Computers and Electronics in Agriculture 161 (2019) 92–103

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Spatial modeling of pigs' drinking patterns as an alarm reducing method II. Application of a multivariate dynamic linear model

K.N. Dominiak[a,b,c,*], J. Hindsborg[a], L.J. Pedersen[b], A.R. Kristensen[a]

[a] Faculty of Health and Medical Sciences, Department of Veterinary and Animal Sciences, University of Copenhagen, Grønnegårdsvej 2, 1870 Frederiksberg C, Denmark
[b] Department of Animal Science, Aarhus University, Blichers Allé 20, 8830 Tjele, Denmark
[c] Livestock Innovation, SEGES, Danish Agriculture and Food Council F.m.b.A., Agro Food Park 15, 8200 Aarhus N, Denmark

ARTICLE INFO

Keywords:
Detection performance
Early warning
Tabular Cusum
Sensor-based
Water consumption

ABSTRACT

The objectives of this paper are to evaluate the detection performance of a previously developed multivariate spatial dynamic linear model (DLM), which aim to predict outbreaks of either diarrhea or pen fouling amongst growing pigs, and to discuss potential post processing strategies for reducing alarms. The model is applied to sensor based water data from a commercial herd of finisher pigs (30–110 kg) and a research facility herd of weaner pigs (7–30 kg). Performance evaluation is conducted by applying a *standardized two-sided Cusum*, on the forecast errors generated by the spatial model. For each herd, forecast errors are generated at three spatial levels: Pen level, section level, and herd level. Seven model versions express different temporal correlations in the drinking patterns between pens and sections in a herd, and the performances of each spatial level are evaluated for every model version. The alarms generated by the Cusum are categorized as true positive (TP), false positive (FP), true negative (TN), or false negative (FN) based on time windows of three different lengths. In total, 126 combinations of herds, spatial levels, model versions, and time windows are evaluated, and the performance of each combination is reported as the *area under the ROC curve* (AUC). The highest performances are obtained at herd level given the longest time window and strongest temporal correlation (AUC = 0.98 (weaners) and 0.94 (finishers)). However, the settings most suitable for implementation in commercial herds, are obtained at section level given the medium-length time window and strongest temporal correlation (AUC = 0.86 (weaners) and 0.87 (finishers)). The combination of a spatial DLM and a two-sided tabular Cusum has high potential for prioritizing high-risk alarms as well as for merging alarms from multiple pens within the same section into a reduced number of alarms communicated to the caretaker. Thus, the spatial detection system described here, and in a previous paper, constitute a new and promising approach to sensor based monitoring tools in livestock production.

1. Introduction

For more than 20 years the development of sensor-based detection models within the field of livestock science has been subject to an increasing scientific focus. However, a general problem for detection models is that they generate too many false alarms (Hogeveen et al., 2010; Dominiak and Kristensen, 2017). False alarms reduce the usefulness of a detection model as a decision-support tool, and represent a major reason for models being unsuited for implementation in modern livestock production herds (Mein and Rasmussen, 2008; Hogeveen et al., 2010).

The number of false alarms declines as the performance of a detection model increases. However, both a review of clinical mastitis (CM) detection models (Hogeveen et al., 2010), as well as a recent review paper focusing on livestock related sensor-based detection models in scientific literature from 1995 to 2015 (Dominiak and Kristensen, 2017), show that it is exceedingly difficult to achieve detection performances so high that the number of false alarms will be acceptable in a real-life production herd.

In future development of sensor-based detection systems, it may therefore be highly relevant to focus equally on both achieving very high detection performances and implementing different methods for prioritizing, sorting, or categorizing the generated alarms. Such methods have not had primary focus throughout the scientific literature (Dominiak and Kristensen, 2017). Thus, only three methods; Fuzzy logic (de Mol and Woldt, 2001), Naïve Bayesian Network (NBN) (Steeneveld et al., 2010), and Hidden phase-type Markov (Aparna et al., 2014), are described as alarm-reducing methods in peer-reviewed papers from 1995 to 2015.

Both de Mol and Woldt (2001) and Steeneveld et al. (2010) combine

* Corresponding author at: Livestock Innovation, SEGES, Danish Agriculture and Food Council F.m.b.A., Agro Food Park 15, 8200 Aarhus N, Denmark.
E-mail address: kand@seges.dk (K.N. Dominiak).

https://doi.org/10.1016/j.compag.2018.10.037
Received 5 September 2017; Received in revised form 6 September 2018; Accepted 25 October 2018
Available online 31 October 2018

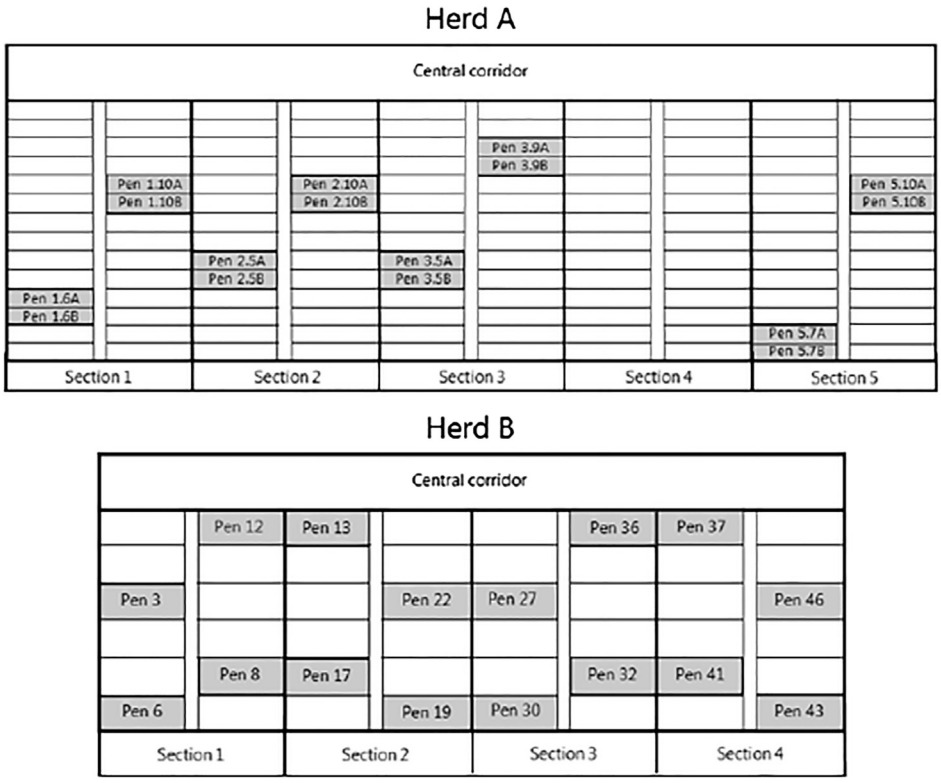

Fig. 1. Structure of Herd A and Herd B. Grey double pens in Herd A and grey pens in Herd B were equipped with sensors and included in the study. From Dominiak et al. (2018).

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.

The three methods mentioned above, all focus on identifying the events of interest at *individual animal level*. However, when considering the production of growing pigs (weaners 7–30 kg and finishers 30–110 kg), both production systems and management routines are based on groups of animals in pens within sections in a herd rather than on individual animals. These systematic production routines mean that the manager at any given time knows the age and size of the pigs in any given section. Furthermore, the pigs within a section are sorted according to criteria like gender or size placing the smaller pigs, which need higher managerial focus, in pens closer to the entrance door, and larger pigs in other pens following a herd-specific strategy. In the everyday management of growing pigs, it will therefore be highly relevant with a spatial detection system, which points out specific pens or sections for the manager to aim focus.

The spatial detection system, evaluated in this paper, incorporates correlations between simultaneously monitored water data from all pens with sensors in a herd, and is therefore able to identify a specific area (pen or section) within the herd where the outbreak of an unwanted condition is about to happen. An area-specific alarm will enable the manager to prioritize which alarm to attend to first according to specific knowledge of the pigs in the alarmed areas. The combination of area-specific alarms and knowledge of the animals in the pointed areas furthermore enables the manager to choose the right intervention for the pigs in the alarmed area, whereby the consequences of a condition can be reduced or prevented.

Danish pig producing units for growing pigs are very well suited for spatial modeling. In order to maintain a high level of bio-security, these units are run with a clear spatial separation between pigs of different age groups (Danish Agriculture and Food Council, 2010). Such a construction of the herd enables a spatial modeling of a production site as one whole production unit (the herd) consisting of a number of identical subunits (sections) each consisting of a number of identical sub-subunits (pens). A warning system based on spatial modeling, can communicate area-specific warnings (Lawson et al., 2016), which will provide the manager with valuable information when choosing an intervention to prevent or reduce the consequences of a condition.

In this paper we will present the second, and final, part of the full description of a spatial approach to the challenge of reducing alarms. The description is initiated in Dominiak et al. (2018) where the development of a multivariate dynamic linear model (DLM) for modeling water consumption of growing pigs in Denmark is described. The structure of the DLM will be described in short terms in Section 2. The objectives of the present paper are to evaluate model performances on test data from two different herds, using Area Under Curve (AUC), and to discuss post processing strategies for reducing the number of alarms that are generated.

2. Herds, data and models

In this section a short background is given to the spatial DLM whose predictive accuracy is evaluated in the present paper. It is based on Dominiak et al. (2018), where a detailed description of the model development can be found.

2.1. Herd description

Water consumption data for this study was obtained from two different herds. Herd A is a Danish commercial finisher herd, and Herd B is an experimental weaner herd, “Grønhøj”, which is owned by the Danish Pig Research Centre.

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. 1 (Herd A)). Each section consists of 28 pens, where 18 pigs are

inserted per pen. Two neighbouring pens share the same water pipe, which supplies one drinking nipple in each of the two pens (36 pigs).

Herd B consists of four sections, each with 12 pens for weaner pigs, where 15 pigs are inserted per pen (Fig. 1 (Herd B)). Each pen is supplied by individual water pipes, hence one water pipe supplies one drinking bowl per pen (15 pigs).

2.2. Sensor data

Water consumption data was obtained by a flow meter (RS V8189 15 mm Diameter Pipe) (Anonymous, 2000), which was placed on the water pipe supplying either drinking nipples (Herd A) or bowls (Herd B) in the pens. The data was converted to litres before it was aggregated per hour, yielding water use in litres per hour as input from each sensor to the DLM.

A total of eight sensors were installed in Herd A with two sensors in each of four sections, and each sensor monitoring the water consumption of two neighbouring pens, a double-pen (36 pigs). Data from one sensor created an individual time series, hence the full data set from Herd A consists of eight time series, or variables, which were monitored simultaneously.

Sixteen sensors in total were installed in Herd B with four sensors in each of four sections. Each sensor monitored the water consumption of one single pen (15 pigs) in individual time series. The full data set from Herd B therefore consists of sixteen time series, or variables, which were monitored simultaneously.

Throughout the rest of the paper, a pen is defined as the area comprising the number of pigs whose water consumption was monitored by a single sensor (see Table 1). A section is defined as all pens with sensors within the same section in the herd, and a herd is defined as all pens with sensors within the farm building.

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

2.3. Modeling drinking patterns

The drinking patterns of both weaners and finishers followed a clear diurnal pattern (Fig. 2), and the underlying level of water consumed increased over time, indicating that pigs drank more as they grew (Madsen et al., 2005).

The diurnal drinking patterns were described by the sum of four dynamic linear models, each constituting a sub-model in a larger DLM. In total, three sub-models for harmonic waves (H1, H2, H3) plus one for the under-lying linear growth (LG) were superpositioned into the final full DLM (Madsen et al., 2005). H1 described a harmonic wave, which peaks every 24 h, whereas H2 described a harmonic wave, which peak every 12 h, and H3 described a harmonic wave, which peaks every 8 h (see Fig. 3).

The amount of water consumed within the last hour at time t for each of the n sensors was expressed in the observation vector $Y_t = (Y_{1t}, ..., Y_{nt})'$. The aim of the DLM was to predict the next observation by estimating the parameter vectors $\theta_1, ..., \theta_t$ from the observations.

Table 1
Characteristics for the two herds in the study. From Dominiak et al. (2018).

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/per section	8/2	16/4
Pigs per pen/per sensor	18/36	15/15
Growth period (batch)	14 weeks	8 weeks
Batches per sensor	7	13[a]
Learning data (hours)	9540	14657
Test data (hours)	4441	3025

[a] 14 for Section 4.

The accuracy of the predictions was expressed through forecast errors e_t, which contain any differences between the predicted observation and the actual observation. As long as the drinking pattern reflected a normal situation and evolved as expected, the prediction of the next observation was close to perfect, and any forecast error would be small. Should the pigs, for some reason, drink more or less than expected, the predictions and the observations diverge, and the errors would be larger. A systematic change in the normal drinking pattern therefore generated a sequence of forecast errors, which lead to an alarm when plotted in a control chart, as described in Section 3.

2.4. Model versions

Each of the four sub-models were defined at herd, section, or pen level to allow for the diurnal pattern to express different degrees of temporal correlation between pens or sections in the herd, thus reflecting pigs drinking at different times during the day. Seven different model versions were defined (Table 2) all in which the LG sub-model was defined at section level in order to reflect the relatively uniform growth rate of pigs within a section.

2.5. Model output

From each of the seven model versions a series of forecast error vectors (e_t) and a series of forecast variance–covariance matrices (Q_t) were generated. The forecast errors and variances were entered as input variables to a *standardized two-sided Cusum control chart*, as described by Montgomery (2013). Systematic changes in the water consumption then generated alarms that were evaluated as an expression of the predictive performance of the model versions.

3. Evaluating model performance

3.1. Events of interest

The events of interest in this study were diarrhea and pen fouling amongst growing pigs. Both diarrhea and pen fouling, which is a change in behaviour where the pigs start to lie on the slatted area of the pen and excrete in the lying area, reduce productivity and animal welfare (Aarnink et al., 2006; Pedersen, 2012).

Every morning, the caretakers at each farm registered whether either of the two events occurred in a pen or not. The routines for assessment of either event were described in a project protocol (Lyderik et al., 2016), and were calibrated by an experienced technician regularly throughout the study period.

If considered necessary by the caretaker, pigs with diarrhea were treated with antibiotics, and pens where fouling had occurred were cleaned. Event registration and treatments were conducted once a day, but because the actual outbreak of the event could happen at any hour between two registrations, an event was defined to last 24 h from midnight to midnight in the present study.

The objective of this paper is, however, to evaluate model performances at different spatial levels of a pig production herd rather than the ability to distinguish between specific conditions. Therefore registrations of both diarrhea and fouling were joined under the common term “event”.

Merging the two types of events is supported by Madsen et al. (2005) and Andersen et al. (2016), who state that changes in drinking patterns can reflect changes in the general wellbeing of pigs. This implies that changes in drinking patterns may not be uniquely related to a specific type of event.

Despite regular calibration of registration routines, significant herd-specific differences in the frequency of event registrations occurred, and two different event definitions were used: In Herd A the daily caretaker was replaced with unexperienced personnel a number of times during

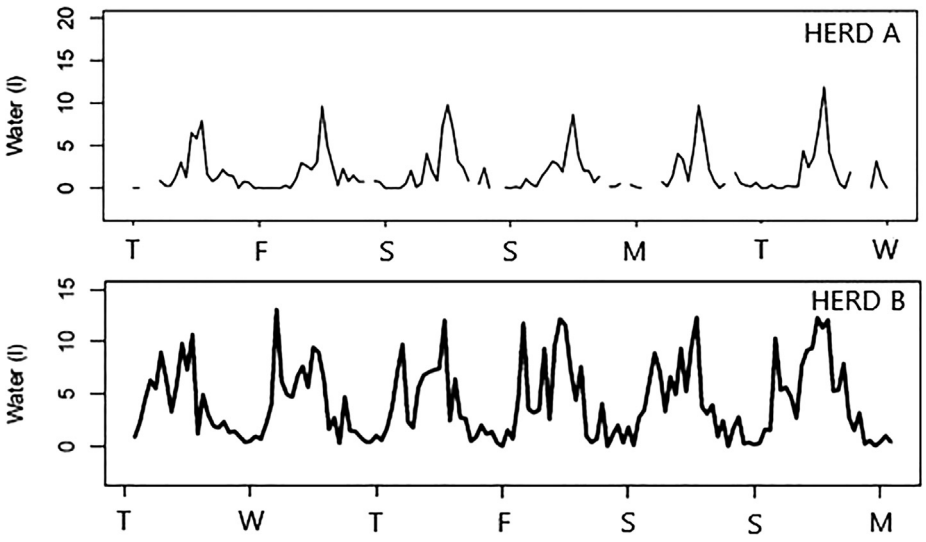

Fig. 2. Diurnal drinking pattern of finishers (Herd A) and weaners (Herd B).

the period of data collection. As a consequence of that, the commitment to register daily events was inconsistent and resulted in periods with no registrations. For performance evaluation on Herd A data, all event registrations available for the herd constituted the gold standard.

In Herd B the threshold for identification of diarrhea was low. This lead to multiple periods with registrations of diarrhea every day for 14–21 days, although only few or no interventions were made during those periods. For performance evaluation on Herd B data, the initiation of an intervention (medical treatment of diarrhea or cleaning of pens with fouling), rather than daily event registrations, constituted the gold standard.

3.2. Time window

By comparing alarms and events occurring at the same moment, the alarms can be categorized true or false, and the performance of the model can be calculated (Hogeveen et al., 2010). But alarms seldom occur at the exact same moment as the events, and if they did, they were of little predictive value, leaving insufficient time to implement preventive interventions. Therefore *time windows* are often used (Hogeveen et al., 2010; Ostersen et al., 2010; Jensen et al., 2017). A time window is a defined period of time associated with a registered event, and any number of alarms occurring within that window are treated as one single alarm, and categorized as detecting the event correctly (Fig. 4).

Time windows can be of varying lengths, and may extent from before an event to after an event (de Mol et al., 1997; Jensen et al., 2017). As Fig. 4 illustrates, the length of a time window has great influence on the categorization of true or false alarms, and therefore on the performance of a model. Longer windows improve model performance, whereas windows extending beyond an event can result in alarms being communicated after the event has occurred. The categorization of alarms as true or false are counted as follows:

- Alarms within a time window are counted as one *true positive* (TP).
- Alarms occurring outside of a time window are counted as *false positive* (FP).
- If no alarms occur within a time window, it is counted as *false negative* (FN).
- Days without alarms and with no time window are counted as *true negative* (TN).

The detection accuracy can then be expressed by *sensitivity* (*Se*) and *specificity* (*Sp*), which are calculated as:

$$Se = \frac{TP}{(TP + FN)} \tag{1}$$

and

$$Sp = \frac{TN}{(TN + FP)} \tag{2}$$

where *TP* denotes the total number of TP cases and accordingly for the other variables.

Three lengths of time windows were applied for the performance evaluation in this paper (see Fig. 4). The longest window included three days before an event plus the day of the event, but zero days after. The two other window lengths include two days and one day before an event respectively plus the day of the event, but none after. The three windows were denoted (3/0), (2/0), and (1/0) respectively, following the terminology of Jensen et al. (2017).

3.3. Standardized two-sided CUSUM

In a Cusum, the deviations from the mean, μ_0, are accumulated over time, and when the sum of accumulated deviations exceeds a defined threshold, the process is considered out of control and an alarm is generated (Montgomery, 2013).

The inputs to the Cusum in this study were series of forecast errors, e_t generated by the DLM. For a pen it was simply the series of forecast errors from the sensor in the corresponding pen (8 in Herd A, 16 in Herd B), whereas the series of forecast errors for a Section (4 in Herd A and in Herd B) was generated by adding the forecast errors of all sensors at time t within the specific section together.

The series of forecast errors for the herd (1 in Herd A and in Herd B) was likewise generated by adding the forecast errors of all sensors in the herd at time t together. In case of missing data at time t, the value of the corresponding forecast error was set equal to zero. Thus, if e_t denotes the full vector of forecast errors at time t, the scalar forecast error e_t^u for the unit u (a specific pen, a specific section or the entire herd) is found as

$$e_t^u = I_u e_t, \tag{3}$$

where I_u is a row vector only consisting of zeros and ones. If u is a specific pen, it means that I_u is a row vector with the element 1 at the position of u in e_t. Accordingly, if u is a section, I_u will have ones at the positions corresponding to pens in the section in question and zeros elsewhere.

The series of forecast variances, Q_t^u, for a given unit were calculated according to standard rules as

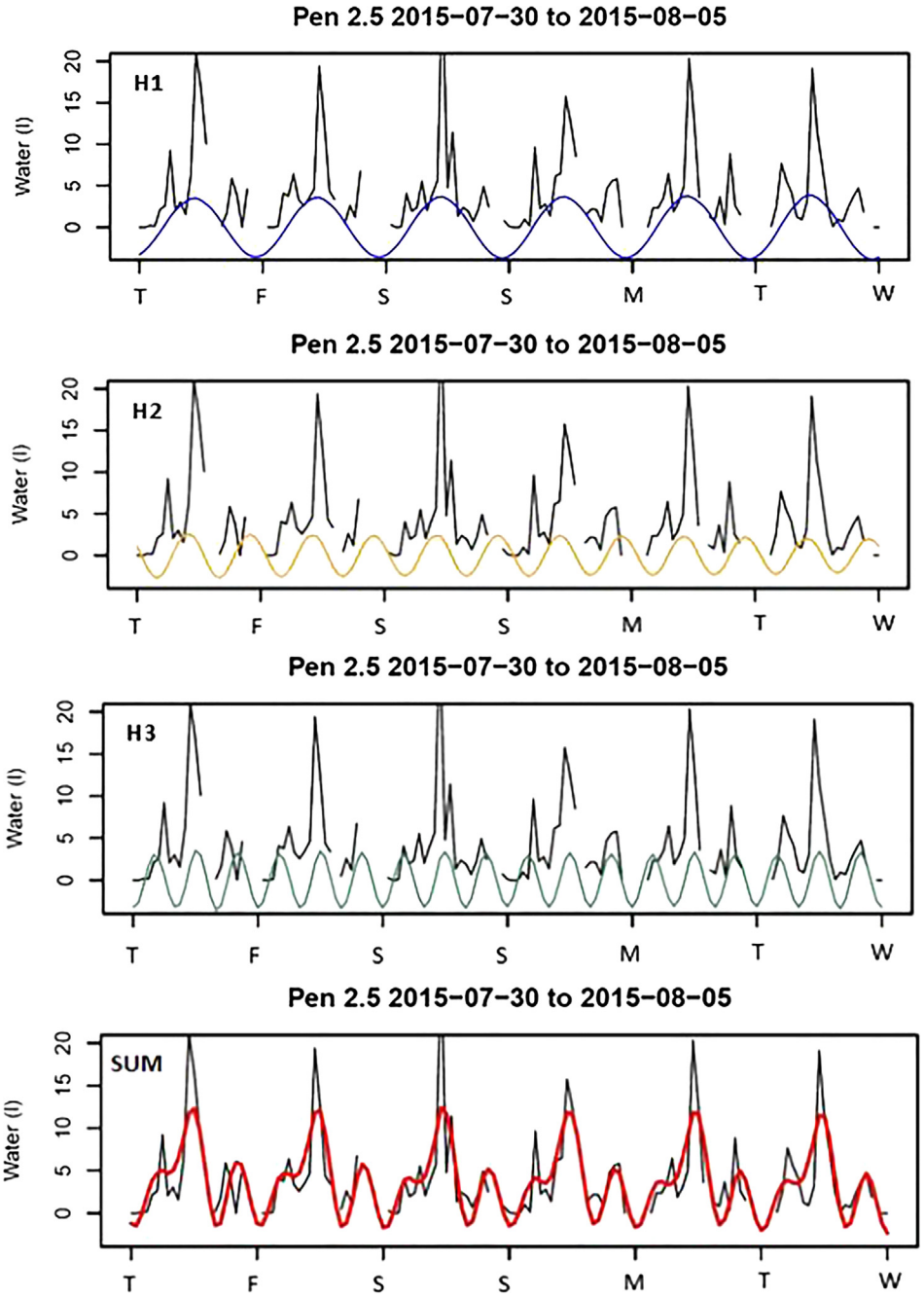

Fig. 3. The diurnal drinking pattern (black line) is shown together with the three harmonic waves; 24 h (H1), 12 h (H2), and 8 h (H3). The sum of the three harmonic waves and the underlying level (which is not depicted) is shown in (SUM). From Dominiak et al. (2018).

$$Q_t^u = I_u Q_t I_u'. \tag{4}$$

In case of missing data at time t, the value of the corresponding forecast variance was set equal to 1. The cumulated sum of the Cusum was reset when an alarm had been generated, and since the test data for both herds covered the length of two batches, the cumulated sum was also reset at the beginning of the second batch.

In the two-sided Cusum, the forecast errors above the mean (zero) are summed separately as *upper Cusum*, and the forecast errors below the mean are summed separately as *lower Cusum*. This two-sided Cusum allows for different interpretations of alarms caused by water consumption higher than expected and lower than expected. Since the underlying level of water consumption increases as the pigs grow, the numerical values of the forecast errors increase as well (Madsen et al., 2005). In order to distinguish between the growth-related increase and increases caused by the process being out of control, the forecast errors are standardized, and a *Standardized* two-sided Cusum control chart is applied, as described by Montgomery (2013).

Since the expected value of e_t^u is 0, the standardized value y_t^u simply becomes

$$y_t^u = \frac{e_t^u}{q_t^u}, \tag{5}$$

where $q_t^u = \sqrt{Q_t^u}$.

Then, the **Upper Cusum** for the unit is the series

$$C_t^{u+} = max\,[0, y_t^u - k + C_{t-1}^{u+}] \tag{6}$$

and the **Lower Cusum** is the series

Table 2
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. 3). 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. From Dominiak et al. (2018).

LG	H1 H2 H3	Interpretation
S	HHH	The full harmonic pattern evolves identically for all pens in the herd
S	HSP	H1 evolves identically for all pens, H2 evolves identically within sections but differently between sections, H3 evolves differently in each pen
S	HSS	H1 evolves identically for all pens, H2 and H3 evolve identically within each section but differently between sections
S	SSS	The full harmonic pattern evolves identically within each section but differently between sections
S	SSP	H1 and H2 evolve identically within sections but differently between sections, H3 evolves differently in each pen
S	SPP	H1 evolves identically within sections but differently between sections, H2 and H3 evolve differently in each pen
S	PPP	The full harmonic pattern evolves differently in each pen

$$C_t^{u-} = max\,[0,\ -k - y_t^u + C_{t-1}^{u-}]. \tag{7}$$

where k is the *reference value*. The reference value allows for a constant level of *slack* or *allowance* to be accepted as an integrated part of the system and it is subtracted from y_t before the summation. The value of k is traditionally chosen relative to the size of the shift to be detected (Montgomery, 2013).

In addition to the reference value, a *decision interval*, or a threshold, h, must be chosen as well. If either C_t^{u+} or C_t^{u-} exceeds the threshold, the process is considered to be out of control, and an alarm is generated. Montgomery (2013) recommends h to be defined at fixed values of 4 or 5 for a standardized Cusum.

Choosing the right settings of the threshold value, h, and the reference value, k, of the Cusum are essential to the number of alarms generated. Therefore the optimal combination of h and k for each vector of forecast errors is chosen by iterations over sequences of h and k values, and a Cusum is run for each generated combination of h and k. Threshold values were iterated from 0 to 5 and reference values were iterated from 0 to 2.

In the supplementary material, the effects of different settings of the Cusum parameters are illustrated.

3.3.1. Evaluating spatial levels

Evaluation of model performance was done for each of the seven model versions (Table 2) separately on data from Herd A and Herd B. All model versions were evaluated for their ability to predict the occurrence of events of interest at either of the three spatial levels; pen level (in a specific pen), section level (in a specific section within the herd), or herd level (in any pen within the herd) using three different lengths of time windows.

Days with events at pen level were the days when events were registered in the pen by the caretakers. Days with events occurring at section level were all days with minimum one event registered in any pen within the section. Thus, if events were registered in two or more pens within the same section at the same day, they count as one event-day at section level. Days with events occurring at herd level were all days where minimum one event was registered in the herd. If events were registered in two or more pens in the herd at the same day, that day counted as one event-day at herd level.

When evaluating spatial level performance, a total of $2 \times 7 \times 3 \times 3 = 126$ model combinations were evaluated based on the

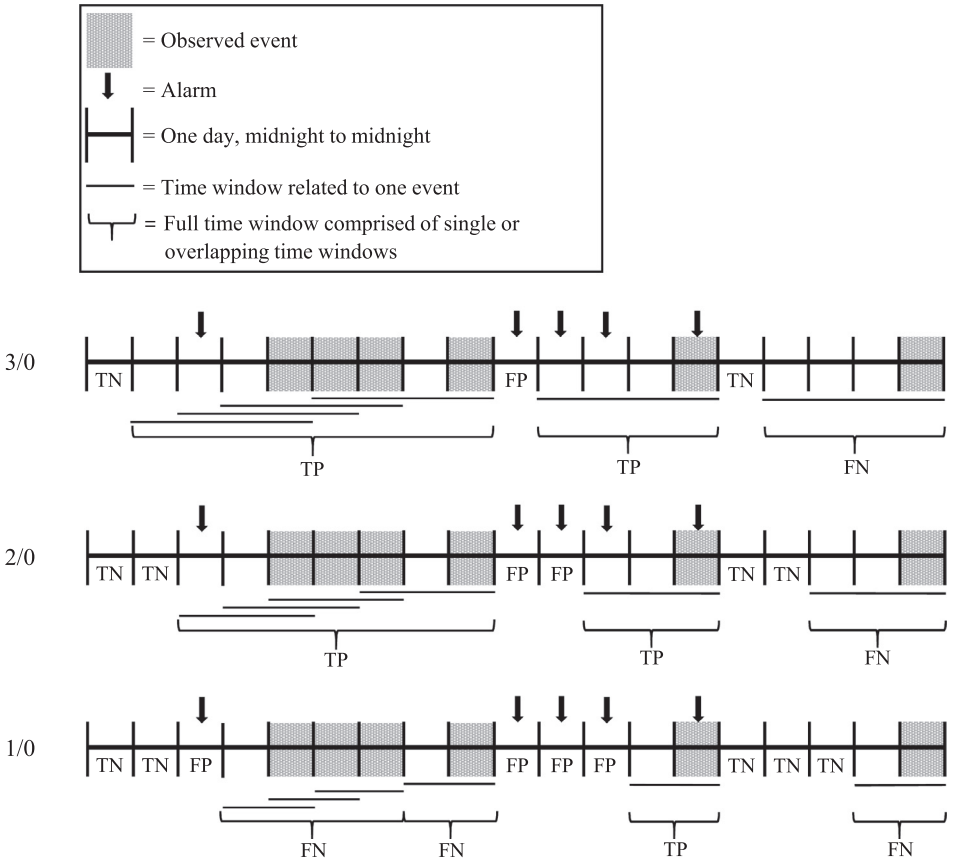

Fig. 4. Example of definitions of true positives (TP), false positives (FP), true negatives (TN), and false negative (FN). All observed events are associated with a time window, and overlapping time windows are merged into longer windows. Three lengths of time windows are illustrated; 3/0 = three days before an event and zero days after, 2/0 = two days before an event and zero days after, 1/0 = one day before an event and zero days after. All alarms occurring within a time window are counted as one TP alarm. If no alarms occur within a time window, it is counted as one FP. Days outside of time windows but with alarms, are counted as FP, whereas days outside of time windows with no alarms are counted as TN. Based on illustration by Jensen et al. (2017).

following:

- Herd (Herd A, Herd B).
- Model version (temporal correlation) (HHH, HSP, HSS, SSS, SSP, SPP, PPP).
- Spatial level (Pen, Section, Herd).
- Time Window (3/0, 2/0, 1/0).

A Cusum was run on each series of standardized forecast errors within the level for all of the 126 model combinations per $h \times k$ combination, and the performance was calculated on the pooled outputs of these Cusums per setting. As an example, let n be the number of units (pens or sections) at the spatial level in question and let the *setting*, $s = (h, k)$, be a unique combination of threshold and reference value. Each Cusum, with a unique setting, s, was run in turns on standardized forecast errors from all units at the spatial level (number of pens in a section or number of sections in a herd). In order to obtain the overall performance of the Cusum setting, the four classification categories (TP, FP, TN, FN) were counted across units as follows:

- $TP_s = TP_{s1} + TP_{s2} + ...+TP_{sn}$
- $FP_s = FP_{s1} + FP_{s2} + ...+FP_{sn}$
- $TN_s = TN_{s1} + TN_{s2} + ...+TN_{sn}$
- $FN_s = FN_{s1} + FN_{s2} + ...+FN_{sn}$,

where e.g. TP_{s1} is the number of true positives for unit 1 under setting s.

Then the conditional prediction accuracy was calculated for each s in terms of sensitivity (Se_s) and specificity (Sp_s) as follows:

$$Se_s = \frac{TP_s}{(TP_s + FN_s)} \tag{8}$$

and

$$Sp_s = \frac{TN_s}{(TN_s + FP_s)} \tag{9}$$

In other words; for each pen, when evaluating pen level performance, or section, when evaluating section level performance, a Cusum was run for each of the combinations of h and k. Then the outputs from all pens, or sections, were pooled for each individual h and k combination, and the performance was calculated for each pooled output. Hereby $h \times k$ performances were generated, each based on outputs from all pens or sections in the herd (see Fig. 5).

For each spatial level of each model version, the performance parameters Se_s and the false positive rate $FPR_s = 1 - Sp_s$ were plotted against each other so that each setting produced an observation in the diagram as shown in Fig. 5. For a given value of FPR_s, the best possible Se_s is desired, implying that the *Receiver Operation Characteristic* curve (ROC curve) was identified by connecting observations that, for each value of FPR_s, maximize Se_s. Thus, the ROC curve is a nondecreasing function of FPR_s.

Table 3
AUC (area under curve) for prediction of events at herd level (in any pen in the herd) in Herd A and in Herd B with three different lengths of time windows applied. 3/0 time window covers tree days before the event and zero days after the event, 2/0 time window covers two days before the event and zero days after the event, 1/0 time window covers one day before the event and zero days after the event. AUC for seven model versions is presented for both Herd A (commercial finishers) and Herd B (research centre weaners). Sensors were evenly distributed between four sections in each of the herds with two sensors per section in Herd A and four sensors per section in 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		Herd A			Herd B		
LG	H1 H2 H3	3/0	2/0	1/0	3/0	2/0	1/0
S	HHH	0.9358	0.9194	0.8013	0.9842	0.9734	0.8878
S	HSS	0.9014	0.8694	0.8287	0.5789	0.6087	0.5369
S	HSP	0.8600	0.8796	0.8307	0.7737	0.7280	0.6761
S	SSS	0.8972	0.8836	0.8083	0.8105	0.8720	0.8438
S	SSP	0.8736	0.8604	0.8052	0.9368	0.9614	0.8473
S	SPP	0.9274	0.9194	0.8036	0.8316	0.8316	0.8253
S	PPP	0.9283	0.9148	0.8090	0.8316	0.8816	0.8395

As the final measure of the predictive performance, as a measure of test accuracy, the Area Under Curve (AUC) was calculated. An AUC = 1 indicates perfect predictive performance, so values close to 1 were preferred. The AUC was calculated in R (R Core Team, 2014) using the function “trapz” from the library “pracma”.

4. Results and discussion

The AUC of the 126 different model combinations can be seen in Tables 3–5. The predictive performance is in general higher at herd level and decreases as the level gets more detailed, which is illustrated in Fig. 6 for model version HHH and time window 3/0. The results also show that the AUC in general is higher when the longer time window (3/0) is used, and decreases as the time window gets shorter.

The overall best predictive performances is reached for Herd B, model version HHH at herd level for time windows 3/0 (AUC = 0.9842). For Herd A, the highest predictive performance is reached by model version HHH with time window 3/0 at herd level (AUC = 0.9358).

Due to the compromised quality of the event registrations in Herd A, the number of false positive alarms may be overestimated. On the other hand, a number of alarms, either true or false, which might have occurred during periods of sensor outages, are not included in the evaluation of the model (because they are not registered) and this may have affected the results as well.

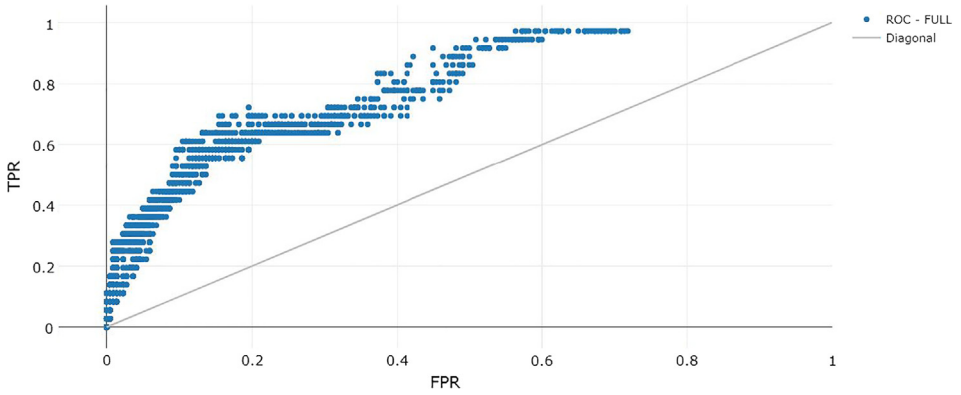

Fig. 5. Illustration of a ROC curve from a section in Herd B with time window 3/0. The prediction accuracy is plotted for each $h \times k$ combination.

Table 4
AUC (area under curve) for prediction of events at section level (in a specific section) in Herd A and in Herd B with three different lengths of time windows applied. 3/0 time window covers tree days before the event and zero days after the event, 2/0 time window covers two days before the event and zero days after the event, 1/0 time window covers one day before the event and zero days after the event. AUC for seven model versions is presented for both Herd A (commercial finishers) and Herd B (research centre weaners). Sensors were evenly distributed between four sections in each of the herds with two sensors per section in Herd A and four sensors per section in 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		Herd A			Herd B		
LG	H1 H2 H3	3/0	2/0	1/0	3/0	2/0	1/0
S	HHH	0.8882	0.8708	0.8144	0.8715	0.8576	0.7705
S	HSS	0.8592	0.8339	0.8135	0.7193	0.6789	0.6444
S	HSP	0.8616	0.8345	0.8105	0.7711	0.7280	0.6850
S	SSS	0.8647	0.8405	0.8084	0.8205	0.8008	0.7641
S	SSP	0.8611	0.8324	0.8098	0.8635	0.8563	0.8020
S	SPP	0.8757	0.8524	0.7959	0.8375	0.8085	0.7643
S	PPP	0.8631	0.8382	0.7825	0.8311	0.8085	0.7667

Table 5
AUC (area under curve) for prediction of events at pen level (in a specific pen) in Herd A and in Herd B with three different lengths of time windows applied. 3/0 time window covers tree days before the event and zero days after the event, 2/0 time window covers two days before the event and zero days after the event, 1/0 time window covers one day before the event and zero days after the event. AUC for seven model versions is presented for both Herd A (commercial finishers) and Herd B (research centre weaners). Sensors were evenly distributed between four sections in each of the herds with two sensors per section in Herd A and four sensors per section in 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		Herd A			Herd B		
LG	H1 H2 H3	3/0	2/0	1/0	3/0	2/0	1/0
S	HHH	0.8878	0.8701	0.8164	0.7671	0.7348	0.6871
S	HSS	0.8599	0.8424	0.8137	0.6468	0.6320	0.6109
S	HSP	0.8598	0.8422	0.8154	0.6747	0.6459	0.6234
S	SSS	0.8583	0.8350	0.8102	0.7535	0.7309	0.7035
S	SSP	0.8585	0.8408	0.8129	0.7682	0.7401	0.6969
S	SPP	0.8782	0.8634	0.8087	0.7750	0.7555	0.7208
S	PPP	0.8644	0.8500	0.7975	0.7671	0.7490	0.7454

The risk of compromised quality of the gold standard may be higher when it is observed by personnel in commercial herds, than when it is observed in more controlled study designs. However, a detection model, which is based on a study design highly similar to real life production conditions, is better suited for future implementation in commercial herds, as discussed by Hogeveen et al. (2010).

4.1. Herd level

Several of the herd level performances (AUC > 0.92) in Table 3 indicate prediction accuracy close to perfect. It is, however, worth remembering that any event at any day within the herd is included when evaluating performances at herd level. Furthermore, overlapping time windows are merged into one window lasting from the first day of the first window to the final day of the last window. This means that for Herd B, a total of 5 time windows (longest = 47 days) cover 106 of 126 days in the test data when window length 3/0 is used.

For Herd A, a total of 10 time windows (longest = 20 days) cover 79 days out of 172 days in the test data when using window length 3/0. The combination of a few long time windows and a few days outside any time windows affects the outcome of the Se_s and Sp_s leaving only few points in the ROC curve. Although the impressive predictive performances are correct, the settings they represent are of little value for the manager in the everyday production.

The main reason for the relatively small managerial value is that an alarm at Herd level is associated with any event in any pen in the herd within the given time window. The manager therefore gets no information on which area (pen or section) is at risk of an outbreak. In addition the lengths of the merged time windows reduce the practical value of a Herd level alarm since the event may occur at any day within the time window, and it covers up to 20 days.

As described in Section 3.3 a Herd level alarm is based on the cumulated sum of forecast errors generated by all sensors in the herd, and a section level alarm is based on the cumulated sum of forecast errors generated by all sensors within the same section. If a Herd level alarm occurs at the exact same time (same hour) as a section level alarm, it may therefore be caused by a very large forecast error from that specific section. Such a Herd level alarm can therefore be used to prioritize the simultaneously occurring section level alarm above those section level alarms which do not raise simultaneous alarms at Herd level.

Such a use of Herd level alarms to prioritize certain section level alarms above other section level alarms is highly relevant for the manager when deciding which alarm to attend to first.

4.2. Section and pen level

The prediction accuracies for Herd A at both section level and pen level are high (AUC = 0.83–0.89 for the two longer time windows) and almost identical with respect to model versions and time windows (see Tables 4 and 5). This indicates that events are registered on the same days in all pens within the same section, and that changes in drinking patterns occur at the same time for all pens within the same section. When events are registered on the same day in all pens within the same section, it may be due to a contagious disease, like diarrhea, affecting multiple pens in the section at the same time. The finding of such a correlation meets the initial expectations for this study.

It should, however, be noted that numerous longer periods of missing data throughout the test data set of Herd A may have reduced any differences in drinking patterns between pens. Herby promoting similarity between pens and sections. Running the model on data from another herd, or redefining learning and test data in the present data set, is needed to confirm this.

The prediction accuracies for Herd B differ between section level and pen level (see Tables 4 and 5). For section level, the accuracies are high (AUC > 0.80 for the two longer time windows) for five of the seven model versions, whereas all AUC's are lower than 0.70 at pen level.

Few and small pigs are monitored in Herd B (15 weaners, 7–30 kg). An irregular drinking pattern from just a single pig in a single pen therefore has larger effect on the water consumption in the pen. Differences between pens are thereby easier generated, and this may be the cause of the different AUC's at pen and section level.

At all times, the manager knows the age and size of pigs in any section. Furthermore an experienced manager is aware of high-risk periods for conditions such as diarrhea and pen fouling. This means that the managerial value of alarms at both section level and pen level is very high. A section level alarm informs the manager which section needs extra managerial focus. Combined with a pen level alarm from the same section, the manager gets more specific information on where in the section the outbreak of an event has originated, or is more severe. A spatial alarm therefore supports the manager in choosing the right intervention for the targeted area timely enough to prevent or reduce an outbreak of an event.

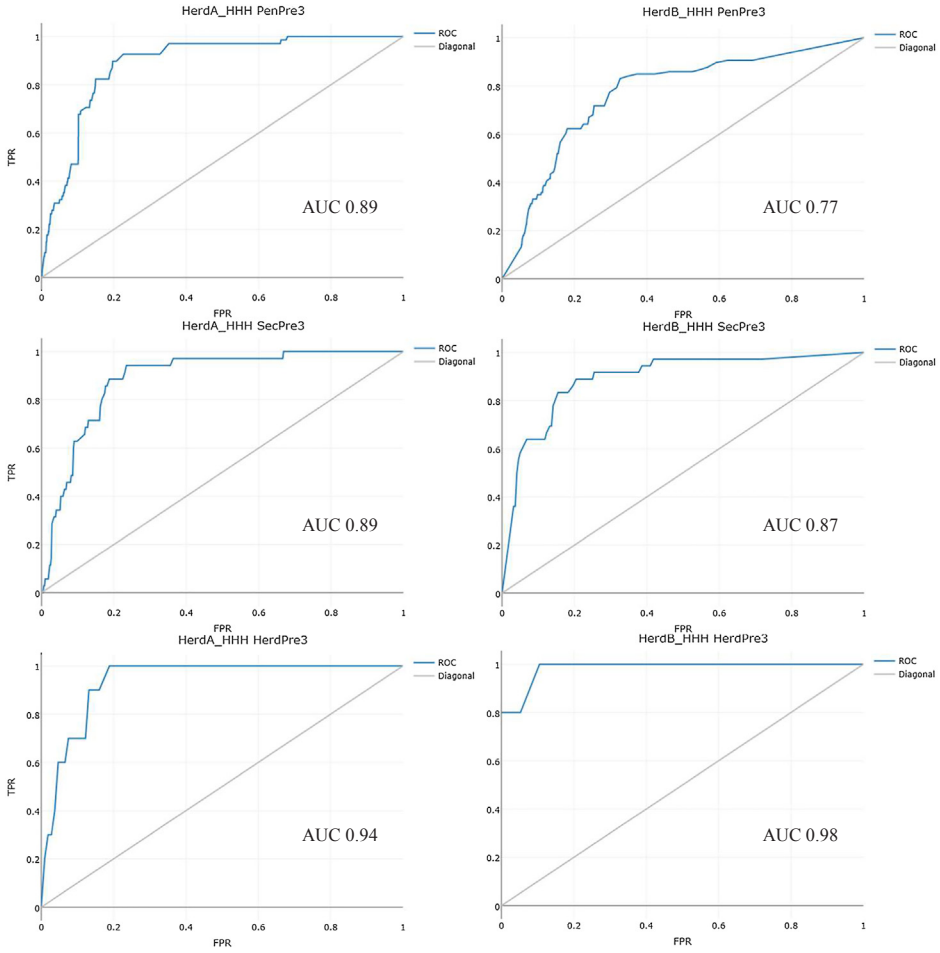

Fig. 6. ROC curves and corresponding AUC from Herd A (left) and Herd B (right), HHH model version and 3/0 time window.

4.3. Time windows

Longer time windows yield higher performances for both herds and all model versions. Although an alarm three days before an event (3/0 window) may be too long for the precise timing of managerial interventions, an alarm two days ahead (2/0 window) might be sufficient in many situations, especially if the predictive accuracy is higher than when shorter windows are applied.

When evaluating the 1/0 window performances for Herd A, the HHH model version (expressing temporal correlations between drinking patterns from all pens in the herd, see Table 2) is able to predict an event in a specific pen with a fairly high predictive accuracy (AUC = 0.8164). Even though the HSP model version (expressing a lower degree of temporal correlation between drinking patterns in the herd, see Table 2) predicts events with a higher accuracy (AUC = 0.8307), this is only obtained at herd level. As discussed above, the herd level is a very general spatial level at which an alarm contains little value in daily management. For Herd B, the highest AUC, given the 1/0 time window, is reached by the HHH model version at all spatial levels. Thus, herd level reaches the highest accuracy (AUC = 0.8878), and then the accuracy is reduced for both section level (AUC = 0.8020), and pen level (AUC = 0.7208).

Since all alarms within a time window detect an event correctly, as described in Section 3.2, there is no information on whether an alarm occurs in the beginning, middle or end of a time window. Based on the results in Tables 3–5, the longer time window (3/0) is the more accurate, closely followed by the 2/0 time window. However, considering the lack of precision *within* the time window, the shorter of the two (2/0) would provide more precise information, and therefore higher managerial value.

4.4. Model versions

The HHH model version provides the highest AUC for predicting events at all spatial levels in Herd A. This indicates that finisher pigs across Herd A show peaks in their drinking pattern at the same time of day (see Table 2) throughout the entire growing period. The HHH model version is presenting the poorest fit of the seven versions on Herd A data when developing the model, whereas the SSS model version in general has poor performance in terms of AUC, but obtained the best fit to data (Dominiak et al., 2018).

No single model version provides the highest AUC across levels in Herd B. For predictions of events at herd level, the HHH model version has the highest accuracy at all time windows, but for predictions at section level and pen level, the model versions with harmonic waves defined at section and pen level provide high accuracies as well.

There is a remarkably clear connection between the level of the harmonic waves in the model versions and the level where events are predicted with the highest accuracy for time window 1/0 in the way that the HHH model version predicts best at herd level (AUC = 0.8878), the SSP model version predicts best at section level (AUC = 0.8020), and the PPP model version predicts best at pen level (AUC = 0.7454).

As for Herd A, it is the HHH model version which has the poorest fit in the DLM and the best prediction accuracy when evaluated on test data and events for Herd B. The SSP version fit the test data better, but it is the PPP version, which fit the best in Dominiak et al. (2018).

An explanation of this inverse relation between fit and prediction accuracy may be that the models with better fit end up overfitting the

training data. Over-fitting models tend to model random noise as well, causing the model versions with higher complexity to make worse predictions, when trying to include random noise in the modeled pattern (Fortmann-Roe, 2012).

It may also be, that models with better fit at the same time have a higher adaptability to changes and irregularities. Instead of generating alarms, a well fitting model will adjust to the changes and accept them as a part of the pattern.

4.5. Ensemble classifying methods

An *ensemble classifier* combines the output of different models and often increase predictive performance over a single model (Witten and Frank, 2005). Alarms communicated from an ensemble are often considered more valid than alarms from individual model versions.

In order to improve predictive performance, the two ensemble classifying methods, *bagging* and *boosting* are tested on the seven model versions for each herd individually. Both methods are machine learning methods, and they combine the decisions of different models by amalgamating the outputs into a single prediction (Witten and Frank, 2005). Kamphuis et al. (2010) applied both bagging and boosting to decision trees in a clinical mastitis detection model. They found bagging to give the better results.

The bagging method lets all model versions vote whether an alarm should be generated or not, on a daily basis. A defined threshold states how many models should agree, and if the threshold is reached or exceeded, the ensemble generates an alarm. The boosting method works on the same principles, only the votes are weighted according to, for example, the performance of each model version. In this test, the specificity of each model was used as weighting factor in the voting under the boosting method.

In our study, neither bagging nor boosting improved the AUC when compared to the AUC of the best of the individual model versions. Thus, the improvement seen in the study by Kamphuis et al. (2010) is not seen here. An obvious reason could be that all seven model versions are based on exactly the same data, and furthermore have many structural similarities. The seven tests can therefore not be seen as independent, which means that the output of the seven model versions neither supplement nor complement each other, and therefore no improvement is found by applying neither of the ensemble methods.

5. Future perspectives

In this section future applications and perspectives of the detection model are presented.

5.1. Alarm prioritizing method

Some alarms occur at the same time t in a pen and in the corresponding section as illustrated in Fig. 7. If the apparent connection between changes in drinking patterns and general wellbeing (Madsen et al., 2005; Andersen et al., 2016) is accepted, then such alarms should be considered true, independently of event registrations. Such an alarm is either caused by a very large deviation in a single pen, caused by for instance a broken water nipple or very sick pigs in that specific pen, or by relatively smaller unidirectional errors in more pens, which could be caused by a simultaneous outbreak of a disease in several pens in the section. Both types of scenarios are severe for both animal welfare and productivity, and therefore an alarm occurring at the same time, t, in a pen and the corresponding section should always be given high priority.

5.2. Alarm reducing method

Alarms from multiple pens within the same section on the same day will be merged and communicated as one alarm for the section rather than multiple individual pen level alarms. This method reduces the number of alarms communicated to the manager. Although the method to some extent devaluate pen-specific information, there is a managerial value in section-specific alarms due to the sectionalized structure in the pig producing units as a whole.

5.3. Alternative post processing methods

DLMs have been used in several previous studies with the purpose of detecting undesired events. The general procedure has been to fit a (univariate or multivariate) DLM to data and, afterwards, to produce series of forecast errors which, in a second step, are post processed in order to produce warnings.

Several different post processing methods have been used previously. Jensen et al. (2017) used a threshold for the Mahalanobis distance (found by Cholesky decomposition of the forecast variance–-covariance matrix) between the multivariate forecast error and the zero vector. In another multivariate study, Jensen et al. (2016) used a Naïve Bayesian Classifier and in Jensen and Kristensen (2016), artificial neural networks were used for post processing.

In univariate studies (Madsen and Kristensen, 2005; Cornou et al., 2008) and studies where a multivariate observation has been transformed to a univariate response (Bono et al., 2012, 2013, 2014) a Cusum in combination with a V-mask (Montgomery, 2013) has been a popular tool for detection of gradual changes in the observed pattern of data. For sudden changes a simple Shewhart control chart (Montgomery, 2013) applied to the forecast errors has often been used (Bono et al., 2012, 2013, 2014; Cornou et al., 2014).

The post processing method used in this study has been the tabular Cusum with various settings but as illustrated by the overview above, many other options exist. It could be argued that using a multivariate approach, and then later only use univariate Cusums for detection of events, considerably reduces the spatial information available in the model. Thus, it would be interesting for future research to study alternative post processing methods.

A first step could be to distinguish alarms generated by the upper Cusum from those generated by the lower Cusum. In case of diarrhea, for instance, an increased water consumption is assumed (Madsen and Kristensen, 2005) and, accordingly, the upper Cusum might generate an alarm. Other disturbances leading to decreased water consumption might produce an alarm generated by the lower Cusum. A more sophisticated approach would be to use a structured Bayesian network for post processing of the forecast error vectors and classify them according to presence or absence of events. That is, however, outside the scope of this article.

5.4. Implementation considerations

Since the alarms from the presented detection system are nonevent-specific, the manager has to add herd-specific knowledge when responding. Keeping in mind, that changes in water consumption may reflect changes in the general wellbeing of the pigs, or other conditions than those of interest in this study, an alarm should be interpreted as “something is wrong” rather than as an alarm for e.g. diarrhea.

The high prediction accuracies obtained in this study implies that the alarms are very trustworthy when pointing out an area. In other words, when a manager gets an alarm targeting a specific pen or section, it is very likely that an event is occurring in that specific area within the next two or three days. This enables the manager to go to the pointed area and check for initial signs of reduced health or welfare amongst the animals, as well as for any signs of problems with climate control, feeding system, and water system.

The high area-specific accuracies make the spatial detection system well suited for implementation in a commercial herd. It is, however, recommended to perform extern validation of the model on data from independent herds in order to confirm the results.

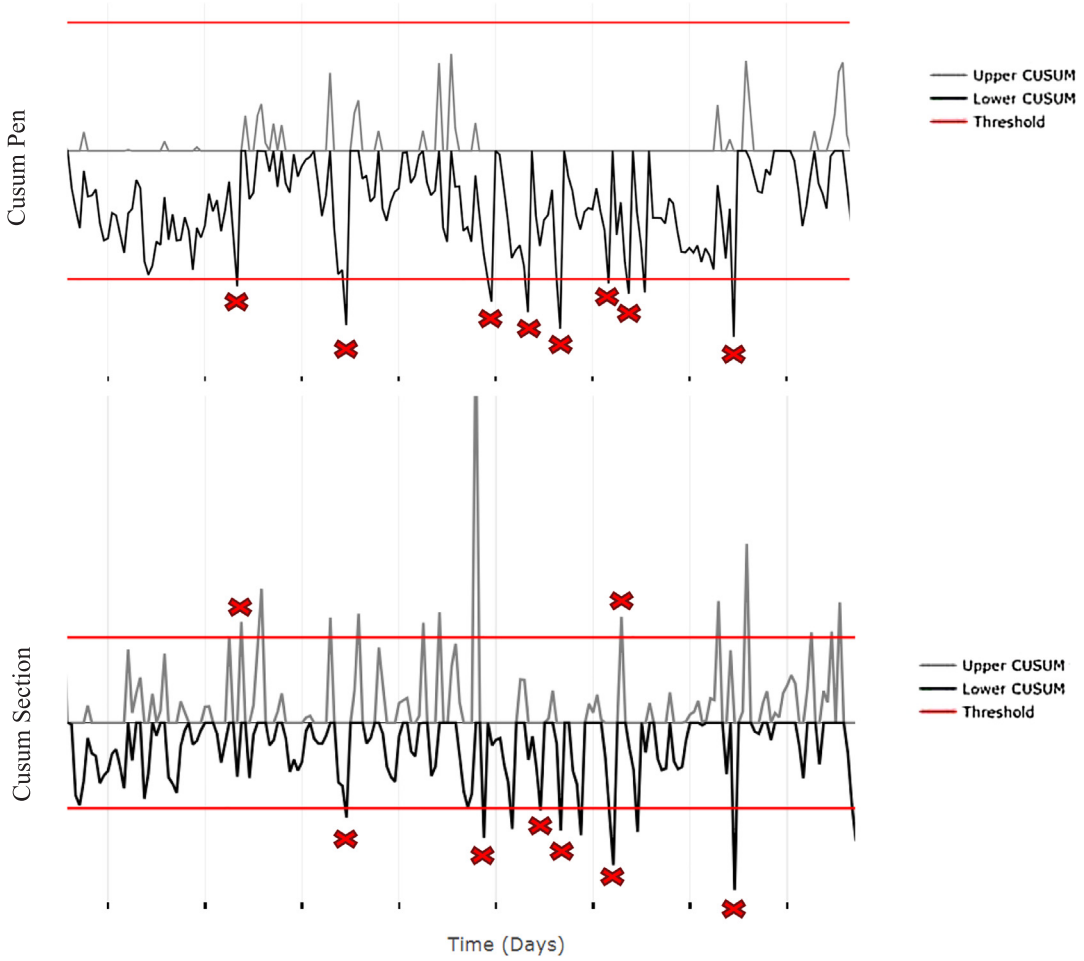

Fig. 7. Example of a Cusum from a pen (top) and the corresponding section (bottom) with simultaneous alarms. Alarms marked with an X occur at the exact same hour in the pen as in the section.

6. Conclusion

The new spatial approach, presented in this paper, makes it possible to predict events at separate spatial levels in herds of growing pigs. The model version expressing highest temporal correlation in drinking patterns between pens and sections in a herd (HHH) has the poorest fit, but tend to predict outbreaks of unwanted events better.

Longer time windows and predictions at herd level yield very high predictive accuracies, but alarms communicated at herd level are of little or no value in a commercial production herd due to very long overlapping time windows and non-specific spatial identification of events.

The predictive accuracies for identifying events in a specific pen or section are high (AUC > 0.80), and given the 2/0 time window the multivariate spatial DLM constitute a new and promising approach to sensor based monitoring tools in livestock production.

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.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.compag.2018.10.037.

References

Aarnink, A.J.A., Schrama, J.W., Heetkamp, M.J.W., Stefanowska, J., Huynh, T.T.T., 2006. Temperature and body weight affect fouling of pig pens. J. Animal Sci.

Andersen, H.M.L., Jorgensen, E., Pedersen, L.J., 2016. Using evolutionary operation technique to evaluate different management initiatives at herd level. Livestock Sci. 187, 109–113.

Anonymous, 2000. Instruction leaflet flow sensors - 15 mm dia. Pipe v8189 05/2000. < http://docs-asia.electrocomponents.com/webdocs/001b/0900766b8001bb47.pdf > .

Aparna, U., Pedersen, L.J., Jorgensen, E., 2014. Hidden phase-type markov model for the prediction of onset of farrowing for loose-housed sows. Comput. Electron. Agric. 108, 135–147 Update Code: 20141203.

Bono, C., Cornou, 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 update Code: 20121205.

Bono, C., Cornou, C., LundbyeChristensen, S., Kristensen, A.R., 2013. Dynamic production monitoring in pig herds ii. Modeling and monitoring farrowing rate at herd level. Livestock Sci. 155 (1), 92–102 update Code: 20130731.

Bono, C., Cornou, C., LundbyeChristensen, S., Kristensen, A.R., 2014. Dynamic production monitoring in pig herds iii. Modeling and monitoring mortality rate at herd level. Livestock Sci. 168, 128–138 Update Code: 20141119.

Cornou, C., Ostergaard, S., Ancker, M.L., Nielsen, J., Kristensen, A.R., 2014. Dynamic monitoring of reproduction records for dairy cattle. Comput. Electron. Agric. 109, 191–194 Update Code: 20150121.

Cornou, C., Vinther, J., Kristensen, A.R., 2008. Automatic detection of oestrus and health disorders using data from electronic sow feeders. Livestock Sci. 118 (3), 262–271 update Code: 20080000.

Danish Agriculture and Food Council, 2010. Danish pig producers and animal welfare.

Tech. rep., Danish Agriculture and Food Council.
de Mol, R.M., Kroeze, G.H., Achten, J.M.F.H., Maatje, 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., Woldt, W.E., 2001. Application of fuzzy logic in automated cow status monitoring. J. Dairy Sci. 84 (2) pID: 51; Note: Includes references; Identifier: false positive results. Diagnostic value.; Category Code: Veterinary Science [L800]. Mathematics And Statistics [X100].; NAL Location: 44.8 J822.; Update Code: 200505.
Dominiak, 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.
Dominiak, K.N., Pedersen, L.J., Kristensen, A.R., 2018;al., submitted for publication. Spatial modeling of pigs' drinking patterns as an alarm reducing method. I. Developing a multivariate dynamic linear model. J. Publ (submitted for publication).
Fortmann-Roe, S., 2012. Accurately measuring model prediction error. http://scott.fortmann-roe.com/docs/MeasuringError.html visited on the 27th of October 2018.
Hogeveen, H., Kamphuis, C., Steeneveld, W., Mollenhorst, H., 2010. Sensors and clinical mastitis - the quest for the perfect alert. Sensors 10 (9), 7991–8009 update Code: 20100000.
Jensen, D.B., Hogeveen, H., Vries, A.D., 2016. Bayesian integration of sensor information and a multivariate dynamic linear model for prediction of dairy cow mastitis. J. Dairy Sci. 99 (9), 7344–7361.
Jensen, D.B., Kristensen, A.R., 2016. Comparison of strategies for combining dynamic linear models with artificial neural networks for detecting diarrhea in slaughter pigs. J. Anim. Sci. 94 (E-Suppl. 5), 84.
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, 51–62.
Kamphuis, C., Mollenhorst, H., Heesterbeek, J.A.P., Hogeveen, H., 2010. Detection of clinical mastitis with sensor data from automatic milking systems is improved by using decision-tree induction. J. Dairy Sci. 93 (8), 3616–3627 iD: 65; Update Code: 20100000.
Lawson, A.B., Banerjee, S., Haining, R.P., Ugarte, M.D., 2016. Handbook of Spatial Epidemiology. Taylor & Francis.
Lyderik, K., Andersen, H.-L., Kristensen, H., Pedersen, L., 2016. Protocol for daily observation in the herds of the pigit project. pigit report no 8. Tech. rep.
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 update Code: 20050000.
Madsen, T.N., Kristensen, A.R., 2005. A model for monitoring the condition of young pigs by their drinking behaviour. Comput. Electron. Agric. 48 (2), 138–154 update Code: 20050000.
Mein, G.A., Rasmussen, M.D., 2008. Performance evaluation of systems for automated monitoring of udder health: would the real gold standard please stand up? In: Lam, T.J.G.M. (Ed.), Mastitis control: from science to practice. Proceedings of International Conference. Mastitis control: from science to practice. Proceedings of International Conference, The Hague, Netherlands. Wageningen Academic Publishers, Wageningen iD: 34; September - 2 October 2008; 2008.:259-266; Update Code: 20090000.
Montgomery, D.C., 2013. Statistical Quality Control A Modern Introduction, seventh ed. John Wiley & Sons Inc.
Ostersen, T., Cornou, C., Kristensen, A.R., 2010. Detecting oestrus by monitoring sows' visits to a boar. Comput. Electron. Agric. 74 (1), 51–58.
Pedersen, K.S., 2012. Smaagrisediarre. Dansk Veterinaer Tidsskrift.
R Core Team, 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. < http://www.R-project.org/ > .
Steeneveld, W., Gaag, L.C.v.d., Ouweltjes, W., Mollenhorst, H., Hogeveen, H., 2010. Discriminating between true-positive and false-positive clinical mastitis alerts from automatic milking systems. J. Dairy Sci. 93 (6), 2559–2568 update Code: 20100000.
Witten, I., Frank, E., 2005. Data Mining: Practical Machine Learning Tools and Techniques. second ed. Morgan Kaufmann, Elsevier.