Original papers

Spatial modeling of pigs’ drinking patterns as an alarm reducing method I. Developing a multivariate dynamic linear model

K.N. Dominiak\textsuperscript{a,b,}, L.J. Pedersen\textsuperscript{b}, A.R. Kristensen\textsuperscript{a}

\textsuperscript{a} Faculty of Health and Medical Sciences, Department of Veterinary and Animal Sciences, University of Copenhagen, Grønneårdvej 2, 1870 Frederiksberg C, Denmark
\textsuperscript{b} Department of Animal Science, Aarhus University, Blichers Allé 20, 8830 Tjele, Denmark

\textbf{A R T I C L E   I N F O}

Keywords:
Spatial model
Early warning
Growing pigs
Simultaneous monitoring
Water consumption

\textbf{A B S T R A C T}

The overall objective of this paper is to present the development of a spatial multivariate dynamic linear model (DLM) modeling the water consumption of growing pigs throughout the entire growth periods. The water consumption from multiple pens in multiple sections are monitored simultaneously by flow meters in both a commercial herd of finisher pigs (30–110 kg) and a research facility herd of weaner pigs (7–30 kg). The diurnal drinking patterns are modeled by a multivariate DLM, which is superpositioned by four sub-models describing three harmonic waves and a growth trend. The overall hypothesis of this paper is that pens and sections in a herd of growing pigs are correlated, and that this correlation can be modeled using model parameters defined at different spatial levels. Therefore seven model versions are defined to reflect a variety of temporal correlation structures between the monitored drinking patterns. The model versions were trained on learning data of the two herds, and run on separate test data sets from the herds. Their ability to fit the test data is measured as mean square error (MSE). Results for the finisher herd indicate that drinking patterns from pens within the same section are correlated (MSE = 13.850). For the weaner herd, results indicate an inverse relation between the degree of correlation and the model fit. Thus, the best fit (MSE = 1.446) is found for the model version expressing least correlation in data from pens across the herd.

1. Introduction

The everyday focus in livestock production is to ensure a profitable production without compromising animal welfare. Over the years, livestock production has been subjected to an increasing industrialization, which has led to larger, centralized production units with less time available for attending the individual animal.

Sensor-based monitoring and early warning systems can aid the daily manager to identify individual animals, or groups of animals, which need high priority attention. Ideally the system can generate a warnings timely enough for the manager to decide for the right intervention and either prevent any welfare reducing condition from progressing, or at least reduce its consequences (Kristensen et al., 2010). Early warning systems, or detection models, for livestock production have been developed for the past twenty years (Dominiak and Kristensen, 2017), and they often aim to detect different conditions in individual animals like Clinical Mastitis (CM) in cows de Mol et al. (1997), Huybrechts et al. (2014)), lameness (Maertens et al., 2011; Garcia et al., 2014) and oestrus (Bressers et al., 1995; Ostensen et al., 2010). Furthermore the modeling of changes in animal behavior monitored by sensors, has had an increased focus within the field of precision livestock farming over the recent years. Modeling animal behavior has for instance been used as early indicators of diseases (Zhuang et al., 2018), unwanted events (Oczak et al., 2013), or as a decision tool for managing groups of growing pigs (Brown-Brandl et al., 2013).

It is, however, a general problem for sensor-based early warning systems that they generate too many false alarms. False alarms are costly and time consuming for the farmer, and excess alarms devaluate the managerial value, and diminish the trust in the warning system (Berckmans, 2014; Dominiak, 2017).

The primary cause for the excess number of false alarms is insufficient detection performance by detection models developed in a twenty-years period between 1995 and 2015, as discussed in the two review papers by Dominiak and Kristensen (2017), Hogevæn et al. (2010). The overall difficulties in obtaining sufficiently high detection performances indicate that the number of alarms cannot be reduced through high performances of the detection models alone. One way of reducing the number of alarms is to incorporate a post-processing method in the algorithm for the full detection system (Dominiak, 2017).
Hereby the alarms get sorted, or prioritized, and reduced before they are communicated to the farmer.

Within the field of livestock precision farming, only three papers, written by Aparna et al. (2014), Steeneveld et al. (2010), de Mol and Woldt (2001), describe methods aimed specifically at reducing the number of false alarms through post-processing methods. Both de Mol and Woldt (2001) and Steeneveld et al. (2010) combine alarms from sensor-based detection models with cow-specific non-sensor information to reduce the number of false alarms, whereas Aparna et al. (2014) prioritize alarms along a time-gradient by predicting the onset of farrowing based on well-defined behavioural phases leading up to the event.

An alternative approach for reducing alarms is to relate alarms to specific areas of the herd using a spatial model. Such a spatial detection model aims to identify specific high-risk areas within the herd, rather than target individual animals. Area-specific alarms enable the manager to include any specific knowledge of the animals in the targeted areas, and hereby choose the best suited intervention under the given circumstances.

For bio-security reasons, modern Danish pig production units for growing pigs are run very disciplined and systematically with a clear spatial separation between pigs of different age groups in closed sections (Cameron, 2000; Danish Agriculture and Food Council, 2010). This separation restricts most diseases from spreading between sections in a herd and, to a certain extent, between pens in a section (Cameron, 2000; Pedersen, 2012; Vils, 2013).

From a modeling perspective, such a construction of the production unit makes it well suited for the development of a spatial model. Hence, the herd can be modeled as a system consisting of one large unit (the whole herd), which consists of a number of identical subunits (sections), with each subunit consisting of a number of identical sub-subunits (pens).

The parameter used in the model must contains relevant information on all animals across the herd in order to reflect the entire modeled system. Madsen et al. (2005) modeled the drinking pattern of a whole section of weaner pigs, and found that changes in the pattern held information on the general wellbeing of the pigs as well as having predictive value for detecting outbreaks of diarrhea. Later Andersen et al. (2016) showed that changes in drinking patterns could indicate stress caused by a variety of factors like stocking density and amount of rooting material supplied. These studies indicate a high level of information in water data, and this is supported in a recent study, where Jensen et al. (2017) found unexpected changes in the pigs’ water consumption to be the one single parameter containing most information in the prediction of outbreaks of either diarrhea or fouling in a pen with finisher pigs.

Previous modeling of water data from growing pigs has been done on individual pens (Kashiha et al., 2013; Andersen et al., 2016; Jensen et al., 2017) or on the total water consumption in a section (Madsen et al., 2005). By modeling pens or sections separately, each modeled unit is considered isolated from other parts of the herd.

On the other hand, the spatial detection model presented in this paper is characterized by the incorporation of correlations between simultaneously monitored water data from multiple pens in multiple sections in a herd. Hereby interactions across the herd are reflected, and the model is able to detect and locate systematic changes in drinking patterns in specific areas (pens or sections) of the herd.

Both Jensen et al. (2017), Madsen et al. (2005) have shown that systematic changes in drinking patterns can be used to predict outbreaks of unwanted events amongst growing pigs. Predictions based on the spatial model developed in this paper, will be communicated as area-specific alarms to the manager, and are able to point out which pens or sections are of higher risk of an outbreak of unwanted events. Such area-specific alarms enable the manager to focus the managerial attention, and choose the right interventions for the pointed area, whereby the consequences of the unwanted events can be reduced or prevented. The application of the spatial model is, however, the objective of a following paper (Dominiak et al., submitted for publication), and will not be discussed further here.

Thus, the objective of this paper is to present the development of a new, spatial approach for modeling drinking pattern of growing pigs throughout the entire growing period using a multivariate dynamic linear model. It is our hypothesis that pens and sections in a herd of growing pigs are correlated, and that this correlation can be modeled using model parameters defined at different spatial levels.

2. Materials and methods

2.1. Herds, sensors and data

2.1.1. Herd description

For this study, water consumption data was obtained from two different herds. Herd A is a Danish commercial finisher herd, and Herd B is an experimental weaner herd, “Granhøj”, owned by the Danish Pig Research centre.

The general routines in Danish weaner and finisher production are structured so that the time of insertions of pigs in the farm, and the length of the growth period run in a cycle (Fig. 1). Such a production
cycle is a part of a larger production plan coordinated with the suppliers of the incoming pigs and the abattoir, when regarding finishers, or buyers, when regarding weaners. All pigs in one section are inserted at the same day, and they are all of same age relative to weaning date. When a section is emptied, it is cleaned and dried out for bio-security reasons before a new batch of pigs are inserted. For Herd A one growth period (30–110 kg) is approximately 14 weeks including one week of cleaning (Fig. 1(A)), and for Herd B one growth period (7–30 kg) is 8 weeks including four days of cleaning (Fig. 1(B)).

Herd A produces 10,000 cross-bred finisher pigs per year, and the herd has five identical sections, of which four are included in this study (Fig. 2(A)). Each section consists of 28 pens, and two neighbouring pens share the same water pipe, which supplies one drinking nipple in each of the pens (see Fig. 3). One pen per section is left empty at insertion, and is used as hospital-pen during the growth period. Hence, approximately 486 pigs are inserted in a section with 18 pigs in each pen, and they are fed with liquid feed three times a day (Krogsdahl, 2014b). From 60 kg bodyweight the pigs are fed restrictively as it is common practice with finisher pigs in order to increase the lean meat percentage (Vils, 2012).

Herd B consists of four sections, each with 12 pens for weaner pigs (Fig. 2(B)). One water pipe supplies one drinking bowl per pen (Fig. 3). 15 pigs are inserted in each pen, and the pigs are fed ad libitum with dry feed three times a day during the whole growth period (Krogsdahl, 2014a).

The main characteristics of the two herds are summarized in Table 1.
Table 1
Characteristics for the two herds in the study.

| Characteristic          | Herd A           | Herd B           |
|-------------------------|------------------|------------------|
| Production type         | Commercial       | Research Farm    |
| Animal group            | Finishers (30–110 kg) | Weaners (7–30 kg) |
| Sections                | 4                | 4                |
| Sensors total/ section  | 8/2              | 16/4             |
| Pigs per pen/ section   | 18/36            | 15/15            |
| Growth period (batch)   | 14 weeks         | 8 weeks          |
| Batches per sensor      | 7                | 13 *             |
| Learning data (hours)   | 9540             | 14657            |
| Test data (hours)       | 4441             | 3025             |

* 14 for Section 4.

2.1.2. Data

Water data was obtained by photo-electric flow sensors (RS V8189 15 mm Diameter Pipe) measuring water flow per millisecond as pulses proportional to the velocity of the water (Anonymous, 2000). The sensors were calibrated between batches, and the number of pulses entered a central data base once every 24 h. For this study the number of pulses were converted to litres and aggregated per hour, yielding water use in litres per hour.

In Herd A a total of eight sensors were installed with two sensors placed in each of four identical sections (Fig. 2 (Herd A)). All sensors were placed on water pipes supplying two neighbouring pens, and therefore each sensor monitored the joint water use of pigs in two pens. Both sections and pens were randomly chosen, and seven batches were monitored per section from May 2014 to March 2016. In this study a batch of pigs is defined as all pigs inserted in the same section at the same day.

The full data set for Herd A consists of eight time series, one per sensor, of length from the first observation in the herd to the last observation in the herd. In total, 16,309 h. Every observation from each sensor is paired with the insertion date of the relevant batch of pigs at any given time.

In Herd B a total of sixteen sensors were installed with four sensors in each of four identical sections (Fig. 2 (Herd B)). Each sensor monitored the water use of one individual pen. The sections included in this study were assigned by the research centre, whereas the pens within each section were randomly chosen. 13 batches (Sections 1, 2 batches (Section 4) were included, and data was collected from October 2014 to December 2016.

The full data set for Herd B consists of sixteen time series, one per sensor. The monitoring period begins with the first global observation and ends with the final global observation. In total, 18,755 h. As in the data set for Herd A, every observation is paired with the insertion date of the relevant batch.

The data sets from both herds are divided into learning data sets and test data sets. The two final batches from each herd are defined as test data sets, and all prior batches are defined as learning data sets. One full batch between learning and test data is left out for each sensor from each herd in order to avoid any observations from the same pigs occurring in both data subsets. This division is chosen in order to provide the larger amount of data for the model learning as recommended by Witten and Frank (2005), while recognizing that test data sets consisting of only one batch may be too small.

The number of batches differs between the two herds (7 batches Herd B, 13/14 batches Herd A), as does the length of each batch (14 weeks Herd A, 8 weeks Herd B). This is reflected as differences in the proportions between training and test data from the two herds (see Table 1). Hence, the learning data set for Herd A consists of four batches (9540 h, 68%), the test data set consists of two batches (4441 h, 32%), and one batch per pen (2328 h) between the two data subsets is left out. On the other hand, the learning data set for Herd B consists of ten batches (14,657 h, 83%), the test data set consists of two batches (3025 h, 17 %), and one batch per pen (1073 h) between the two data subsets is left out.

During cleaning periods between batches, no sensor observations were made. Such periods were considered planned periods of missing data as opposite to any occasional missing observations or sensor outages during the growth periods.

As only actual water flow is measured, it is not possible to distinguish periods with no water consumption from (short) sensor outages. Since water consumption is typically very low during the night, it was decided to interpret missing observations of a duration of less than 5 h between 10:00 PM and 4:00 AM as zero observations. All other missing observations are considered as sensor outages.

2.2. Model description

In this section the structure of the developed model is described. The model is developed as a general tool, which in theory can be applied to any herd with either weaner pigs or finisher pigs. Nomenclature and abbreviations used in the model description are defined or described in Table 2.

Table 2
Nomenclature for the model description.

| Symbol/ abbreviation | Name                  | Definition/description |
|----------------------|-----------------------|------------------------|
| $Y_t = (Y_{t1}, \ldots, Y_{tn})'$ | Observation vector   | Water consumed within the last hour at time $t$ for each of the $n$ sensors |
| $\theta_0 = (\theta_0, \ldots, \theta_0)'$ | Parameter vector     | Latest variables describing the observations at time $t$ |
| $\hat{Y}_t$          | Observation equation  | $F\hat{Y}_t + \nu_t$ |
| $\hat{\theta}_t$     | System equation      | $G\hat{\theta}_t - \omega_t$ |
| $\nu_t$              | Observation error term| $\sim \mathcal{N}(0, \Sigma_t)$ |
| $\omega_t$           | System error term    | $\sim \mathcal{N}(0, \Sigma_t)$ |
| $F_t$                | Design matrix        | Dimensions $n \times r$ |
| $G_t$                | System matrix        | Dimensions $n \times n$ |
| $\mathcal{V}_t$      | Observation variance-covariance matrix | Dimensions $r \times r$ |
| $W_t$                | System variance-covariance matrix | Dimensions $n \times n$ |
| $\lambda$            | Mean of the one-step-ahead forecast for $Y_{t+1}$ given all observations until time $t$ | The forecast is provided by the Kalman filter $^*$ |
| $e_t$                | Vector of forecast errors | $Y_t - \hat{Y}_t$ |
| MSE                  | Mean Square Error    | $\sum_{t=1}^T e_t^2$ |
| $H_1$                | Cyclic model 24 h    | See Eq. (3) |
| $H_2$                | Cyclic model 12 h    | See Eq. (3) |
| $H_3$                | Cyclic model 8 h     | See Eq. (3) |
| $L_G$                | Linear growth sub-model | See Eq. (4) |
| $H$                  | Herd level           | All pens with sensors in the herd |
| $S$                  | Section level        | All pens with sensors within the same section |
| $P$                  | Pen level            | A pen with a sensor |

* See West and Harrison (1999)

2.2.1. General dynamic linear model

The water consumption over time is modeled simultaneously for all sensors in the herd. The observation vector $Y_t = (Y_{t1}, \ldots, Y_{tn})'$ is the water consumed within the last hour at time $t$ for each of the $n$ sensors. It is modeled by the matrix quadruple $F_t, G_t, \mathcal{V}_t, W_t$ where, following the description by West and Harrison (1999):

- $F_t$ is a known $(n \times r)$ design matrix;
- $G_t$ is a known $(n \times n)$ system matrix;
- $\mathcal{V}_t$ is a known $(r \times r)$ observation variance-covariance matrix;
- $W_t$ is a known $(n \times n)$ system variance-covariance matrix.
The four matrices, $F$, $G$, $V$, and $W$, define the way $Y_t$ relates to an underlying parameter vector $\theta_t$ at time $t$, and how the system evolves over time in the two equations:

**Observation equation**

$$Y_t = F_t \theta_t + \nu_t, \quad \nu_t \sim \mathcal{N}(0, V).$$

**System equation**

$$\theta_t = G_t \theta_{t-1} + \omega_t, \quad \omega_t \sim \mathcal{N}(0, W).$$

The aim of the dynamic linear model (DLM) is to estimate the parameter vectors $\theta_t$ from the observations $Y_t$ by sequential use of the Kalman filter. Let $D_0$ denote the initial information before any observations are made so that $\theta_{t=1} \sim \mathcal{N}(m_0, C_0)$. Furthermore, let $D_{t=1} = D_0 \cup \{Y_1, ..., Y_{t=1}\}$ denote all available information before time $t$ so that $\theta_{t=1} \sim \mathcal{N}(m_{t=1}, C_{t=1})$. When a new observation $Y_t$ becomes available, the Kalman filter will update the conditional distribution from $\theta_{t=1} \sim \mathcal{N}(m_{t=1}, C_{t=1})$ to $\theta_t \sim \mathcal{N}(m_t, C_t)$ as described by West and Harrison (1999).

2.2.2. Model construction

When looking at Figs. 4 and 5 it can be seen that the water consumption of growing pigs has a clear diurnal pattern. Furthermore, Fig. 6 illustrates how the underlying level of daily water consumption increases as the pigs grow.

Madsen et al. (2005) found that the drinking pattern of a whole section of 405 weaner pigs could be described in a DLM composed of four smaller DLMs, describing three harmonic waves of lengths 24 h, 12 h, and 8 h, and a growth trend. The same four sub models describe the diurnal drinking pattern of a pen of weaners or finishers very good as well, as illustrated in Fig. 7, and the development of the full multivariate model will be described in the following subsections.

2.2.3. Cyclic models

The diurnal drinking pattern is modeled by three cyclic models, each describing a harmonic wave. Harmonic waves can be expressed in a DLM using trigonometric functions in the *Fourier form representation of seasonality* (West and Harrison, 1999; Madsen et al., 2005), where each wave takes up two parameters, representing the phase and amplitude of the cosine waveform. According to West and Harrison (1999), the harmonic waves can be described with the design matrix $F_t^h$ and system matrix $G_t^h$ defined as:

$$F_t^h = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad G_t^h = \begin{pmatrix} \cos(\omega) & \sin(\omega) \\ -\sin(\omega) & \cos(\omega) \end{pmatrix},$$

with $\omega = 2\pi/24$ yielding a wave with a period of 24 ($M_{H1}$), $\omega = 2\pi/12$ a wave with a period of 12 ($M_{H2}$), and $\omega = 2\pi/8$ a wave with a period of 8 ($M_{H3}$).

2.2.4. Linear growth model

The underlying level of water consumption can be described by a linear function, and the increase over time is included by combining the linear function with a growing trend in a linear growth model, modeling the increase from time $t-1$ to $t$.

The general description of a dynamic linear growth model, as based on West and Harrison (1999), is characterized by the following design and system matrices:

$$F_t^g = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad G_t^g = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$  

The parameter vector $\theta_t$ consists of a level parameter $\theta_{t1}$ and a growth parameter $\theta_{t2}$. Thus, the expected level at time $t$, $\theta_{t1}$, will be the sum of the level at time $t-1$ and the growth parameter.

2.3. Full model - univariate

For a single sensor, the univariate model consisting of four sub models; one linear growth model $M_{LG}$ and the three cyclic models $M_{H1}$, $M_{H2}$ and $M_{H3}$, is characterized by the design matrix

$$F_t^u = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}^T \quad \text{and} \quad G_t^u = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$$

and the system matrix

![Pen 5.7 2015-08-11 to 2015-08-17](image)

**Fig. 4.** Diurnal drinking pattern, finishers Herd A.
where $\omega = 2\pi/24$.

The observation variance-covariance matrix reduces in the univariate case to a scalar, $V_{tu}^o$, whereas the size of the system variance-covariance matrix, $W_{tu}^s$, is of size $8 \times 8$. The parameter vector $\Theta_t$ has eight elements; one for level, one for growth, and two for each of the three harmonics.

2.3.1. Full model - multivariate

The simplest possible multivariate model for $n$ sensors would be to define a system matrix $G_t$ of size $8n \times 8n$ as a block diagonal matrix where each of the $n$ blocks along the diagonal is identical to $G_t^s$ from Eq. (6). Similarly, the observation variance-covariance matrix would be a diagonal matrix having all diagonal elements equal to $V_{tu}^o$. The system variance-covariance matrix would be a block diagonal matrix where each block along the diagonal would be equal to $W_{tu}^s$. Finally, the design matrix $F_t$ would be a $8n \times n$ matrix with $n$ blocks each equal to $F_t^o$. The underlying assumption behind such a model would be that the observations from the $n$ sensors were completely independent.

A multivariate model as described would, however, not add anything to a scenario with $n$ univariate models running separately in parallel. Therefore, the model has to be modified to allow for interactions between sensors. This can, basically, be achieved by direct modeling of the interactions in the design and system matrices and/or by estimating full variance-covariance matrices $V_t$ and $W_t$ allowing for correlations between sensors (as opposed to block diagonal matrices).

In this study both approaches will be used. The interactions between the elements of the parameter vector will be directly modeled in the design and system matrices, and the interactions between the observation errors will be modeled by a full variance-covariance matrix.

2.3.2. Modeling interactions in the design and system matrices

When modeling the spatial structure of the two herds included in this study, three spatial levels, Pen, Section and Herd, are defined as follows:

- **Pen Level** describes individual sensors, each monitoring either the joined water consumption of pigs in two neighboring pens (Herd A), or in a single pen (Herd B).
- **Section Level** describes each of the physical sections of a herd as one individual unit including all sensors within the section.
- **Herd Level** describes the physical building in which the sections are placed, and includes all sensors in all pens and sections.

In order to describe any herd-specific interactions, the four sub...
models can be defined individually at either of the three spatial levels with the following properties:

- Defined at Pen Level the sub model evolves differently in each pen over time. One block is formed per pen in $G_i$.
- Defined at Section Level the sub model evolves identically in all pens within the section over time, expressing interactions between pens in the same section. One block is formed per section in $G_i$.
- Defined at Herd Level the sub model evolves identically in all pens in the herd over time, expressing interactions between all pens in the herd. One block including the whole herd is formed in $G_i$.

Let $M_{iLv}$ denote a sub model, $i$, where $i \in \{LG, H1, H2, H3\}$ indicates either the linear growth ($LG$), or one of the three cyclic models ($H1, H2, H3$). Furthermore, let a sub model be defined at level $Lv$, where $Lv \in \{h, s, p\}$ for either herd ($h$), section ($s$), or pen ($p$). Finally, $n_h$ and $n_p$ denote the number of sections in the herd and the number of pens in the herd, respectively.

As an example of the direct modeling of interactions in the design and system matrices; take a herd with two sections ($n_s = 2$) each with two sensors (so that $n_s = 2n_p = 4$), and the following definitions of interactions for each sub model:

- $M_{LGs}$ is defined at Section Level
- $M_{H1s}$ is defined at Herd Level
- $M_{H2s}$ is defined at Section Level

Fig. 7. The diurnal drinking pattern (black line) is shown together with the three harmonic waves; 24 h (H1), 12 h (H2), and 8 h (H3). The sum of the three harmonic waves and the underlying level (which is not depicted) is shown in (SUM).
\( M_{3I3p} \) is defined at Pen Level

The design matrix \( F^r \) will then have dimensions \( 18 \times 4 \) with the following structure:

\[
F^r = \begin{pmatrix}
1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0
end{pmatrix}
\]

(7)

And the corresponding system matrix \( G^r \) will have the dimensions \( 18 \times 18 \), consisting of nine blocks in the following block diagonal structure:

\[
G^r = \begin{pmatrix}
G_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & G_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & G_{13} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & G_{14} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & G_{22} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & G_{23} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & G_{24} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & G_{33} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & G_{34}
end{pmatrix}
\]

(8)

where 0 denotes a \( 2 \times 2 \) sub matrix only consisting of zeros.

2.3.3. Modeling interactions in the Observation variance-covariance matrix, \( V \)

The observational variances, \( V \), express the naturally complex random variation of water consumption as well as any uncertainty in the measurements, or observations, of the water data. The full variance-covariance matrix for the observation error \( v \) will be defined from three separate variance components corresponding to herd level, section level and pen level, respectively. This corresponds to an error structure as follows for the \( j \)th sensor (placed in Pen \( P_j \), located in Section \( S_j \) of Herd \( H_j \)):

\[
v_{ij} = \nu_{ij} + \nu_{sj} + \nu_{pj},
\]

(9)

where

\( \nu_{ij} \sim \mathcal{N}(0, \sigma_j^2) \) is an error term which is common for the entire herd;

\( \nu_{sj} \sim \mathcal{N}(0, \sigma_s^2) \) is an error term which is common for all pens in a section;

\( \nu_{pj} \sim \mathcal{N}(0, \sigma_p^2) \) is an error term which is specific for a pen.

As an example, consider again a herd with two sections, each holding two pens with a sensor in each of them. The distribution of the observation error then becomes

\[
v_{ij} \sim \mathcal{N}\left(\begin{pmatrix}
\sigma_1 + \sigma_2 + \sigma_3 \\
\sigma_1 + \sigma_2 + \sigma_3 \\
\sigma_1 + \sigma_2 + \sigma_3 \\
\sigma_1 + \sigma_2 + \sigma_3
end{pmatrix}
\right)
\]

(10)

All three variance components of \( V \) are assumed constant over time for all batches, but different between herds. Leaking drinking bowls or drinking nipples often occur for a shorter period of time in one or a few pens, and this is likely to affect the pen level variance, which is also assumed for any inaccuracies of the flow meters. Both mixing of liquid feed (in Herd A) and washing of sections between batches occur with the same frequency for all batches, affecting all pigs in the herd, and therefore presumably the herd level variance. Also a possible influence of weather conditions is assumed to be expressed at herd level.

2.3.4. Modeling interactions in the System variance-covariance matrix, \( W \)

The system variance, \( W \), expresses any uncertainty about the changes of the state vector from time \( t-1 \) to \( t \), and thereby determines the stability of the system over time. As described by West and Harrison (1999) the system variance \( W \) can be expressed as a fixed proportion of the estimated variance \( C \) of \( \delta \) given all observations until (and including) time \( t \) by a discount factor, \( \delta \), which by definition, satisfies the condition \( 0 < \delta \leq 1 \) (West and Harrison, 1999). Given \( \delta \) and \( C \), the whole system \( W \) can be identified as follows for each time step \( t \):

\[
W = \frac{1-\delta}{\delta} P,
\]

(11)

where

\[
P = G_t C_{t-1} G_t = V [G_t \delta_{t-1} D_{t-1}],
\]

(12)

However, one single discount factor is not recommended for a super positioned model (West and Harrison, 1999). Instead each sub model should be allowed to express different rates of change in stability over time through individual discount factors. In a diurnal pattern the harmonic characteristics are often more durable than the growth trend, which further emphasizes the need for a discount factor for each sub model (West and Harrison, 1999). We will therefore need several discount factors to express the system variance in the present model.

Based on West and Harrison (1999) the system variance of a super positioned DLM, can be defined using several discount groups as follows: Let \( M_i \) denote a sub model (i.e. a certain range of parameters of the parameter vector \( \theta \)). Let \( \gamma \geq 1 \) denote the number of sub models and let \( n_i \) denote the number of parameters of \( M_i \), so that \( \sum_{i=1}^\gamma n_i = n \) is the dimension of the full super positioned DLM.

A block diagonal approach is then applied where, for instance, \( G_n \) denotes the \( th \) block diagonal element of \( G \). Thus, for sub model \( M_i \), we have

\[
P_i = G_n C_{t-1} G_n = V [G_n \delta_{t-1} D_{t-1}], \quad i = 1, ..., \gamma
\]

(13)

and

\[
W_i = \frac{1-\delta_i}{\delta_i} P, \quad i = 1, ..., \gamma
\]

(14)

where \( \delta_i, ..., \delta_i \) are any set of discount factors, \( 0 < \delta_i \leq 1, i = 1, ..., \gamma \), with \( \delta_i \) being the discount factor associated with the sub model in question.

As described in Section 2.3.2, the spatial level of a sub model determines the number of blocks in the system matrix \( G \), and thereby the number of blocks in the system variance matrix \( W \) if there is no correlation between the sensors and the levels. The system variances for all blocks from the same sub model belong to the same discount group, meaning, they are expressed through the same discount factor. Since, however, some sensors are placed within the same section, and all sensors are placed within the same herd, some interactions between the changes of the corresponding state variables in the state vector are assumed:

\bullet \text{Pens are assumed dependent within the same section.}

\bullet \text{Sections are assumed dependent within the herd.}

\bullet \text{Pens are assumed independent between sections except for the correlation expressed through the sections.}

\bullet \text{The linear growth model and the cyclic models are assumed independent of each other.}

As a consequence of the defined correlation structure, the block structure for the example of Section 2.3.2 with two sections each with two sensors will result in \( \gamma = 5 \) blocks where
\[ G_{tt} = \begin{pmatrix} G_{t0}^{11} & 0 \\ 0 & G_{t0}^{22} \end{pmatrix}, \quad G_{tH} = G_{tH}^{11}, \quad G_{th} = \begin{pmatrix} G_{th}^{11} & 0 \\ 0 & G_{th}^{22} \end{pmatrix}. \]

Thus, \( G_{tt} \) is a 2 \( \times \) 2 matrix, whereas \( G_{tt}, G_{th}, G_{tt} \) and \( G_{th} \) are all 4 \( \times \) 4 matrices. The system variance matrix will have the same block structure with 5 diagonal blocks of the same dimensions (but without the zeros).

With the structure defined above it would be possible to use five different discount factors, but throughout the present study, the same discount factors are used for all harmonic models with the same wave length. This means that Blocks 4 and 5 in the example have the same discount factor. Hereby the number of different discount factors is reduced to four, with each discount factor relating to one of the four sub models (LG, H1, H2, H3) which constitute the full model.

2.4. Applying the DLM

2.4.1. Defining a herd

Throughout the development of the multivariate spatial DLM just described, it has been the aim to build a very general model in a way so it can be applied to a variety of herds, not regarding the number of pens and sections, levels of sub models, lengths of batches, sizes of pigs etc. In this section we will describe how the model is applied to the data, but also illustrate the flexibility of the general model by applying it to the data sets from both a finisher herd with eight sensors, Herd A, and a weaner herd with sixteen sensors, Herd B.

In the model, a herd is defined by the number of sensors monitoring water consumption. Input to the model is the number of sections with sensors, the number of sensors within the section, and a list of which pen and section each sensor is placed in. In order to distinguish batches from each other, the insertion date of all batches related to each sensor is also given as input to the model.

2.4.2. Handling missing observations

Both data sets are characterized by periods of missing observations, which can involve everything from one to all sensors in the herd, and last from one hour to a whole batch. Any period of missing observations is registered as NA observations in the data set, and is handled individually for each sensor by the model. Some NA observations are related to the cleaning period between two batches, where no observations are registered from any sensor in the empty section. These periods are regarded as planned periods of missing observations, and at the insertion of a new batch the model is reset as will be described in Section 2.4.3. Some NA observations are, however, unplanned missing observations. A period of unplanned NA observations can be caused by the pigs not drinking any water, or by sensor outages for a shorter or longer period. During a sequence of NA observations, the system keeps evolving and \( C_t \) increases.

2.4.3. Resetting between batches

All data from all batches monitored by one sensor represent one long time series. However, each new batch is likely to evolve differently over time. Therefore each time series must be divided into subseries with the length of the specific batch. If a new batch is inserted at time \( t_n \), the learned values for conditional mean, \( m_t \), and variance-covariance matrix, \( C_t \), from the previous batch are discarded, and the values are reset to \( m_t = m_0 \) and \( C_t = C_0 \).

2.4.4. Estimating variance components

As described in Sections 2.3.3 and 2.3.4, the observation variances are modeled through full variance-covariance matrices, whereas the system variances are modeled as a fixed proportion of the posterior variances, \( C_t \), using discount factors.

Any difference between the predicted multivariate observation and the actual observation is contained in a vector of forecast errors \( e_t \), which is defined as \( e_t = y_t - \hat{y}_t \), where \( \hat{y}_t \) is the mean of the one-step-ahead forecast for \( y_t \) given all observations until time \( t-1 \) (the forecast is provided by the Kalman filter).

If the pigs follow their normal drinking pattern and drink as much water as expected, the prediction of the next observation is close to perfect, and any forecast error, \( e_t \), will be small. If, on the other hand, something is causing the pigs to drink more or less than expected, the forecast error will be larger. The mean square error, MSE defined as \( \frac{1}{T} \sum_{t=1}^{T} e_t^2 \), will be used to measure the predictive performance of each model for comparison of the different versions.

All observation variances and discount factors are estimated on learning data by numerical optimization using the Nelder-Mead algorithm implemented in the optim function in R (R Core Team, 2014). The criterion of optimality is minimization of MSE. Thus, the observation variances and discount factors, which minimize the MSE for the learning data, are the results of the estimation. After insertion of a batch and after a sequence of NA observation lasting more than five hours, the model parameters need to adjust to the observations before reliable forecasts can be produced. Therefore, forecast errors for the first 24 h after insertion and after such an NA sequence are ignored in the evaluation of the MSE.

2.4.5. Model scenarios

Now the general structure of the spatial DLM is described, and seven different versions of the full DLM will be applied to each of the two data sets in order to identify, which versions describe any spatial correlation in the drinking pattern of finishers and of weaners the best. Each model differs with regard to the defined levels of the cyclic models (see Table 3), and the two data sets are divided into learning sets and test sets, as described in Section 2.1.2. The observation variances, \( V_t \), and the discount factors, \( \delta \), of each model will be estimated on learning data, and the estimated values will be entered as input to the model, when run on test data.

As previously described, Danish pig producing units for growing pigs are generally run with a sectionized structure. Therefore pigs of different ages and sizes are located in different, and separate, sections of a herd. In order to reflect that, the linear growth sub model will be

| Model version | Description |
|---------------|-------------|
| HH | The full harmonic pattern evolves identically for all pens in the herd |
| HS | H1 evolves differently for all pens |
| HS | H2 evolves identically within sections but differently between sections |
| HS | H3 evolves differently in each pen |
| HH | The full harmonic pattern evolves identically within each section but differently between sections |
| SSP | H1 and H2 evolve identically within sections but differently between sections |
| SPP | H2 and H3 evolve differently in each pen |
| PPP | The full harmonic pattern evolves differently in each pen |

Table 3

Model versions applied to data sets from Herd A and Herd B. The Linear Growth sub model is defined at Section level in all models, whereas different combinations of level definitions are made for the cyclic sub models H1, H2, and H3 (see Fig. 7). Notations: LG = Linear Growth model, H1 = Cyclic model of length 24, H2 = Cyclic model of length 12, H3 = Cyclic model of length 8. H = Herd level, S = Section level, P = Pen level.
Table 4
Estimated observation variances and discount factors for seven model structures for Herd A. Learning data: 9540 h (68 %). Test data: 4441 h (32%). Notations: LG = Linear Growth model, H1 = Cyclic model of length 24, H2 = Cyclic model of length 12, H3 = Cyclic model of length 8. H = Herd level, S = Section level, P = Pen level.

| Model Structure | Observation variances a | Discount factors b |
|-----------------|-------------------------|-------------------|
| LG H1 H2 H3     | c_1^S, c_2^S, c_3^S    | d_1, d_2, d_3     |
| S HHH           | 5.84 14.672 13.34      | 0.993 0.988 0.994 |
| S HSS           | 9.29 176.09 0.12       | 0.993 0.989 0.993 |
| S HSP           | 0.60 90.39 172.45     | 0.993 0.989 0.994 |
| S SSS           | 0.10 104.41 0.10      | 0.992 0.994 0.999 |
| S SSP           | 3.59 37.37 121.72     | 0.994 0.996 0.994 |
| S SPP           | 1.28 12.87 152.15     | 0.994 0.990 0.995 |
| S PPP           | 0.13 3.56 E-6 78.09   | 0.994 0.991 0.995 |

a c_1^S: Observation variance, herd effect, c_2^S: Observation variance, section effect, c_3^S: Observation variance, pen effect
b d_i: LG, d_1: H1, d_2: H2, d_3: H3

defined at section level for all seven versions in each herd, but the cyclic sub models of the DLM will be defined at different level combinations, which can be seen in Table 3.

For both herds, and all model versions, the same four discount groups were defined; namely one for each of the four sub models.

3. Results and discussion

The estimated variance components and the predictive performance (evaluated as MSE) for each of the seven model versions are presented for both herds.

3.1. Estimated variance components

Tables 4 and 5 show the estimated variance components for each of the seven models applied to data from Herd A and Herd B. In the majority of the models, the observational variances, c_1^S, c_2^S and c_3^S, are very high, especially c_2^S for Herd B, but also c_3^S and c_3^F for Herd A. In general, high observational variances in a DLM can indicate that all variation in the data is expressed through the observation variances, apparently because of a very rigid and non-flexible system.

However, a rigid system would favour very high discount factors (i.e. δ very close to 1), leaving little room for any instability in the system variances (West and Harrison, 1999). This is not the case for the discount factors of this model, as it can be seen in Tables 4 and 5. Especially for Herd B, the estimated discount factors are fairly low, which results in a highly flexible model on all parameters, capable of adjusting very well to the learning data set (Witten and Frank, 2005).

The high flexibility, which characterize the models for both herds, can also be caused by too high a complexity of the models, that is an excess of parameters used to describe the data, hereby leading to an overfitting of data (Hawkins, 2004; Witten and Frank, 2005; Elith et al., 2008; Torgo, 2017). Overfitting often occurs in models with highly correlated parameters (Hawkins, 2004; Witten and Frank, 2005). The high observational variances we see in the DLM presented here, can be the result of a so-called bias-variance tradeoff, which is associated with overfitting (Torgo, 2017). The bias-variance tradeoff describes how a model either adjusts too well to the training data, hereby reducing the bias (MSE) and increasing the variance, or decreases the variance by a reduced sensitivity to the learning data, which results in a higher bias (Witten and Frank, 2005; Torgo, 2017). Since the estimation of variance components in the presented DLM were aiming for the lowest MSE, and the variables in the model are highly correlated, it is very likely, that the model is overfitting the learning data.

Although overfitting in general is sought avoided and does not add to an increased performance (Hawkins, 2004), it does not necessarily impede the predictive performance either (Elith et al., 2008; Lieberman and Morris, 2014). Overfitting is a well known challenge when handling correlated variables or variables analyzed using regression techniques (Hawkins, 2004). But Elith et al. (2008) found that the predictive performance of an overfitted model was unaffected when using Boosted Regression Trees and evaluating using cross-validation. Likewise Lieberman and Morris (2014) concluded that multi-collinearity in cross-validated models, was irrelevant if prediction performance was the goal of the model.

3.2. Predictive performance

In Tables 6 and 7 the MSE of both learning data and test data from each herd is presented. The predictive performance of a model should not be evaluated on learning data (Elith et al., 2008), but we still choose to present the models’ MSE on learning data in order to show how the MSE is lower on learning data due to possible overfitting, as described above.

When comparing the MSE of the learning data to the MSE of the test data for Herd A, it can be seen, that the MSE of the test data is 3.5 times the MSE of the learning data on average (see Table 6), whereas the MSE on learning data is only 1.3 times higher than for test data for Herd B on average (see Table 7). The relative large ratio between learning-MSE and test-MSE for Herd A may indicate differences in the learning data and the test data. Initial evaluation of the full data set does, however, not indicate this, and the ratio is more likely to be the result of the variance-estimating method, which aim to minimize the learning-MSE (James et al., 2013).

Table 6
MSE (mean squared error) for learning data and test data for seven model structures for Herd A. Notations: LG = Linear Growth model, H1 = Cyclic model of length 24, H2 = Cyclic model of length 12, H3 = Cyclic model of length 8. H = Herd level, S = Section level, P = Pen level.

| Model Structure | LG H1 H2 H3 Learning data a | Test data b |
|-----------------|-----------------------------|-------------|
| S HHH           | 4.003 15.687               |             |
| S HSS           | 3.923 14.535               |             |
| S HSP           | 3.921 14.612               |             |
| S SSS           | 3.954 13.850               |             |
| S SSP           | 3.850 13.976               |             |
| S SPP           | 3.800 13.924               |             |

a 9540 h (68%)
b 4441 h (32%)
drink during the night time in order to get their need for water satisfied. Batches are likely to underperform. The herd level model versions to underperform. This abnormal pattern in pen 1.6, peaking every eight hour instead of once per 24 h as in the undisturbed pattern in pen 1.6 is disturbed, which results in the water consumption a week in four pens in four sections in Herd A. The diurnal drinking pattern would characterize pigs. Andersen et al. (2016), and installing an extra drinking nipple might be by the same cyclic sub model. The planned periods of missing data during cleaning periods be ned at herd level as well. Our initial assumption was that a distinct hypothesis of a correlation between pens and sections in a herd, differences between pens in the herd are too large to be described by the same cyclic sub model. This is well illustrated in Fig. 8, which shows the drinking pattern of a week in four pens in four sections in Herd A. The diurnal drinking pattern in pen 1.6 is disturbed, which results in the water consumption peaking every eight hour instead of once per 24 h in the undisturbed diurnal patterns of pens 2.5 and 3.5. This abnormal pattern in pen 1.6, in combination with a longer sensor outage in pen 5.7, decreases any diurnal patterns of pens 2.5 and 3.5. This abnormal pattern in pen 1.6 is disturbed, which results in the water consumption a week in four pens in four sections in Herd A. The diurnal drinking pattern would characterize pigs. Andersen et al. (2016), and installing an extra drinking nipple might be by the same cyclic sub model. The planned periods of missing data during cleaning periods be ned at herd level as well. Our initial assumption was that a distinct hypothesis of a correlation between pens and sections in a herd, differences between pens in the herd are too large to be described by the same cyclic sub model. This is well illustrated in Fig. 8, which shows the drinking pattern of a week in four pens in four sections in Herd A. The diurnal drinking pattern in pen 1.6 is disturbed, which results in the water consumption peaking every eight hour instead of once per 24 h in the undisturbed diurnal patterns of pens 2.5 and 3.5. This abnormal pattern in pen 1.6, in combination with a longer sensor outage in pen 5.7, decreases any correlation between sections, and situations like this are likely to cause the herd level model versions to under perform.

The planned periods of missing data during cleaning periods between batches are likely to affect the model versions with parameters defined at herd level as well. Our initial assumption was that a distinct diurnal pattern would characterize pigs’ drinking pattern throughout the growth period. However, the pigs in pen 1.6 were inserted 10 weeks before the depicted week, and show no such diurnal pattern at this time.

Such a disturbed diurnal pattern indicates that some pigs have to drink during the night time in order to get their need for water satisfied. Drinking activity in finisher pigs during the night was also found by Andersen et al. (2016), and installing an extra drinking nipple might be necessary to restore the diurnal pattern, and supply sufficient amounts of water to the pigs. This is especially profound for finisher pigs that are fed liquid feed, since the restrictively feeding from 60 kg reduces the amount of water assigned through the feed and increases the demand for water from the drinking nipples. The results in Table 6 also show that the model version with all four sub models defined at section level outperforms any model version including cyclic sub models defined at pen level. This supports the initial hypothesis of a correlation between pens and sections in a herd, which can be modeled in a spatial model.

### 3.4. Model versions - Herd B

In Table 7 we see that the best performing model version is the one with all cyclic sub models defined at pen level. This result indicates that any correlation between pens and sections in a herd may be described solely through the correlation structure of the system variance-covariance matrix, \( W \), as described in Section 2.3.4 and through the observation variance-covariance matrix described in Section 2.3.3.

In general, model versions which included cyclic sub models at section level, performed better than any model version including herd level sub models. The herd level versions may fail in Herd B for the same reasons as in Herd A, although the cleaning periods between batches were shorter in Herd B.

The the poorer performance of models with higher degrees of spatial correlation is surprising considering the very disciplined sectionized management on the research farm and the uniformity of the pigs both within sections and within the herd relative to a finisher herd. The main reason for the result may be the overfitting of the model as previously discussed. Since overfitting models tend to model the random noise of a system, as described by Witten and Frank (2005), they are more flexible and quickly adjust to smaller changes (Torgo, 2017). This is in compliance with the results of the estimation of variance components. An adjustment to smaller changes can cause the model to emphasize the importance of random noise within the single pen, hence failing to recognize any correlation between pens.

#### 3.5. Estimation procedure

Since the end goal of the project is to use the developed DLMs for early warning about undesired events it could be argued that the discount factors and observation variances should have been estimated on a learning data set where such events did not occur at all. Such an approach was, for instance, used by Jensen et al. (2017) and the obvious advantage would be that if parameters are estimated under “normal” conditions, deviations from the normal pattern in case of undesired events are more easily detected.

In this study, the learning data set was simply the first batches of the study period and the test data set was the last two batches. The reason for this choice was that we present a framework for simultaneous monitoring of the entire herd. Thus, it is not possible to find batches without any undesired events in any pen of the herd as it was for Jensen et al. (2017) who basically modeled a single pen. Thus, the advantage of fitting the model to “normal” conditions is lost which, potentially, will make it more difficult to detect deviating data patterns caused by undesired events.

An indirect estimation technique through discount factors was used for assessing the system variance-covariance matrices of the models. Several initial attempts were done to estimate all variance components more directly by the EM algorithm (see for instance Dethlefsen, 2001) as previously done in multivariate DLMs by Bono et al. (2012) and Jensen et al. (2017) but the iterative algorithm failed to converge and some of the variances drifted out of scope over the iterations. Therefore, the indirect approach with discount factors was used instead.

The full univariate model (i.e. for one sensor) as presented in Section 2.3 is directly inspired by the work of Madsen et al. (2005) who concluded that a super positioned model consisting of a linear growth component and three harmonics described the diurnal pattern well. It is interesting that Madsen (2001, Chapter 8) reported that the EM algorithm also failed for the univariate model despite several attempts. Thus, it seems to be a pattern that the EM algorithm is not well suited for estimation in models with harmonics based on Fourier form representation of seasonality. The observation is supported by unpublished work by the authors in relation to similar models with diurnal patterns.

Madsen (2001) mentioned that the system variance-covariance matrix must be expected to change over time as the pigs grow.

---

**Table 7**

MSE (mean squared error) for learning data and for test data for seven model structures for Herd B. Notations: LG = Linear Growth model, H1 = Cyclic model of length 24, H2 = Cyclic model of length 12, H3 = Cyclic model of length 8. H = Herd level, S = Section level, P = Pen level.

| Model Structure | Learning data | Test data |
|-----------------|---------------|-----------|
| LG H1 H2 H3     |               |           |
| S               | 1.350         | 1.750     |
| S               | 1.327         | 1.727     |
| S               | 1.325         | 1.712     |
| S               | 1.235         | 1.621     |
| S               | 1.234         | 1.559     |
| S               | 1.229         | 1.556     |
| S               | 1.194         | 1.466     |

* 14,657 h (83%)
* 3025 h (17%)

Since the ratio between the two MSE values is not relevant for choosing the model version with the better predictive performance (James et al., 2013), we will only refer to the MSE of the test data in the following discussion of the results.

In general, the MSE values found for the seven model versions are surprisingly close within each of the two herds. It is therefore not possible to point out the one model version which will perform the better when predicting outbreaks of diarrhea or pen fouling based on the current results. Therefore all seven model versions will be evaluated in the second, and final, part of the full description of the spatial detection system presented in Dominiak et al. (submitted for publication).

### 3.3. Model versions - Herd A

When comparing the MSE of the different model versions from Herd A (see Table 6), it shows that all model versions which include cyclic sub models defined at herd level have the highest MSE. This indicates that differences between pens in the herd are too large to be described by the same cyclic sub model.

This is well illustrated in Fig. 8, which shows the drinking pattern of a week in four pens in four sections in Herd A. The diurnal drinking pattern in pen 1.6 is disturbed, which results in the water consumption peaking every eight hour instead of once per 24 h in the undisturbed diurnal patterns of pens 2.5 and 3.5. This abnormal pattern in pen 1.6, in combination with a longer sensor outage in pen 5.7, decreases any correlation between sections, and situations like this are likely to cause the herd level model versions to under perform.

The planned periods of missing data during cleaning periods between batches are likely to affect the model versions with parameters defined at herd level as well. Our initial assumption was that a distinct diurnal pattern would characterize pigs’ drinking pattern throughout the growth period. However, the pigs in pen 1.6 were inserted 10 weeks before the depicted week, and show no such diurnal pattern at this time.

Such a disturbed diurnal pattern indicates that some pigs have to drink during the night time in order to get their need for water satisfied. Drinking activity in finisher pigs during the night was also found by Andersen et al. (2016), and installing an extra drinking nipple might be necessary to restore the diurnal pattern, and supply sufficient amounts of water to the pigs. This is especially profound for finisher pigs that are fed liquid feed, since the restrictively feeding from 60 kg reduces the amount of water assigned through the feed and increases the demand for water from the drinking nipples.

The results in Table 6 also show that the model version with all four sub models defined at section level outperforms any model version including cyclic sub models defined at pen level. This supports the initial hypothesis of a correlation between pens and sections in a herd, which can be modeled in a spatial model.
Therefore, it was argued that an “online” estimation technique based on discount factors should actually be preferred to estimation by the EM algorithm. It is not clear whether the system variance-covariance actually do change over time, but if Madsen (2001) is right, the discount factor approach will also be a good choice for the present study.

It was also argued that the direct link to the variability of data as expressed by Eqs. (13) and (14) makes the approach less herd specific, because no new estimation is needed for each herd (as long as the estimated discount factors are valid across herds). The results from this study (cf. Tables 4 and 5) do not confirm that discount factors are identical for different herds (at least not for a multivariate spatial model) so it is expected that an estimation step is needed for each herd before the system is ready for use.

When looking at the observation variances of Tables 4 and 5 it is quite obvious that the values are not estimates for the true observation variances. They are far too big and they should just be seen as the values optimizing the fit in combination with the resulting discount factors. The very big values seen for many of the models have as a consequence that the models become less adaptive which may actually be an advantage given that the ultimate use of the models is to produce early warnings of deviating patterns caused by undesired events.

A similar behavior could have been achieved with discount factors closer to 1, but the best fit was apparently achieved by very high “observational” variance and smaller discount factors. It is, however a

---

**Fig. 8.** The water consumption from four pens in four different sections of Herd A in the same week. The pigs in each of the four pens are of different ages with the oldest in pen 1.6. The diurnal pattern is disturbed in pen 1.6, but intact in pens 2.5 and 3.5. In pen 5.7 there is no data due to sensor outage.
question whether a generalized (to the multivariate case) version of the Kalman filter with unknown observational variance should have been developed for this study. Madsen et al. (2005) used that approach for the univariate case so a suggestion for future research is to see whether the method can be extended to the multivariate case.

Even though the observational error structure described in Eqs. (9) and (10) intuitively seems natural, it can be discussed whether it actually over-parameterizes the observation errors. Particularly, it is a question whether the herd level variance (σr²) should have been taken out. For Herd A, this term only contributed with up to 5% of the total observation variance and for Herd B it never even reached 1% of the variance (percentages calculated from Tables 4 and 5). Thus, in future studies, it is recommended not to include the herd level variance.

4. Conclusion

We can conclude, that it is possible to develop a spatial DLM for the modeling of drinking patterns across a herd of growing pigs. In Herd A, the model version expressing the strongest correlation in the drinking patterns between pens within a section (the SSM model) obtains the highest fit. Model versions which include parameters at pen level (SSP, SPP, and PPP) fit almost as well, whereas model versions with parameters expressed at herd level (HHH, HSS, and HSP) fit the worse. Based on model fit, correlation between pens do occur in Herd A, but primarily between pens within the same section. In Herd B, a distinctive inverse relation between model fit and degree of correlation in the drinking patterns are found. This results in the model version with the least correlation between the pens (the PPP model) obtaining the highest fit, and the model version with the highest degree of correlation (the HHH model) obtaining the poorest fit. Thus, based on model fit, little or no correlation between pens occur in Herd B. For both herds, overfitting of test data may influence the results.

5. Perspectives

The overall motivation for the development of the presented model is to investigate spatial modeling of water consumption as a strategy for a future detection system in commercial pig production. However, the ability of the model to identify unwanted events in the herd must be evaluated before taking any further steps. Such an evaluation of the detection performance will be conducted in a following paper. Thus the forecast errors generated by all seven model versions are monitored in a control chart, and the ability of the detection system to predict breaks of either diarrhea or fouling is described. Dominik et al. (submitted for publication).

Acknowledgements

This research was funded by the Danish Council for Strategic Research (The PigIT project, Grant No. 11-116191). We further wish to thank the anonymous farmers, the Danish Pig Research Centre and the technical staff (particularly Mads Ravn Jensen) at Aarhus University for installing and supervising the sensors and for taking care of the daily observations in the herd.

References

Andersen, H.M.L., Jørgensen, E., Pedersen, L.J., 2016. Using evolutionary operation technique to evaluate different management initiatives at herd level. Livestock Sci. 187, 109–113.
Anonymouse, 2000. Instruction leaflet flow sensors - 15 mm dia. pipe v8189. URL http://doc.asa.electrocomponents.com/webdoc/01bi9/090766880116b47.pdf.
Aparna, U., Pedersen, L.J., Jørgensen, E., 2014. Hidden phase-type markov model for the prediction of onset of farrowing for loose-housed sows. Comput. Electron. Agric. 108, 135–147.
Berkvens, D., 2014. Precision livestock farming technologies for welfare management in intensive livestock systems. (special issue: Animal welfare: focusing on the future.). Revue Scientifique et Technique - Office International des Epizooties 33 (1), 189–196.
Bono, C., Cornous, C., Kristensen, A.R., 2012. Dynamic production monitoring in pig herds I: modeling and monitoring litter size at herd and sow level.Livestock Sci. 149 (3), 289–300.
Bressers, H.P.M., Brake, J.H.A.T., Noordhuizen, J.P.T.M., 1995. Automated oestrus detection in group-housed sows by recording visits to the boar. Livestock Prod. Sci. 41 (2), 183–191.
Brown-Brandl, T., Rohrer, G., Eigenberg, R., 2013. Analysis of feeding behavior of group housed growing-finish pigs. Comput. Electron. Agric. 96, 246–252.
Cameron, R.D.A., 2000. A review of the industrialisation of pig production worldwide with particular reference to the asian region. Food and Agriculture Organization of the United Nations. URL http://www.fao.org/ag/aginos/index-en/publications/agapubs/awi_concept_pig_product.pdf.
Danish Agriculture and Food Council, 2015. Danish Pig Producers and Animal Welfare. Tech. rep. Danish Agriculture and Food Council.
de Mol, R.M., Kroese, G.H., Achten, J.M.F.H., Maatte, K., Rossing, W., 1997. Results of a multivariate approach to automated oestrus and mastitis detection. Livestock Prod. Sci. 48 (3), 219–227.
de Mol, R.M., Wold, W.E., 2001. Application of fuzzy logic in automated cow status monitoring. J. Dairy Sci. 84.
Desjen, C., 2001. Space Time Problems and Applications (Ph.D. Thesis). Aalborg University.
Dominik, K.N., 2017. Spatial Modeling of Drinking Patterns as a Tool for Reducing Alarms in Pig Production (Ph.D. Thesis). The Graduate School of Health and Medical Sciences. Aarhus University.
Dominik, K.N., Kristensen, A.R., 2017. Prioritizing alarms from sensor-based detection models in livestock production - a review on model performance and alarm reducing methods. Comput. Electron. Agric. 133, 46–67.
Dominik, K.N., Pedersen, L.J., Kristensen, A.R., submitted for publication. Spatial modeling of pigs' drinking patterns as an alarm reduction method II. Application of a multivariate dynamic linear model.
Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. J. Anim. Ecol. 77 (4), 802–813.
Garcia, E., Klaua, I., Amigo, J.M., Bro, E., Enevoldsen, C., 2014. Lameness detection challenges in automated milking systems addressed with partial least squares discriminant analysis. J. Dairy Sci. 97 (12), 7476–7486.
Hawkins, D.M., 2004. The problem of overfitting. J. Chem. Inf. Comp. Sci. 44 (1), 1–12. http://dx.doi.org/10.1021/ci0342472.
Hoveheen, V., Hampuiski, C., Steeneveld, W., Mollenhorst, H., 2010. Sensors and clinical mastitis - the quest for the perfect alert. Sensors 10 (9), 7991–8009.
Huybrechts, T., Mertens, K., Baardezaeker, J.D., Ketelaere, B.D., Saeya, W., 2014. Early warnings from automatic milking yield monitoring with online syringe control. J. Dairy Sci. 97 (6), 3371–3381 6.
James, G., Witten, D., Hastie, T., Tibishirani, R., 2013. An Introduction to Statistical Learning. Springer.
Jensen, D.B., Toft, N., Kristensen, A.R., 2017. A multivariate dynamic linear model for early warnings of diarrhea and pen fouling in slaughter pigs. Comput. Electron. Agric. 135, 1–62.
Kashiha, M., Bahr, C., Hardeshad, S.A., Ott, S., Moons, C.P.H., Niewold, T.A., Odberg, T.O., Berckmans, D., 2013. The automatic monitoring of pigs water use by cameras. Comput. Electron. Agric. 94, 164–169.
Kristensen, A.R., Jørgensen, E., Toft, N., 2010. Herd Management Science I. Basic concepts. Academic books, Copenhagen.
Krogsdal, J., 2014a. Description of Grønhøj farm. PigIT Report 6, The PigIT project.
Krogsdal, J., 2014b. Department of Large Animal Sciences, University of Copenhagen. Lieberman, M., Morris, J.D., 2014. The precise effect of multicollinearity on classification accuracy. Comput. Electron. Agric. Multiple Linear Regression Viewpoints 40 (1), 5–10.
Madsen, T.N., 2001. Tools for Monitoring Growing Pigs (Ph.D. Thesis). The Royal Veterinary and Agricultural University. Madsen, T.N., Andersen, S., Kristensen, A.R., 2005. Modelling the drinking patterns of young pigs using a state space model. Comput. Electron. Agric. 48 (1), 39–62.
Maertens, W., Vangeyte, J., Baert, J., Jantuan, A., Mertens, K.C., de Campeneere, S., Pluik, A., Opromoler, G., van Weyenberg, S., Nuffield, A., 2011. Development of a real time cow gait tracking and analysing tool to assess lameness using a pressure sensitive walkway: the gaitwise system. Biostat. Eng. 110 (1), 29–39.
Oczak, M., Imsaylovila, G., Costa, A., Viazio, S., Sonoda, L.T., Fels, M., Bahr, C., Hartung, J., Guatino, M., Bεrkvens, D., Vrancken, E., 2013. Analysis of aggressive behaviours of pigs by automatic video recordings. Comput. Electron. Agric. 99, 209–217.
Ostersen, T., Cornou, C., Kristensen, A.R., 2010. Detecting oestrus by monitoring sows’ visits to a boar. Comput. Electron. Agric. 74 (1), 51–60.
Pedersen, K.S., 2012. Smaagresidieare. Dansk Veterinar Tidskrift 6, 18–22.
R Core Team, 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.
Steeneveld, W., Gagg, L.C.v.d., Barkema, H.W., Hoveheen, H., 2010. Simplify the interpretation of alert lists for clinical mastitis in automatic milking systems. Comput. Electron. Agric. 71 (1), 50–56.
Torgo, L., 2017. Data Mining with R Learning with Case Studies. Chapman & Hall.
Vilu, E., 2012. Vaadalderkuver. Tech. Rep. SEGES. Vilu, E., 2013. Manual om vaadaldermanagement. Tech. Rep. SEGES, Danish Pig Research Centre.
West, M., Harrison, J., 1999. Bayesian Forecasting and Dynamic Models. 2nd ed. Springer.
Witten, I., Frank, E., 2005. Data mining: practical machine learning tools and techniques, second ed. Morgan Kauffman, Elsevier.
Zhuang, X., Bi, M., Guo, J., Wu, S., Zhang, T., 2018. Development of an early warning algorithm to detect sick broilers. Comput. Electron. Agric. 144, 102–113.