STATENS VETERINÄRMEDICINSKA ANSTALT

Date: 12-09-2022

Final report

The final report should be completed and filed no later than 6 months after termination of the project

TIDIGA LARM OM PRODUKTIONSSTÖRNINGAR FÖR SNABBARE ÅTGÄRD OCH MINSKADE FÖRLUSTER I GRISFÖRETAG

Project number: O-17-20-978

Budget: 2 909 000

Project period: 2018-2020 (extended to 2021)

Main applicant:

Fernanda Dórea, Swedish National Veterinary Institute
fernanda.dorea@sva.se

Co-applicant(s):

Jenny Frössling, Swedish National Veterinary Institute
Cecilia Hultén, Swedish National Veterinary Institute
Marie Sjölund, Swedish National Veterinary Institute
Lena Eliasson-Selling, Gård och Djurhälsan

Part 1.1: Summary/Abstract

This project aimed to extract more value from data collected in pig farms, constructing indicators of production and reproduction performance which could be monitored continuously. By developing methods for early detection of trends in these indicators, this project aimed to avoid losses due to silent and undetected disease and production disturbances. Health-related indicators were created using data from the production management software WinPig (PigVision), used by the majority of swine farms in Sweden. From a methodological perspective, the research advanced methods to deal with new types of statistical indicators. From a practical perspective, a prototype system was developed with the help of pig health advisors. System performance was evaluated using simulated outbreaks of porcine respiratory and reproduction syndrome (PRRS). The system proved capable of providing early detection of unexpected trends.

Projektet syftade till att skapa ytterligare nytta av data som vanligen registreras på grisgårdar genom att konstruera indikatorer för produktion och fruktsamhet som kan övervakas kontinuerligt. Målet var att bidra till att ge producenter en förbättrad bevakning av gårdens djurhälsa genom att utveckla metoder för tidig upptäckt av trender i dessa indikatorer. Hälsorelaterade indikatorer skapades med hjälp av data från WinPig (PigVision), som används av många grisgårdar i Sverige. Ur ett metodologiskt perspektiv har forskningen tillfört

The project has been financed by:

Stiftelsen Lantbruksforskning

avancerade metoder för att hantera nya typer av statistiska indikatorer. Rent praktiskt utvecklades ett prototypsystem med hjälp av grishälsorådgivare. Systemets prestanda utvärderades med hjälp av simulerade utbrott av grissjukdomen Porcine Reproductive and Respiratory Syndrome (PRRS). Systemet visade sig kunna ge tidig upptäckt av oväntade trender och fungerade som ett användbart verktyg för övervakning av hälso- och produktionsförluster.

Part 1.2: Main report

Introduction

As a response to the need for timely disease detection, and enabled by growing data digitalization, syndromic surveillance (SyS) has been used more and more frequently in animal health (Dórea et al., 2011). The foundational assumption of any SyS system is that a given indicator of population health changes when a health hazard is introduced, and these changes can be noticed if routinely collected data are analysed continuously (Mandl et al., 2004).

As electronic records are reaching farms, data recorded at the herd level, or even individually for each animal regarding different phases of productivity (hereby generalised as “production data”), have appeared as a growing data source for animal health surveillance. These data, if correctly managed, allow large population coverage and the shortest time lag between a health event and its potential detection (Dórea and Vial, 2016). Production records are generated continuously, not only on the occasion of a disease event. This offers advantages such as timeliness, but also results in a lack of specificity, in particular representing a challenge to the definition of the events to monitor and what deviations should be characterised as abnormal. These methodological issues have not yet been extensively addressed, as exploration of farm-level data for SyS has so far been limited, likely due to their distributed nature and governance.

This project aimed to investigate whether production data recorded regularly in reproduction pig farms could be used to construct health and performance indicators; and whether these indicators could be subjected to near-real time temporal monitoring (syndromic surveillance) to detect early signals of reproduction failure, which could be caused by undetected diseases or management errors.

Materials and methods

Data sources:

The majority of swine farms in Sweden manage their records using the production management software WinPig (also called PgVision internationally, Agrosoft®, http://agrosoft.eu/, 2014). To explore the structure of data recorded using this software, data from three large sow producing farm, shared voluntarily and anonymously were used. Dtaa form one of the farms (the largest one) will be used to present results in details and discuss.

For back-end data management, PigVision creates a local server which hosts a SQL (Structured Query Language) database. Data from this SQL database were accessed within the statistical programming environment R (R Core Team, 2020) using the package {RODBC} (Ripley and Lapsley, 2020).

Data organization into indicators and participation of stakeholders

The data available in PigVision were used to construct production and reproduction performance indicators. Pig health advisors from Gård&Djurhälsan participated in the process through two different workshops carried out in May 2018, and October 2019. In the first workshop, they helped interpret the data and construct the indicators. In the second workshop, they helped select indicators that could be useful to monitor health, as well as production and

The project has been financed by:

Stiftelsen Lantbruksforskning

reproduction management. Stakeholders also advised that the best unit of time for counting events and running detection would be weekly. The indicators chosen are listed in Table 1.

Some indicators could be easily grouped as "number of events per time-point (weekly)", which is the type of time-series traditionally used in SyS (discrete time-series). Some other indicators were better expressed as continuous time-series, that is, series in which every new event represents a new observation. A specific value associated with these observations is monitored individually, rather than grouped in discrete time-steps. Examples are pregnancy length and number of live born piglets, both associated with every new farrowing event. The column "Limit(s)" in Table 1 identifies whether indicators were relevant for detection of potential increases (upper control limits are used), decreases (lower control limits) or both.

Time-series analysis

Most of the SyS systems developed to date are based on discrete time-series (Unkel et al., 2012; Dórea and Vial, 2016). Time-series analyses of the continuous time-series was investigated to the extent possible, as described in each step. The entire workflow was programmed using the statistical environment R, and is available publicly (https://github.com/SVA-SE/PigPeaks).

Table 1. List of performance indicators constructed from production data, and chosen by stakeholders as relevant for monitoring.

Category	**Indicators**	**Time-series**	**SyS**[a]	**Limit(s)**[b]	**Utility**
Gilts	Age at first service	Continuous	No	---	Management
	Age at first farrowing	Continuous	No	---	Management
Empty Sows	Number of sows empty longer than 4 days	Weekly	Yes	UCL [c]	Health
Services	Reservices per week	Weekly	Yes	UCL	Health
	Time to reservice	Continuous	Yes	UCL | LCL [d]	Health
	Percentage of reservices after 4 weeks	Weekly	Yes	UCL	Health
	Percentage of failures after 4 weeks	Weekly	Yes	UCL	Health
Pregnancy	Pregnancy length	Continuous	Yes	LCL	Health
	Time to abortion	Continuous	No	---	Health
	Abortions per week	Weekly	Yes	UCL	Health
Farrowing	Days between farrowings	Continuous	Yes	UCL | LCL	Health
	Total piglets per farrowing	Continuous	Yes	UCL | LCL	Health
	Live piglets per farrowing	Continuous	Yes	LCL	Health
	Percentage of dead piglets per farrowing	Continuous	Yes	UCL	Health
	Mummified piglets per farrowing	Continuous	Yes	UCL	Health
Weaning	Days to weaning	Continuous	No	---	Management
	Litters weaned per week	Weekly	No	---	Management
	Piglets weaned per week	Weekly	Yes	LCL	Health
	Piglets weaned per litter	Continuous	Yes	LCL	Health
	Deaths in weaning piglets per week	Weekly	Yes	UCL	Health
	Weaning weight per litter (average per piglet)	Continuous	Yes	LCL	Health
Post-Weaning	Deaths in weaners per week	Weekly	Yes	UCL	Health
Exit	Last event before exiting per week	Weekly	No	---	Management
	Dead sows per week	Weekly	Yes	UCL	Health

[a] Syndromic Surveillance applied; [b] Temporal aberration detection limit(s); [c] Upper control limit; [d] Lower control limit.

The project has been financed by:
Stiftelsen Lantbruksforskning

The presence of temporal effects (Lotze et al., 2008) was evaluated for all weekly time-series using the R package {vetsyn} (Dórea et al., 2015). To create an outbreak-free historical baseline for each indicator in which SyS was applied, the non-parametric method of moving percentiles was chosen, using a 95th moving percentile.

In the absence of significant temporal effects that needed to be modelled and extracted from the discrete time-series data, control charts were applied directly to the indicator time-series as the methods of aberration detection. This even for the continuous time-series, which although novel in veterinary SyS, more closely resemble the industrial processes which control charts were originally intended to monitor.

To ensure detection of sudden spikes (potentially indicative of management errors) as well as slow increases/decreases in the observed values (potential outbreak signals), both the Shewhart and Exponentially Weighted Moving Average (EWMA) control charts were applied (Yahav et al., 2011). We previously demonstrated that these control charts have complementary performance in detecting outbreaks of different shapes (Dórea et al., 2014).

Algorithms' parameterization and implementation

Both detection algorithms chosen were applied to each indicator. For the weekly indicators, the current algorithms implementations from the {vetsyn} package were used, which is informed by our previous work (Dórea et al.,2013a).

The {vetsyn} workflow was designed specifically for discrete time-series analysed in the SyS context, as algorithms are applied recursively in each time-point to be evaluated.

This project allowed us to advance methods for analysis of continuous time-series. The original control charts implementations in the package {qcc} (Scrucca, 2004) were used, in which there is no separation between baseline and evaluation window. The entire historical series is used to calculate the central, upper and lower control limits. All observations in the time-series are then evaluated against these limits.

For the choice of detection limits, we used an approach that our group had developed (Dórea et al., 2013b) of using a scoring system to combine multiple detection limits to generate a "severity score" between 0 (no alarm) and 3 (highest alarm)

Evaluation of performance in a practical scenario

No outbreaks of regulated diseases were known to be present in the historical data, so we opted for outbreak simulation. It was of specific interest to inject outbreaks of PRRS since Sweden is free of the disease, but the virus still circulates in Europe (Carlsson et al., 2009). Based on Pejsak and Markowska-Daniel (1997) and Valdes-Donoso et al. (2018), ten indicators were deemed most suitable to inject outbreak signals:

- *Reservices per week*, *Abortions per week*, *Piglets weaned per week*, *Deaths in weaning piglets per week*, and *Dead sows per week*; all of which are discrete, *weekly* time-series.
- *Pregnancy length*, *Live piglets per farrowing*, *Percentage of dead piglets per farrowing*, *Mummified piglets per farrowing*, and *Piglets weaned per litter*; all of which are *continuous* time-series.

For each of the 10 indicators, twelve copies of each indicator series were created, and a single outbreak signal was injected in each of those copies, always in a different quarter (12 outbreaks simulated). Detection performance was evaluated based on three main performance measures: false alarm rate, timeliness and sensitivity. False alarm rate was assessed based on the number of alarms triggered when the control charts were applied to the original observed data, without any injected outbreaks. The false alarm rate was calculated based on the lowest detection threshold. Timeliness of detection was assessed by plotting the simulated outbreak signal curves against the density of alarms generated per week, for each indicator and each

The project has been financed by:

Stiftelsen Lantbruksforskning

algorithm. The system's sensitivity was evaluated based on the total number of alarms generated by each of the two control charts for each of the 12 simulated outbreaks.

Communication of results to stakeholders

Two dashboards for visualization of the indicators and trends were created: a simple and a full dashboard. The simple dashboard, which does not include all indicators, was meant to be used for communication purposes when discussing with health advisors and farmers what the indicators mean, and how they can be monitored. A full dataset was created artificially to create a simulated farm, and allow this dashboard to be made publicly without exposing the data from the contributing farmers who shared their data anonymously to support the project.

Results

Performance of algorithms to detect outbreaks

Figs. 1 and 2 show the total number of alarms triggered in the weekly time-series by the EWMA and Shewhart control charts, respectively, in each simulated outbreak week, for all twelve simulated quarters simultaneously (superimposed). For instance, week 1 shows the sum of alarms triggered in week 1 summing all twelve simulated outbreak quarters. EWMA results are shown for continuous time-series in Fig. 3.

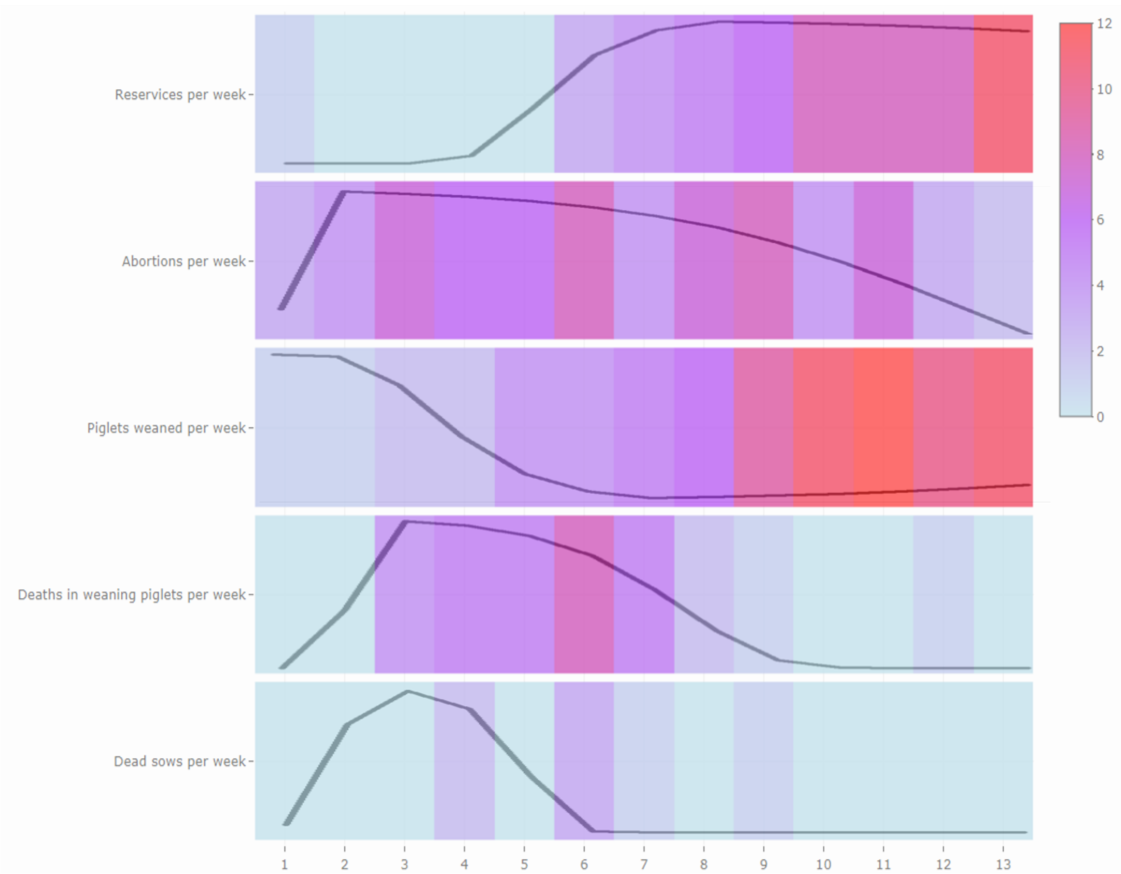

Figure 1. Total number of alarms triggered by the EWMA control chart in each week of simulated outbreaks in weekly indicators, superimposed for all twelve quarters simulated. The corresponding log-normal curves used for aberration injection are also depicted.

The project has been financed by:

Stiftelsen Lantbruksforskning

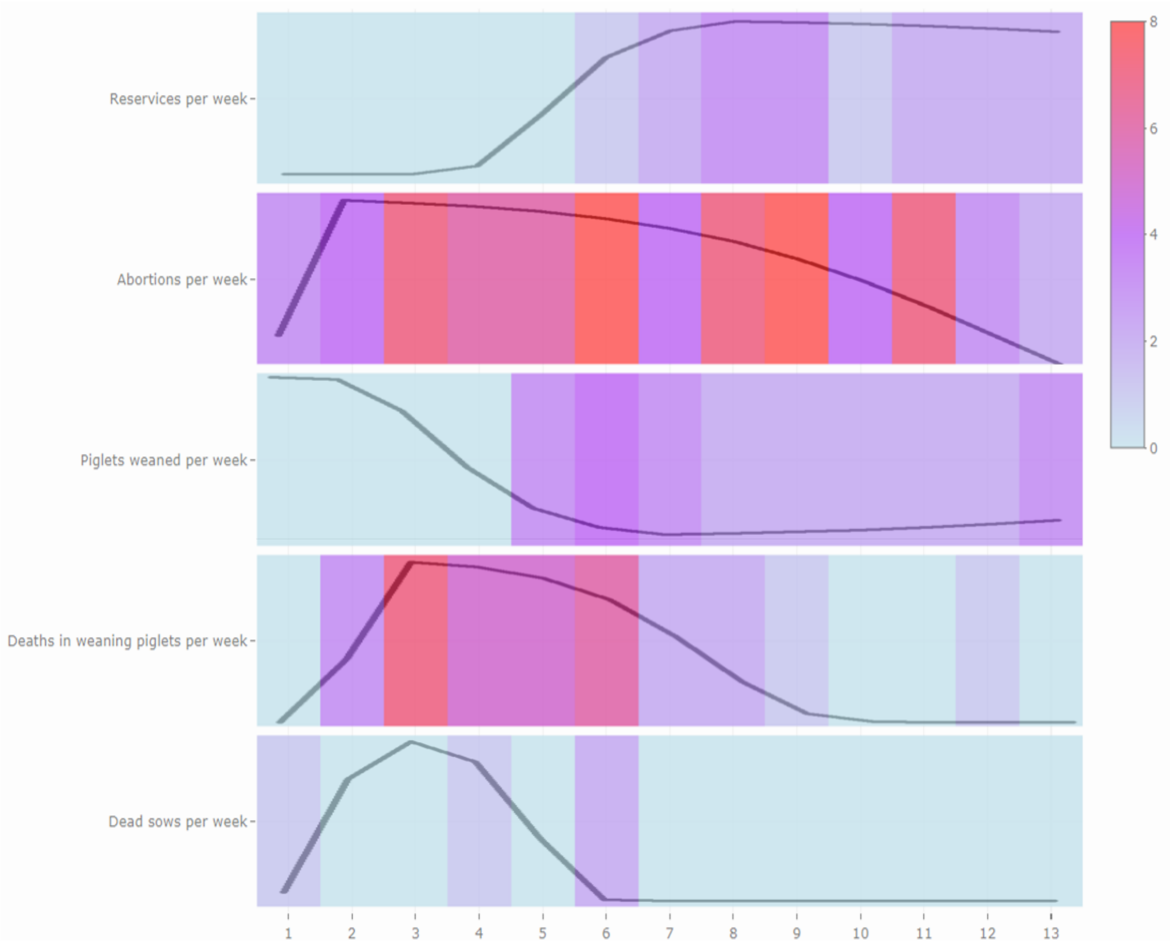

Figure 2. Total number of alarms triggered by the Shewhart control chart in each week of simulated outbreaks in weekly indicators, superimposed for all twelve quarters simulated. The corresponding log-normal curves used for aberration injection are also depicted.

Regarding weekly indicators, and as illustrated in Figs. 1 and 2, in 3 out of 12 simulated outbreak quarters the system generated an alarm already in week 1 for the indicator *Abortions per week* with both control charts. Indicators varied in their sensitivity and how early they generated alarms, but by outbreak week 6 all 5 weekly indicators monitored would already have generated at least one alarm in all simulated scenario

For the continuous indicators, as shown in Fig. 3 the indicator which triggered the most EWMA alarms in an earlier phase was *Mummified piglets per farrowing*. Timeliness of continuous indicators with the Shewhart algorithm was poorer in comparison to EWMA.

The project has been financed by:

Stiftelsen Lantbruksforskning

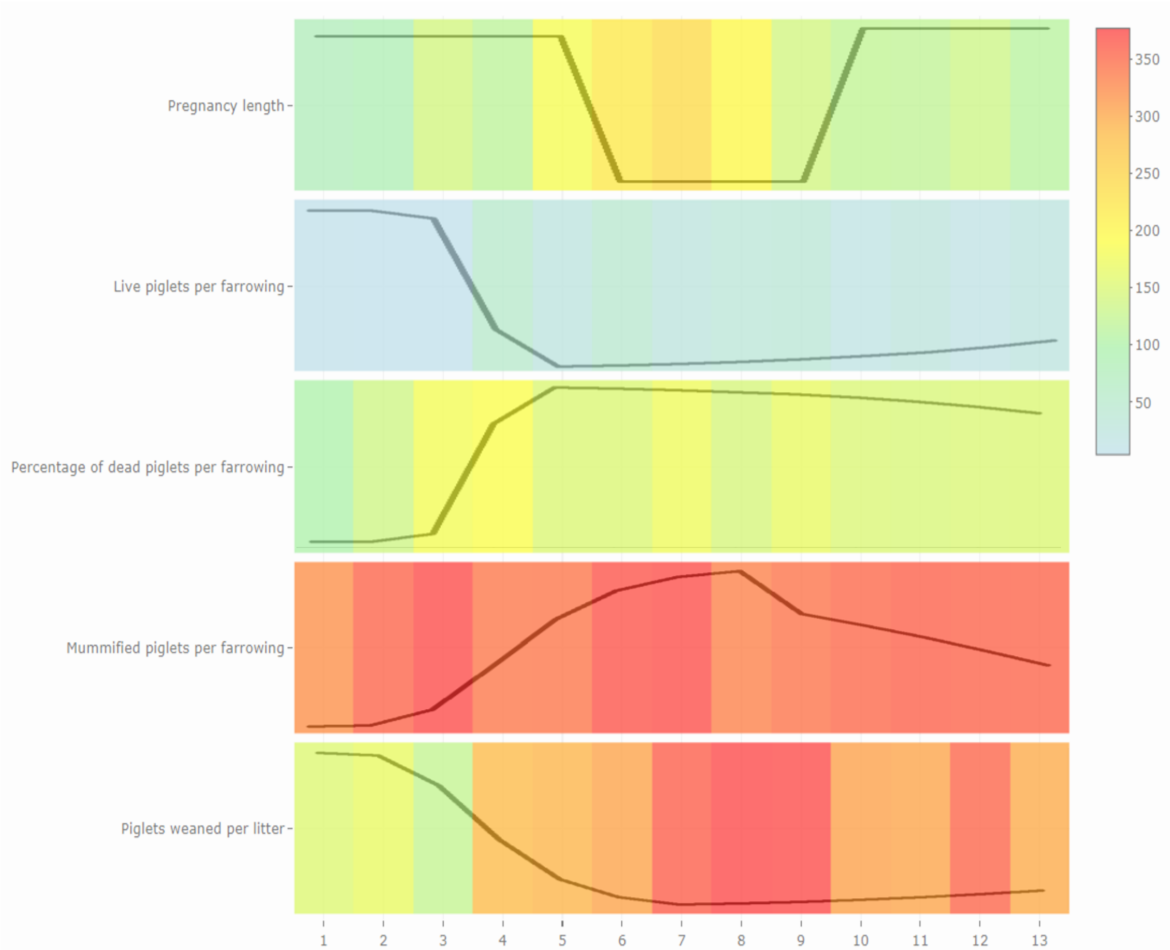

Figure 3. Total number of alarms triggered by the EWMA control chart in each week of simulated outbreaks in continuous indicators, superimposed for all twelve quarters simulated. The corresponding log-normal curves used for aberration injection are also depicted.

Dashboard for communication

The simple dashboard is available at (https://sva-se.github.io/PigPeaks). The screenshot in Figure 4 shows the first panel on the dashboard, which shows 4 mortality indicators. The version online has a second panel with service success indicators (reservices per week, percentage of services that result in reservice, and service failure percentage). The dashboard is fully interactive, allowing users to zoom, see more information for each week upon hovering the mouse, and select which sews categories to show from 4 age categories: gilts, young, prime and mature.

The project has been financed by:

Stiftelsen Lantbruksforskning

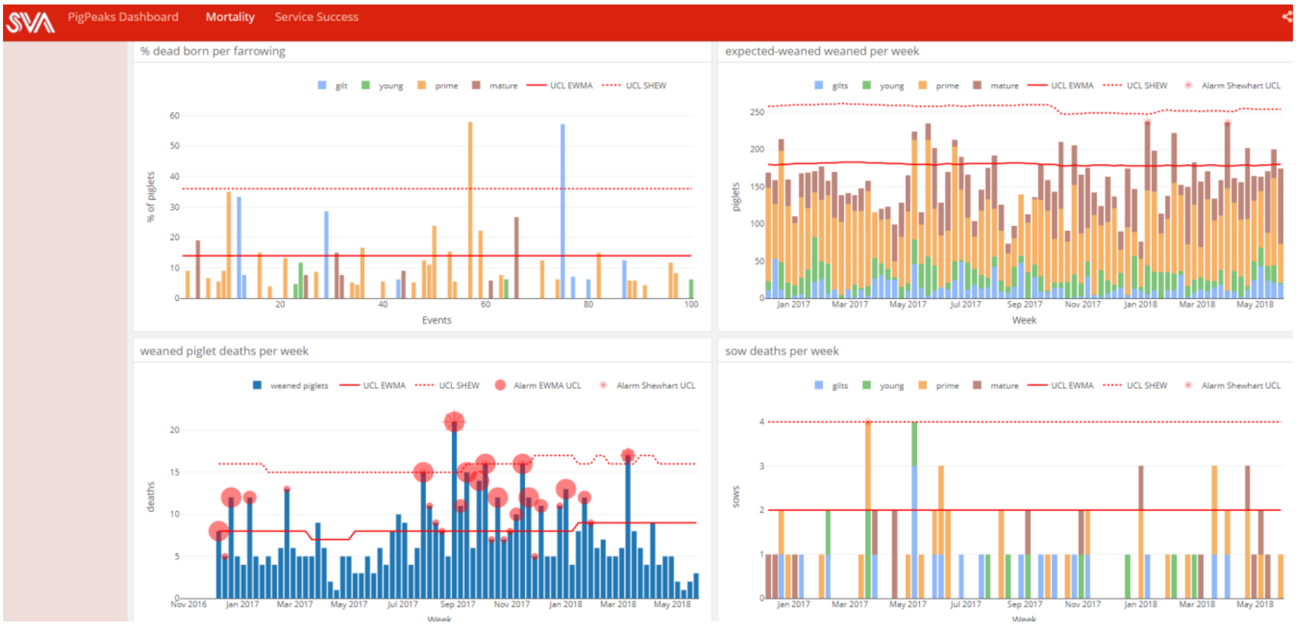

Figure 4. Dashboard showing 4 indicators of mortality, available at https://sva-se.github.io/PigPeaks.

A version with all indicators listed in Table 1, separated in panels for each production category shown in Table 1, was also constructed. All codes are publicly available for reuse here: https://github.com/SVA-SE/PigPeaks.

Discussion

The project was successful in creating an automated workflow of data analysis that allows data to be extracted from the software WinPig, and compile these data as continuous indicators for monitoring of productive and reproductive performance. This workflow will enable the development of data-driven tools to support herd management, as well as disease control and surveillance in the future.

In this specific project, we focused on creating data workflows to apply continuous monitoring of production and reproduction performance, and showed that the application of syndromic surveillance algorithms to the data was capable of providing early detection of signals associated with PRRS infection.

The system investigated solves the practical issue of data access by applying SyS directly at the data source, using data already routinely collected by the leading swine herd management software in Sweden. Implementation in practice and in real-time will however require evaluation of a greater number of farms, and in particular development of the monitoring system to work prospectively, while all analyses reported here were performed retrospectively.

As the number of indicators brought up was large, the main challenge for adoption of the system in practice will be interpreting all the system outputs, and incorporating them into every day management decisions. In the future, research should focus on how to combine the evidence of multiple indicators to have one overall alarm/assessment of the risk of an ongoing outbreak.

Conclusions

The value of data already collected on pig farms as part of every-day operations can be greatly increased if these data are used to routinely monitor production and reproductive performance. This project showed, from a methodological perspective, that this is possible. Practical use will require further collaboration with stakeholders to transform the system into a true “decision support system”.

The project has been financed by:
Stiftelsen Lantbruksforskning

Relevance and recommendations

From a practical perspective, the system advanced options for the implementation of syndromic surveillance in practice, and created a workflow that can be reused in the future to provide Swedish stakeholders with data-driven methods to support decision in animal health and animal welfare.

From a research and methodological perspective, the project advanced syndromic surveillance methods by uncovering “continuous time-series indicators”, and proposing methods to incorporate them into automated monitoring.

Interaction with farmers and health advisors, the potential users of the system, was poorer than planned due to the COVID-19 outbreak. At the same time, we knew this project would only achieve a pilot using data from a small number of farms, and a next step of large development and implementation of the developed prototype would be needed. During the third year of the project, the research team started a collaboration with social scientists in the University of Nottingham and University of Zurich, which is focused on the very issue of creating and testing the use of data tools to improve disease control and management at the farm level. That collaboration has now been funded through a large Horizon 2020 project, DECIDE (https://decideproject.eu/), including partners from around Europe, also testing data tools for control of non-regulated diseases. The results of PigPeaks are directly feeding into DECIDE, which will allow further testing and implementation of the dashboard as a tool for decision in practice.

References

Carlsson, U., Wallgren, P., Renström, L.H.M., Lindberg, A., Eriksson, H., Thorén, P., Eliasson-Selling, L., Lundeheim, N., Nörregard, E., Thörn, C., Elvander, M., 2009. Emergence of porcine reproductive and respiratory syndrome in Sweden: Detection, response and eradication. Transbound. Emerg. Dis. 56, 121–131. https://doi.org/10.1111/j.1865-1682.2008.01065.x.

Dórea, F.C., Vial, F., 2016. Animal health syndromic surveillance: a systematic literature review of the progress in the last 5 years (2011-2016). Vet. Med. Res. Reports. 7, 157–170. https://doi.org/10.2147/vmrr.s90182.

Dórea, F.C., Sanchez, J., Revie, C.W., 2011. Veterinary syndromic surveillance: Current initiatives and potential for development. Prev. Vet. Med. 101, 1–17. https://doi.org/10.1016/j.prevetmed.2011.05.004.

Dórea, F.C., McEwen, B.J., McNab, W.B., Revie, C.W., Sanchez, J., 2013a. Syndromic surveillance using veterinary laboratory data: Data pre-processing and algorithm performance evaluation. J. R. Soc. Interface. 10. https://doi.org/10.1098/rsif.2013.0114.

Dórea, F.C., Lindberg, A., McEwen, B.J., Revie, C.W., Sanchez, J., 2014. Syndromic surveillance using laboratory test requests: A practical guide informed by experience with two systems. Prev. Vet. Med. 116, 313–324. https://doi.org/10.1016/j.prevetmed.2014.04.001.

Dórea, F.C., Widgren, S., Lindberg, A., 2015. Vetsyn: An R package for veterinary syndromic surveillance. Prev. Vet. Med. 122, 21–32. https://doi.org/10.1016/j.prevetmed.2015.10.002.

Keeling, M., Rohani, P., 2008. Modeling Infectious Diseases in Humans and Animals, in: Introduction to Simple Epidemic Models, Princeton University Press, New Jersey, United States, pp. 15-52. https://doi.org/10.2307/j.ctvcm4gk0.

Lotze, T., Murphy, S., Shmueli, G., 2008. Implementation and Comparison of Preprocessing Methods for Biosurveillance Data. Adv. Dis. Surveill. 6, 1–20.

Mandl, K., Overhage, J.M., Wagner, M.M., Lober, W.B., Sebastiani, P., Mostashari, F., Pavlin, J.A., Gesteland, P.H., Treadwell, T., Koski, E., Hutwagner, L., Buckeridge, D.L., Aller, R.D., Grannis, S., 2004a. Implementing Syndromic Surveillance: A Practical Guide Informed by the Early Experience. J. Am. Med. Informatics Assoc. 11, 141–150. https://doi.org/10.1197/jamia.M1356.

Pejsak, Z., Markowska-Daniel, I., 1997. Losses due to porcine reproductive and respiratory syndrome in a large swine farm. Comp. Immunol. Microbiol. Infect. Dis. 20, 345–352. https://doi.org/10.1016/S0147-9571(97)00010-6.

Scrucca, L., 2004. qcc: An {R} package for quality control charting and statistical process control. R News 4, 11–17.

Silva, G.S., Schwartz, M., Morrison, R.B., Linhares, D.C.L., 2017. Monitoring breeding herd production data to detect PRRSV outbreaks. Prev. Vet. Med. 148, 89–93. https://doi.org/10.1016/j.prevetmed.2017.10.012.

Unkel, S., Farrington, C.P., Garthwaite, P.H., Robertson, C., Andrews, N., 2012. Statistical methods for the prospective detection of infectious disease outbreaks: a review. J. R. Stat. Soc. 175, 49–82. https://doi.org/10.1111/j.1467-985X.2011.00714.x.

Valdes-Donoso, P., Alvarez, J., Jarvis, L.S., Morrison, R.B., Perez, A.M., 2018. Production losses from an endemic animal disease: Porcine reproductive and respiratory syndrome (PRRS) in selected Midwest US Sow Farms. Front. Vet. Sci. 5, 1–10. https://doi.org/10.3389/fvets.2018.00102.

Yahav I., Lotze T., Shmueli G., 2011. Algorithm Combination for Improved Performance in Biosurveillance, in: Castillo-Chavez C., Chen H., Lober W., Thurmond M., Zeng D. (Eds.), Infectious Disease Informatics and Biosurveillance, Integrated Series in Information Systems, vol 27, Springer, Boston, MA, pp. 173-189. https://doi.org/10.1007/978-1-4419-6892-0_8.

Result dissemination:

Scientific publications, *published*	*C. Merca, I. Clemensson Lindell, L. Ernholm, L. Eliasson Selling, T.P. Nunes, M. Sjölund, F.C. Dórea. Veterinary syndromic surveillance using swine production data for farm health management and early disease detection, Preventive Veterinary Medicine, Volume 205,2022,ISSN 0167-5877. https://doi.org/10.1016/j.prevetmed.2022.105659. (https://www.sciencedirect.com/science/article/pii/S0167587722000927*
Conference publications/ presentations	*C. Merca, I. Clemensson Lindell, L. Ernholm, L. Eliasson Selling, T.P. Nunes, M. Sjölund, F.C. Dórea 2022, Title: Application of syndromic surveillance to routinely collected swine production data for farm health management and early disease detection Society for Veterinary Epidemiology and Preventive Medicine (SVEPM) Annual conference, 23-25 March 2022, Belfast, Ireland https://svepm2022.org/programme/*
Oral communication*, to sector, students etc.*	*Workshop to pig health advisors from Gård och Djurhälsan, May 29, 2018, Stockholm. Diskussionsmöte om nya lösningar för översikt av djurhälsa/produktion och larm om störningar. 12 participants.*
	Workshop to pig health advisors from Gård och Djurhälsan, October 17, 2019, Stockholm. 9 participants.

The project has been financed by:

Stiftelsen Lantbruksforskning