be described mainly by harmonics 1 and 2.

(a)



(b)

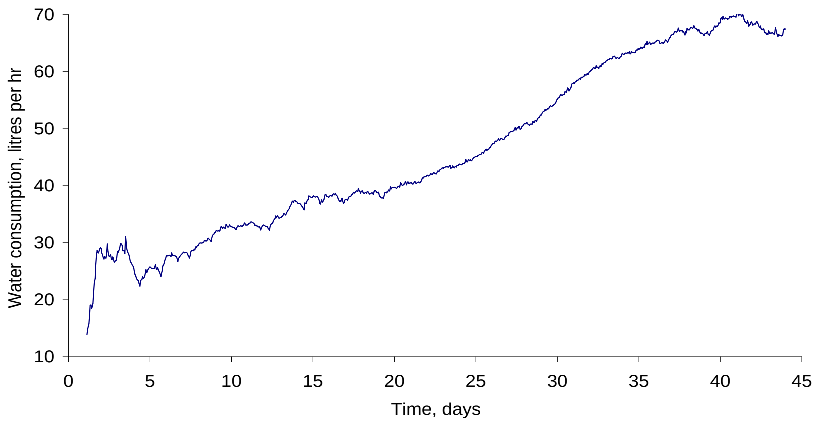


Fig. 6. (a) The raw one-hour observations of water consumption from a group of young pigs. (b) The level component from the reduced model with 3 harmonics.

4.4 Model check

Error analysis can be performed on fitted errors by measuring the retrospective departure of the data from the model (see, e.g. Box and Jenkins, 1976). But since the interest is in the predictive fit of the model, embodied in the sequence of one-step ahead forecast distributions, the one-step forecast errors e_t will be used for model assessment. The analysis will be based on an examination of the forecast errors e_t from the reference data set. Under the assumptions of the model, the forecast

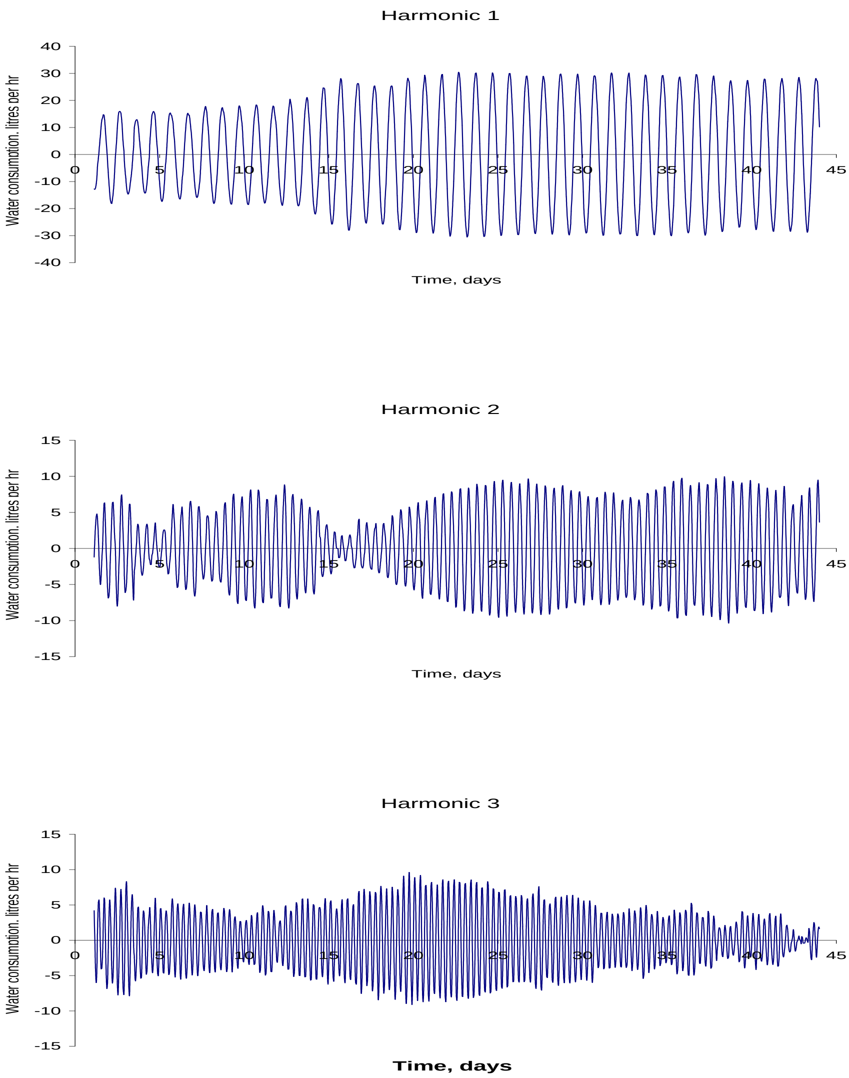


Fig. 7. The 3 harmonic components from the reduced model. It is clear that the first harmonic describes a great part of the basic daily pattern, whereas harmonics 2 and 3 handle the minor changes in the cycles.

distributions are $(Y_t|D_{t-1}) \sim T_{n_{t-1}}[f_t, Q_t]$, so the random error term e_t has the predictive distribution

$$(e_t|D_{t-1}) \sim T_{n_{t-1}}[0, Q_t].$$

Observed deviations in the error sequence away from the predicted behaviour are indicative of either irregularities in the data series or model inadequacies. For the analysis, the raw errors e_t are standardized with respect to the variance of the forecast

$$e_t^* = e_t/\sqrt{Q_t}, \tag{11}$$

which is distributed as

$$(e_t^*|D_{t-1}) \sim T_{n_{t-1}}[0, 1], \quad (t = 1, 2, \ldots).$$

Figure 8 (a) shows a time series plot of the standardized forecast errors. It does not

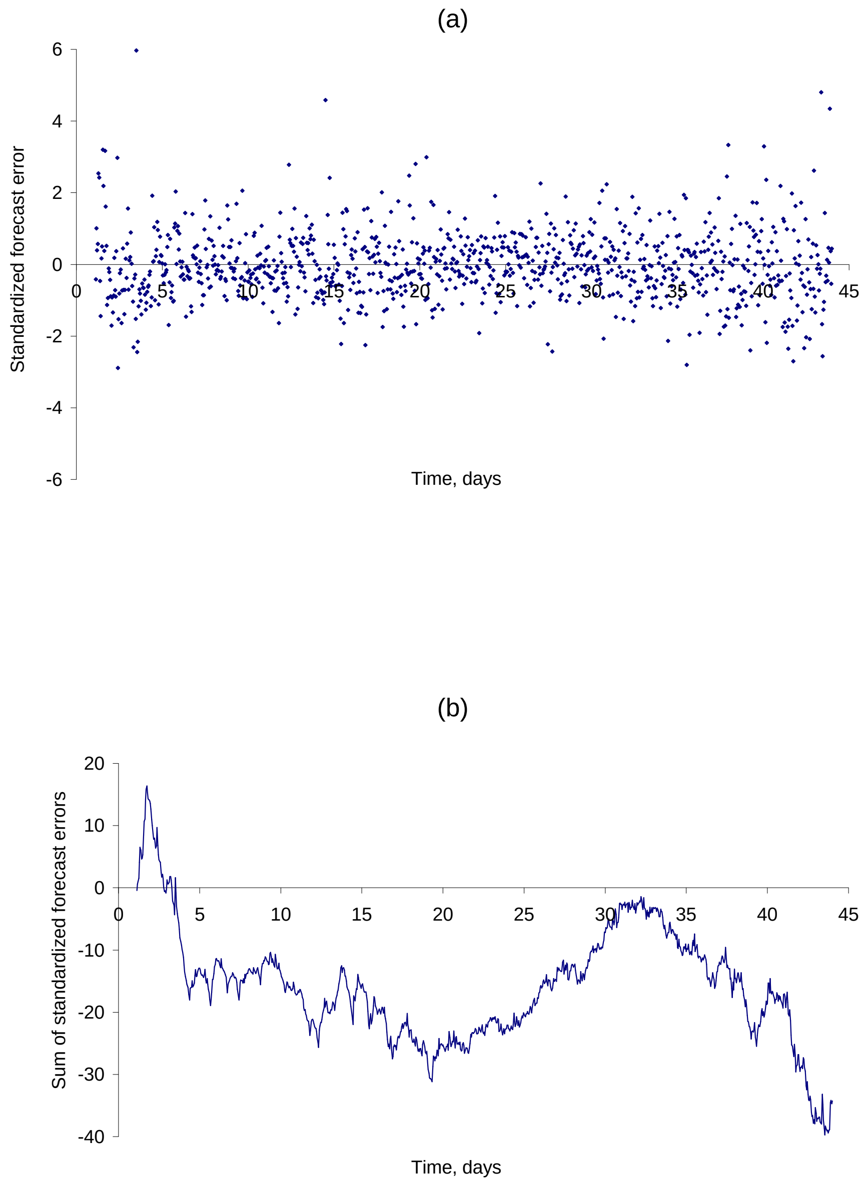


Fig. 8. (a) Time series plot of the standardized forecast errors. The plot does not indicate any changes over time. (b) Sum of the standardized forecast errors.

indicate any changes over time.

The forecast errors are expected to be mutually independent. This is probably the case most of the time, when the pigs' drinking pattern is natural without any influence of diseases or deviating quality in the feed. But in situations with disturbances, e.g. outbreak of a disease that affects the pigs' water consumption, there will be sequences where the model does not match the observed data (refer to Madsen and Kristensen, 2005, for a graphical illustration). This phenomenon will imply correlated forecast errors in the number of time steps the model needs to adapt to the new pattern. For identification of correlated errors it can be helpful to look at a plot of the cumulative sum

$$C_t = \sum_{i=1}^{t} e_i^* \tag{12}$$

against the time t. Such a plot is shown in Figure 8 (b). The plot indicates strong correlation in residuals within the first 5 days, which is probably caused by the abnormal or non-stable drinking behaviour right after the pigs are weaned, as it appears from Figure 6. There is a positive drift in the sum of errors from day 20 to 32, and from day 35 the sum is declining again. In the same interval the level curve (Figure 6 (b)) tends to be S-shaped with increasing slope in the interval with positive forecast errors and decreasing slope in the interval with negative errors.

This could indicate that the correlation in forecast errors is an effect of the model structure only allowing for linear growth from time $t-1$ to t. An obvious solution to that problem could be to use a quadratic growth model, which is identical to the model described above, except for the addition of a growth term to the trend-part of the model. With the same notation as in equation 5 the quadratic growth model is described as:

$$\begin{aligned} Y_t &= \mu_t + v_t \\ \mu_t &= \mu_{t-1} + \beta_t + w_{t,1} \\ \beta_t &= \beta_{t-1} + \gamma_t + w_{t,2} \\ \gamma_t &= \gamma_{t-1} + w_{t,3} \end{aligned} \tag{13}$$

where the quantities μ_t, β_t and γ_t, respectively, represent level, growth and change in growth at time t. Compared to the analogue model for the continuous time scale, β_t represents the first derivative with respect to time, and γ_t the second derivative of the expected level.

The quadratic growth model has been combined with the cyclic model (with 3 harmonics) by the same principles as described in Section 3.2. To ensure optimal flexibility, γ_t was given its own discount factor. The setting of the 3 discount factors were optimized by minimizing MSE in the same way as described in Section 3.6. Optimal choice of discount factors were $\delta_{level} = 0.99$, $\delta_{growth} = 0.99$ and $\delta_{harmonics} = 0.96$. The cumulative sum of forecast errors from the quadratic growth model is shown in Figure 9, from which it is appears that the model is more adaptive to changes in growth rate. The quadratic growth model in particular seems to handle the irregular drinking behaviour during the first 5 days better than the model with only a linear growth term.

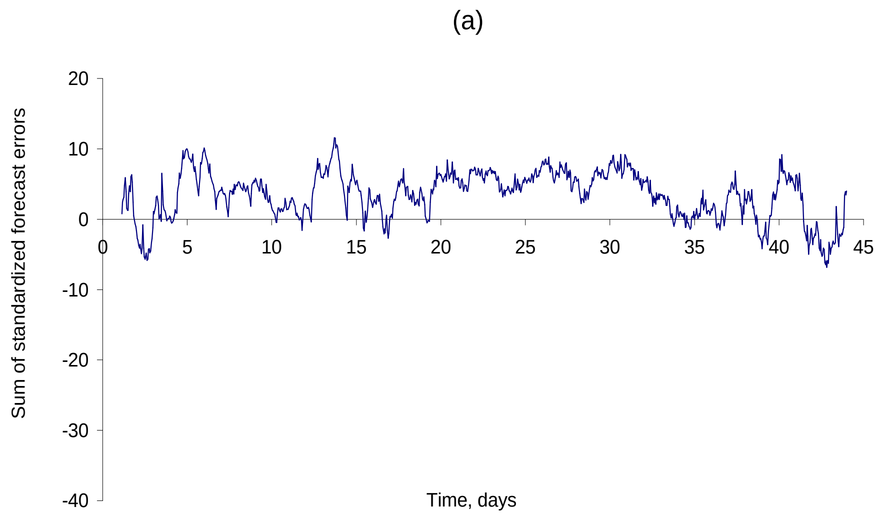


Fig. 9. Sum of standardized forecast errors from the quadratic model.

Figure 10 (a) shows a histogram of standardized forecast errors from the quadratic growth model. The solid line is the probability density function of the fitted student t distribution. It is seen that there is a tendency for the observations to lie slightly to the left of the expected mean, 0. The median is found to be -0.10. A log-transformation of data did not reduce the skewness in the histogram significantly.

In Figure 10 (b) the quantile-quantile plot of the observed standardized errors against the fitted student t distribution is shown. Here it is seen that the major part of the observations are close to a straight line, although at both ends there are observations that lie far from the straight line. These can be caused by outliers or by model collapse in time intervals where the pigs suddenly change their drinking behaviour.

5 Discussion

A Dynamic Linear Model including cyclic components has proven to be well suited for modelling the water consumption data from a batch of young growing pigs. The combination of a growth model and a cyclic model seems to fit the characteristics of the growing pigs' diurnal drinking pattern, and by decomposition, it furthermore allows monitoring of the trend and the cyclic pattern individually.

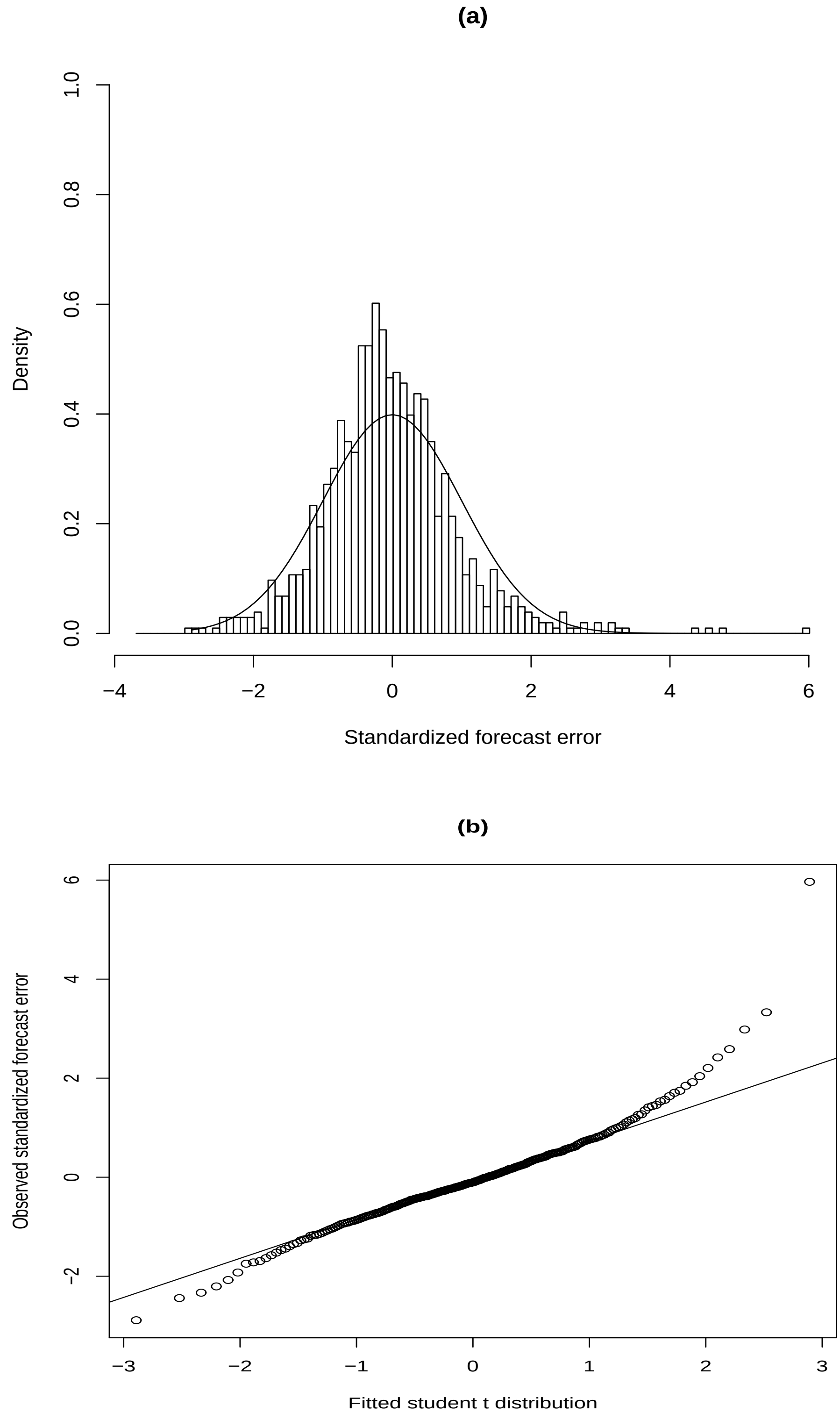


Fig. 10. (a) Histogram of standardized forecast errors. The solid line is the probability density function of the fitted student t distribution. (b) Quantile-quantile plot for the observed standardized forecast errors against the fitted student t distribution.

Despite a lot of effort, it was not possible to get any reasonable estimates of the variance components of the model. One of the reasons for that could be that the assumption that both the observational and the system variance are constant might be wrong. Instead a model that describes the evolution in the system variance by discount factors was chosen. The model with discount factors is possibly more durable and robust, in case of nonconstant system variance $\mathbf{W}$, because $\mathbf{W}$ is defined as a relative proportion of the state vector variance, thereby allowing the variance matrix to vary with time. Furthermore, even if the assumption about constant variance components should be right and their values were known, there seems to be very little lost with the discount model in terms of a potentially improved description. For example, Harrison (1967) showed that, considering the second-order polynomial DLM, the maximum loss in one-step-ahead prediction for typical general settings is an increase in standard deviation of less than 1%.

Optimal discount factors were chosen by minimizing the mean square of the forecast error. The relatively high optimal settings of the discount factors (0.97-0.99) indicate that the model is in good agreement with the structure of the reference data set.

Error analysis was used to verify the model. Strong serial correlation in the standardized step-ahead forecast errors indicated that the second-order polynomial submodel, describing the growth in the overall level, is not flexible enough. The model only allows for linear growth, which makes it difficult to adapt to data in periods with non-linear growth. The growth model was replaced by a third-order polynomial model which allows for quadratic growth. This improved model performance, i.e. it reduced the serial correlation in the standardized forecast errors.

The slope of the overall level curve might, to some extent, be interpreted as a mark of the growth rate since the growth rate and feed consumption as well as feed- and water consumption are correlated. Experiments have proved that as much as 75% of the pigs' daily water intake is closely associated with eating bouts (Bigelow and Houpt, 1988). Therefore one would expect the overall water curve to follow some kind of exponential progress as is the case with the growth curve, see, e.g. Whittemore (1993). For that reason it seems reasonable to recommend the quadratic growth model when modelling the water consumption of growing pigs. On the other hand, if the model is used as part of a monitoring system, as is the case in Madsen and Kristensen (2005), it might be desirable that the system react to changes in growth rate of water consumption. One way of doing this is to assume only linear growth in the model, and then react when a series of serial correlated forecast errors indicates that the model is inadequate because of changing growth rate. This issue is discussed further in Madsen and Kristensen (2005).

Besides monitoring the forecast errors, visual inspection of model components can provide valuable information to the manager. Monitoring the trend or level curve gives a comprehensive view of how the production progresses.

Finally it should be noted that the model only has been verified by a single data set and that the results concerning the optimal number of harmonics and the setting of the discount factors should be interpreted with care. An analysis based on the complete data set described in Section 2 can be found in Madsen and Kristensen (2005).

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Transaction on Automatic Control AC -19, 716 –723.

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.

Box, G., Jenkins, G., 1976. Time series analysis: Forecasting and Control, 2nd Edition. Holden-Day, San Francisco.

Cook, D. F., Massey, J. G., McKinney, C., 1992. A knowledge-based approach to statistical process control. Computers and Electronics in Agriculture 7, 13–22.

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.

Goodall, E. A., Sprevak, D., 1985. A bayesian estimation of the lactation curve of a dairy cow. Animal Production 40, 189–193.

Harrison, P., 1967. Exponential smoothing and short-term forecasting. Man. Sci. 13, 821–842.

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.

Kai, P., Madsen, T., Pedersen, B., Ruby, V., Boykel, I., Nielsen, N., 1999. Elektronisk overvågning af produktionen i stalde med alt-ind alt-ud produktion. Repport 420, The National Committee for Pig Production, Danish Bacon & Meat Council.

Kalman, R., 1960. A new approach to linear filtering and prediction problems. Journal of Basic Engineering 82.

Lundbye-Christensen, S., 1991. A multivariate growth curve model for pregnancy. Biometrics 47, 637–657.

Lundbye-Christensen, S., Christensen, M., 1995. Modelling the growth of human placental lactogen in maternal serum during late pregnancy. Revista Brasileira de Probabilidade e Estatistica 9, 1–24.

Madsen, T. N., 2001. Tools for monitoring growing pigs. Dina research report no. 91, The Dina Research School, The Royal Veterinary and Agricultural University.

Madsen, T. N., Kristensen, A. R., 2005. A model for monitoring the condition of young pigs by their drinking behaviour. Computers and Electronics in Agriculture, Submitted.

Nienaber, J. A., McDonald, T. P., Hahn, G. L., Chen, Y. R., November-December 1990. Eating dynamics of growing-finishing swine. Transactions of the ASAE 33 (6), 2011–2018.

Pole, A., West, M., Harrison, J., 1994. Applied Bayesian Forecasting and Time Series Analysis. Chapman & Hall, New York.

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.

Thysen, I., 1992. Detection of changes in daily milk production by a multi-state kalman filter. Dina Rapport No. 3.

Thysen, I., 1993. Monitoring bulk somatic cell counts by a multi-process kalman filter. Acta Agriculturæ Scandinavica, Section A, Animal Science 43, 58–64.

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

Whittemore, C., 1993. The Science and Practice of Pig Production. Longman Scientific & Technical.