Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Preventive Veterinary Medicine 154 (2018) 23–28

Contents lists available at ScienceDirect

Preventive Veterinary Medicine

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

Exploring milk shipment data for their potential for disease monitoring and for assessing resilience in dairy farms

Nils Fall[a,*], Anna Ohlson[b], Ulf Emanuelson[a], Ian Dohoo[c]

[a] Department of Clinical Sciences, Swedish University of Agricultural Sciences, PO Box 7054, SE-75007 Uppsala, Sweden
[b] Växa Sverige, PO Box 210, SE-101 24 Stockholm, Sweden
[c] Centre for Veterinary Epidemiological Research, University of Prince Edward Island, Charlottetown, PEI, Canada

ARTICLE INFO

Keywords:
Milk shipment data
Disease monitoring
Resilience, organic
Dairy

ABSTRACT

The use of routinely recorded data for research purposes and disease surveillance is an attractive proposition. However, this requires that the validity and reliability of the data be evaluated for the purpose for which they are to be used. This manuscript reports an evaluation of milk shipment data for evaluating their usefulness in disease monitoring and the resilience of organic and conventional dairy herds in Sweden. A large number of inconsistencies were observed in the data, necessitating substantial efforts to "clean" the data. Given that the selection of rules used in the cleaning process was subjective in nature, a sensitivity analysis was carried out to determine if different cleaning routines produced substantially different results. Despite the cleaning efforts we observed far more large residuals at the shipment level than expected. Thus, it was concluded that the data were too "noisy" to be used for identification of short term impacts on milk production.

Resilience was evaluated by examining the residual variance in milk shipped per cow per day under the assumption that herds with high resilience would have lower residual variance. The effects on residual variance of organic status or whether or not the herd used an automatic milking system were evaluated in models in which the residual variance was stratified or not by these factors. We did not find consistent evidence to suggest that organic herds had higher resilience than conventional herds, but this could be partly due to using residual variance as the measure indicating resilience.

1. Introduction

Milk recording data are a common source of information when it comes to measuring milk production, milk composition and udder health in dairy production. Milk recording is based on individual cow measurements taken, usually, monthly. An alternative, potential source of information on milk production and milk composition is the milk shipment data, which has the benefit of being recorded frequently, i.e. at each bulk tank milk shipment, usually every one or two days. Milk shipment data are, however, aggregated on a herd level and have therefore less detail. Because most serious diseases afflicting dairy cattle affect milk production (Wüthrich et al., 2016; Cumming et al., 2005; Charfeddine and Pérez-Cabal, 2017; Toftaker et al., 2017), these regularly collected milk production data have potential for use in research and for disease surveillance through monitoring of fluctuations in production. Milk shipment data have previously been used for research purposes (Toftaker et al., 2017), but a thorough evaluation of the usefulness of the data source still is needed.

The concept of resilience in ecological systems was first described by the Canadian ecologist Holling (1973) and describes the capacity of an ecosystem to tolerate disturbance without collapsing into a qualitatively different state that is controlled by a different set of processes. A resilient ecosystem can withstand shocks and rebuild itself when necessary. In modern literature, resilience is defined as the capacity of a system to absorb disturbance and re-organize while undergoing change so as to still retain essentially the same function, structure, identity and feedbacks (Folke, 2006). Elgersma et al., 2018;Döring et al. (2014) suggest that there are different definitions of resilience but distinguish three steps that all covered in almost all definitions: disturbance, response and outcome. In dairy production resilience can be defined as the capacity to cope with disease and other production disturbances, a concept akin to general adaption (Elgersma et al., 2018;Döring et al., 2014; Elgersma et al., 2018). Health, as stated by the IFOAM principles of organic agriculture, is the wholeness and integrity of living systems. It is not simply the absence of illness, but the maintenance of physical, mental, social and ecological well-being. Immunity, resilience and

Abbreviations: AMS, automated milking system

* Corresponding author.
E-mail addresses: nils.fall@slu.se (N. Fall), anna.ohlson@vxa.se (A. Ohlson), ulf.emanuelson@slu.se (U. Emanuelson), dohoo@upei.ca (I. Dohoo).

https://doi.org/10.1016/j.prevetmed.2018.03.012
Received 28 November 2017; Received in revised form 14 March 2018; Accepted 17 March 2018
0167-5877/ © 2018 Elsevier B.V. All rights reserved.

regeneration are key characteristics of health (IFOAM, 2018). Organic philosophy implies that natural behaviour, optimal feed and low stress levels are disease preventing factors and that they also lead to better resilience. Thus, it can be hypothesised that organic dairy herds would have less severe, or shorter duration of infections or other diseases/disturbances.

However, the concept of resilience in organic dairy production is challenging to assess and at least one previous attempt has been carried out. Elgersma et al. (2018) used variability of milk production in individual cows as an indicator of resilience. If milk shipment data can be shown to be a reliable means for disease surveillance, we hypothesize that they can be used to assess resilience by monitoring the variability in the routine bulk tank milk shipment volumes.

The first objective of this study was to explore the challenges and possibilities of using milk shipment data as a source of information for disease outbreak surveillance. A secondary objective was to use shipment data to assess the potential difference in resilience in organic versus non-organic dairy production.

2. Materials and methods

2.1. Study population

This study was based on a data from a research project with focus on monitoring outbreaks with Bovine Respiratory Syncytial Virus and Bovine Corona Virus in organic and conventional dairy herds in Sweden (Wolff et al., 2015). The sampling frame was all dairy herds with an average herd size of at least 50 cows and enrolled in the Swedish Official Milk Recording Scheme. Geographically, all Swedish counties except for the most southern (Skåne) were included. Skåne was excluded because it was known that there are very few Bovine Respiratory Syncytial Virus or Bovine Corona Virus negative herds in that region. Organic herds were defined as herds with KRAV-certified dairy production (www.krav.se). KRAV is the major Swedish certification body for organic production. In Sweden all farms have grazing systems according to legal requirements. The main Swedish breeds are Swedish Red and Swedish Holstein, where the yearly average milk yield in 2017 was 9156 kg and 10,274 kg, respectively (Veldhuis et al., 2016;Växa Sverige, 2018). All dairy companies have individual rules on bulk milk somatic cell counts for premium payments and for rejection of deliverance of milk. A simple random sample of 400 conventional and all eligible organic herds (n = 244) were sent a written invitation to the study in May 2011. The number of invited herds was based on previous experiences of about 30% willingness among Swedish dairy farmers to participate in similar observational studies. In total, 69 (17%) and 75 (31%) farmers with conventional and organic herds, respectively, agreed to participate in the project. However, only 93 herds (54 organic and 39 conventional) had all of the required milk shipment and milk size data and were included in the analyses. In 27 out of 93 herds the milking system changed from conventional milking systems to AMS during the study period.

2.2. Data sources

Both milk shipment and herd size data were available for the period of 1 Apr 2012 to 26 Nov 2015 and these were the data used in the analyses. The shipment data were obtained from the Swedish central milk shipping register, but were collected in two different ways. Firstly, herd-level data were collected from farmers, through their system interface, and secondly, from the central register itself. The central register only saves shipment data for the past three years. Data collected from farmers through their system interfaces stretch further back in time compared to the data retrieved directly from the central register, as interface collection was executed at an earlier point in time. Despite two different collection methods, all shipment data derived from the same source and all variables are the same.

Test day records (approximately monthly) of herd size for the period of 1 July 2010 to 26 November 2015 had been obtained from the Swedish Official Milk Recording Scheme. In order to estimate the number of cows milked on each milk shipment date, linear interpolation of the number cows milking on the previous and subsequent test dates was used. Total milk shipped was converted to milk per cow per day based on the length of the interval between shipments and the estimated number of cows milking.

A herd management data file was obtained from the Swedish Official Milk Recording Scheme. This file included production system, meaning whether the herd was organic or conventional (non-organic), and whether or not it had AMS. If the herd switched AMS category during the study period, the date of transition was noted and the herd's AMS status determined for each shipping date. Region of the country (1 = south, 2 = central, 3 = north), breed of the herd (1 = > 80% Swedish Red, 2 = > 80% Swedish Holstein, 3 = mix Red and Holstein, 4 = other breeds/mix or not recorded) were recorded.

2.3. Data compilation, cleaning and new variable calculations

A number of steps were taken during compilation and amalgamation of the data. These are described in Table 1.

It was clear that there were quite a few large changes in the amount of milk shipped that could not possibly be explained by changes in actual milk production in the herd. In order to remove these “artifacts” a set of rules were created to identify records that had changes in milk shipment volume that were unlikely to have come from a real change in production. These rules were:

Suspect rule 1 - Identify observations where shipping interval changed length (e.g. 2 days to 1 day).

Suspect rule 2 - Identify observations where the total Kg of milk shipped per day changed by more than 20% in small herds (< 50 cow herds), 15% in medium herds (50–99 cow herds) and 10% in large herds (> = 100 cow herds)

Table 1
Steps in preparation of data set cln20.

Data compilation steps	# of herds	# of observations
Initial milk shipment data file	137	95858
Remove exact duplicate records (n = 3243)	137	92615
Combine multiple shipments within one day into a single daily total (5393 combined into 2615 records)	137	89837
Drop duplicate records on consecutive days (n = 164)	137	89673
Drop all milk shipment data after 26 November 2015 (n = 9257) and (38) herds (n = 20,760) with no herd size information	99	59656
Merge with herd managment file and drop (5) herds (n = 646) with no management information	94	59010
Data cleaning steps		
Identify observations meeting supect rule #1 (n = 886)		
Identify observations meeting supect rule #2 (n = 2226)		
Identify observations meeting supect rule #3 (n = 2806)		
Drop observations flagged by any of the preceding three rules plus two shipments on either side (n = 7240)	94	51770
Drop one herd subsequently identified as having seasonal calving (n = 627)	93	51143

Suspect rule 3 - Identify observations where the total milk shipped per cow per day changed by more than 20%, 15% or 10% in small, medium and large herds, respectively.

For each suspect observation, that observation plus two shipments before and after were dropped from the data set. This cleaning process produced a data set subsequently referred to as cln20.

However, the removal of suspect artifacts in the milk shipment data was very subjective in nature, so two additional cleaning processes were carried out. The first used more relaxed thresholds for identifying herds with suspect changes in milk shipped (25% − 20% − 15% for small, medium and large herds) respectively and produced a data set referred to as cln25. The second used a more stringent set of cleaning rules (15% − 10% − 5%) and produced a data set referred to as cln15. Because of the potential for the period immediately after a transition from non-AMS to AMS milking to be highly variable in terms of milk shipped per cow per day, the cln20 data set was modified so that production data for 60 days post-transition were removed (data set called cln60/20). A sensitivity analysis of the cleaning process was carried out by running comparable models on the different data sets.

2.4. Descriptive statistics

Histograms of the proportional change in milk production (milk shipped per cow per day) between shipment dates, before and after the removal of suspect shipments were generated (Fig. 1). Descriptive statistics were computed for the data set using both the milk shipment file (n = 51,143) and a file based on herd averages (n = 93) (Table 1).

2.5. Development of regression models

We explored the use of lowest level of variance of bulk tank milk shipment volumes as a definition of resilience. Random effects linear regression models were used to evaluate the variance in milk shipped per-cow-per-day and to assess if this differed between production system. This evaluation was carried out by comparing residual variance estimates derived from a variety of ways of stratifying the residual variance. Fixed effects for production system, herd size, year (4 levels), season (sine/cosine function), region (3 levels) and breed (4 levels) were forced into all models, as was a random effect for herd. Given that the herd's AMS status might have a substantial impact on milk shipment weights and that the status changed in 27 out of 93 herds during the study period, a new variable was created that split herds which switched status into 2 sub-herds. Given that there were only (on average) 1.29 sub-herds per herd, all of these split records were treated as separate herds and are referred to as such from here on.

To account for the repeated measures nature of the data, a correlation structure was applied to the residuals. For reasons of computational efficiency an auto-regressive (AR1) structure, which assumes equally spaced shipments, was used for all model exploration and model building. Final models were re-fit using an exponential (power) correlation structure which accounts for unequal inter-interval lengths.

In order to determine if production system affected within-herd variation in milk shipment volumes, a model in which the lowest level variance was not stratified was compared to those in which the variance was stratified by either production system or both production system and AMS status.

Models were compared using Akaike information criteria and nested models were compared with likelihood ratio tests. Finally, distributions of residuals were examined to evaluate model assumptions. This included identifying outlying points and use of plots to evaluate normality and homoscedasticity of residuals at both the herd and milk-shipment level.

Initial regression diagnostics revealed some serious problems with the models. First, all models had more "large" (> 3 or < -3) shipping day residuals than expected (0.27% are expected if the data follow a normal distribution) (Table 3). Second, there were 6 herds with extremely large random herd effects (5 with < -5 and 1 with > 5) indicating herds that were shipping much less or much more milk than expected. Given the potential for these herds to be very influential in the regression model, they were removed and two new data sets created (cln20r and cln15r).

3. Results

3.1. Descriptive statistics

Prior to data cleaning, but after merging of the different data sets, the data set cln20 consisted of 59,010 shipment dates from 94 herds over the period of 1 April 2012 to 26 November 2015. Data cleaning removed 7867 (13%) observations and one complete herd but the time

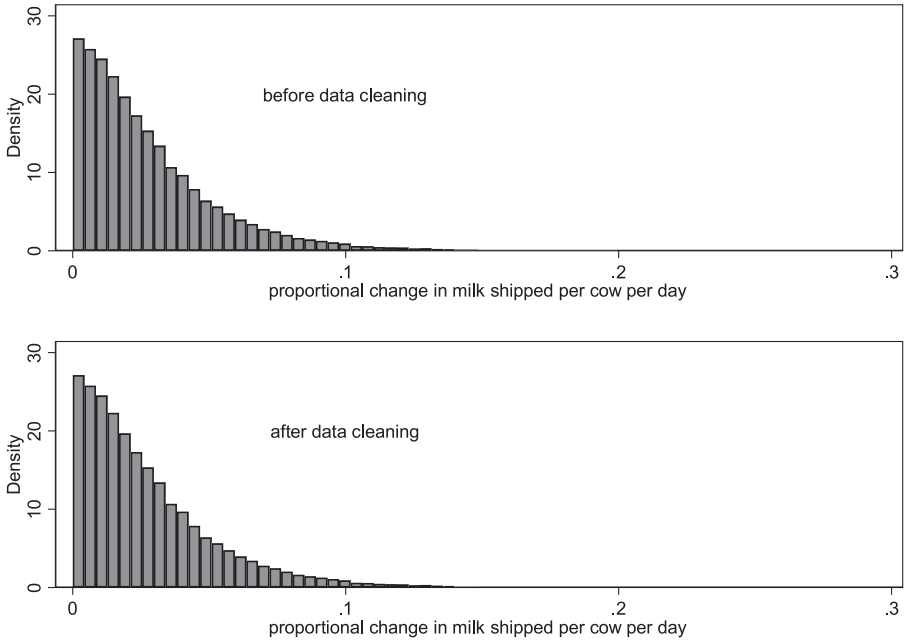

Fig. 1. Proportional change in milk shipped per cow per day before and after data cleaning. (See text for description of data cleaning process).

Table 2
Herd level characteristics of the 93 study herds in data set cln20. Herd size, milk shipped and within herd SD are averages over the study period.

		automated milking system (ams)			
Production system	parameter	no	transition during study	AMS for full period	Total
Non-Organic	number of herds	29	8	2	39
	herd size[a]	89	111	204	99
	milk shipped[b]	28.91	30.35	28.68	29.19
	within-herd SD[c]	2.11	2.28	2.07	2.14
Organic	number of herds	20	19	15	54
	herd size[a]	80	76	72	76
	milk shipped[b]	26.26	26.6	27.19	26.64
	within-herd SD[c]	2.19	2.06	2.36	2.19
Total	number of herds	49	27	17	93
	herd size[a]	85	87	87	86
	milk shipped[b]	27.83	27.71	27.36	27.71
	within-herd SD[c]	2.14	2.13	2.33	2.17

[a] estimated number of cows milked.
[b] milk shipped per cow per day.
[c] within herd standard deviation of milk shipment weights.

interval remained unchanged. Histograms of the distribution of the outcome variable of interest (milk shipped per cow per day) are shown in Fig. 1.

Table 2 provides summary, herd-level statistics relating to milk shipped from the 54 organic and 39 conventional herds in the final data set. Organic herds were smaller (average of 76 vs 99 cows), produced approximately 2.5 kg less milk per cow per day but appeared to have similar within-herd variability in terms of milk shipped.

3.2. Regression models

Resilience was evaluated by examining the residual variance in milk shipped per cow per day under the assumption that herds with high resilience would have lower residual variance. A summary of the residual variance estimates is presented in Table 3. When comparing methods of stratification of the lowest level residuals, there was very little difference in the variance estimates among the three different stratification schemes evaluated (none, production system, AMS and production system). However, the log likelihoods of the three models were very different (likelihood ratio test comparing models 1 and 2 $P < 0.001$, likelihood ratio test comparing models 2 and 3 $P < 0.001$) suggesting that stratification by both AMS and production system status resulted in a model which fit the data much better.

To adjust for seasonal variation in milk production, time of year was included as a sine/cosine function. Other fixed effects included production type, herd size, year, region and breed.

The various data sets (cln20, cln25, cln15 and cln60/20) evaluated in the sensitivity analysis did not show any profound differences in results. The fixed effects were all very similar (results not shown). The variance estimates, measuring the resilience, (shown in Table 3) were not profoundly different among the 4 data sets. Relaxing the cleaning thresholds (data set cln25) retained an extra 1937 (3.8%) records but resulted in very little change in the variance estimates. Strengthening the thresholds (data set cln15) removed and extra 7725 (15.1%) records and substantially lowered the residual variance in non-organic, AMS herds (3.82 to 3.39). A similar, but smaller reduction (3.82 to 3.65) was seen by eliminating the data from the 60 day period following a transition from non-AMS to AMS (data set cln60/20) in non-organic herds while no reduction was seen in organic herds. This suggests that the transition may result in a period of greater instability in non-organic herds than in organic.

A comparison of models utilizing an exponential correlation structure compared to one using an ar1 structure showed a better fit (Akaike information criteria reduced from 142,623 to 142,032). However, the model with the exponential correlation structure took 40× as long to run (22 h hrs vs < 0.5 h on a Surface Pro 3 computer).

The final model (based on the cln20 data set), with the lowest level variance stratified by AMS and production system and utilising an exponential correlation structure is shown in Table 4. Among the fixed effects, year, season, breed and production system all had significant effects on milk shipped. On the other hand, region, herd size and AMS status were not statistically significant. The herd level variance (6.76) was much larger than any of the lowest level variances suggesting that there was much more variation between herds, even after adjusting for the fixed effects, than there was between shipment dates within a herd.

Using the more stringently cleaned dataset (cln15r) did slightly reduce the number of large shipping day residuals compared to the expected number (~2.4× compared to ~3.1× for cln20r), but there was still a lot of unexplained "noise" in the data. Removing the herds with large herd-level residuals did not have a substantial impact on the residual variance estimates.

Table 3
Estimated shipment-day variance from random effects linear regression models of milk shipped per cow per day, as a measure of resilience. Model 1–3 applied different stratifications, model 4-6 compared different cut-offs for removing suspect records from the data set, models 7 and 9 applied an alternative correlation structure while poorly fitting herds were removed in models 8 and 10.

Model	Dataset[a]	Correlation structure	n	Lowest level variance			
				non-organic		organic	
				no ams	ams	no ams	ams
1	cln20	ar1	51014	3.74			
2	cln20	ar1	51014	3.74		3.76	
3	cln20	ar1	51014	3.70	3.82	3.69	3.87
4	cln25	ar1	52951	3.76	3.99	3.81	4.00
5	cln15	ar1	43289	3.75	3.39	3.58	3.89
6	cln60/20	ar1	50269	3.69	3.65	3.69	3.92
7	cln20	exp	51014	3.76	4.00	3.82	3.97
8	cln20r	exp	45110	3.00	3.74	3.63	4.28
9	cln15	exp	43289	3.70	3.39	3.75	3.84
10	cln15r	exp	38702	3.71	3.84	3.58	4.14

[a] see text for description of construction of the various data sets.

4. Discussion

To be able to use secondary databases with large amounts of easy accessible data is a great asset for research purposes. Existing documentation must, however, be critically reviewed to assess the quality of the data for the intended use, and if such documentation does not exist, the researcher must evaluate the data source (Sorensen et al., 1996). Arts et al. (2002) concluded that it is unrealistic to aim for for a registry database completely free of errors and that the intended use of registry data must determine the necessary properties of the data. Exploring the raw milk shipment data revealed a number of weaknessess. Errors such as duplicate records and multiple shipments reported for the same day were of a technical nature and possible to address with a proper data cleaning procedure. It would, however, be advantageous if the data were more stringent already from the source. Large differences in milk volume between adjacent deliveries was another challenge in the data. It seemed unlikely that all changes in the delivered volume came as a result of real changes in milk production per cow milked. Some of the noise could be explained by inaccurate information on the number of cows at each delivery – a number that was extrapolated from the number of cows reported from the monthly test milking. For

Table 4
Final linear regression model evaluating the effect of selected factors on milk shipped per cow per day, based on data set cln20, with the lowest level variance stratified by both AMS and organic and imposing an exponential correlation structure on the lowest level residuals.

Variable	Level	Estimate	S.E.	95% Confidence interval
Fixed effects				
Intercept		27.53	0.78	26.01 ; 29.05
Year	2012	baseline		
	2013	0.22	0.07	0.09 ; 0.36
	2014	0.67	0.08	0.51 ; 0.82
	2015	1.17	0.09	0.99 ; 1.34
Cows[a]		0.02	0.03	−0.04 ; 0.08
Ams (vs non-AMS)	Yes	0.69	0.52	−0.34 ; 1.71
Organic (vs non-organic)	Yes	−2.81	0.53	−3.85 ; -1.78
Time of year	Cos	0.29	0.04	0.2 ; 0.37
	Sin	1.29	0.04	1.2 ; 1.37
Region	South Sweden	baseline		
	Central Sweden	−0.5	0.59	−1.67 ; 0.66
	North Sweden	−0.48	0.68	−1.82 ; 0.85
Breed	Swedish Red	baseline		
	Swedish Holstein	3.74	0.83	2.11 ; 5.38
	Mixed Red and Holstein	0.98	0.7	−0.39 ; 2.34
	Other	0.28	0.82	−1.33 ; 1.89
Random effects				
Herd		6.756	0.931	5.157 ; 8.850
Non AMS, non organic	Rho	0.947	0.002	0.944 ; 0.950
	Residual variance	3.754	0.119	3.528 ; 3.995
Non AMS, organic	Rho	0.939	0.002	0.935 ; 0.942
	Residual variance	3.819	0.116	3.598 ; 4.052
AMS, non organic	Rho	0.894	0.007	0.880 ; 0.906
	Residual variance	3.996	0.238	3.556 ; 4.491
AMS, organic	Rho	0.905	0.003	0.900 ; 0.911
	Residual variance	3.973	0.114	3.756 ; 4.204

[a] per 10 cows.

instance, changes in volume depending on cows that underwent medical treatment, were traded or slaughtered created changes that could not be adjusted for on a daily basis. Moreover, tie stalls practicing regular milking times, could be hypothesized to have the same number of milkings between deliveries whereas the number of milkings in AMS-herds most likely have a larger variation between subsequent milk deliveries. Another factor that potentially could add to the variability is herds practicing milk feeding of calves, where the drain of bulk milk for the calves can be done irregularly with reference to milk delivery times. In Sweden, own processing of milk is unusual and would most likely not have had any big influence on the study results if present.

Stratification of residual variance was done for AMS and production system, which improved the model fit markedly (visualized in Table 3, models 1–3). Hence, in line with the findings of Folke, 2006;Felleki et al. (2012) we concluded that, also when evaluating milk production, attention needs to be paid to factors which influence the variance as well as the fixed effects in the model. Further, the model fit was improved by using an exponential correlation structure compared to the more commonly used AR1 correlation structure, but the demand for computational power was high with a 40-fold increase in time required to fit a model. The results from the sensitivity analyses showed that the shipping-day variance did not change substantially whether we used the stricter or more liberal cleaning procedures. A limitation in the analysis was that the 27 herds that transitioned from non-AMS to AMS during the study period were treated as separate herds before and after the transition as they were too few to add another level to the random structure in the model. Nonetheless, given the small number of conversion herds they were not likely to have had much impact on the variance estimation.

Evaluating resilience

The market share for organic milk is increasing steadily so it is important to evaluate the impact of organic production on animal health and resilience. Previously a variety of longevity measures have been evaluated as they reflect all functional traits (Fall et al., 2008; Ahlman et al., 2011). The ability to adapt to disturbances, resilience, is different, but a rather interesting approach. Few attempts have, however, been done to evaluate resilience in dairy production. ResilienceIt is challenging to identify a suitable measure of resilience, but in line with Elgersma et al. (2018), the underlying thought behind the present approach was that disturbances would affect milk production volume less in the more resilient herds, in our case organically managed herds. For a complex dynamic system, such as milk production, it is hard to separate a genuine response to a specific disturbance from fluctuations in functionality of the system and the approach needs to be system specific and may involve a degree of subjectivity (Cumming et al., 2005; Elgersma et al., 2018;Döring et al., 2014). In the present study, one can argue that the development of rules for identification of erroneous production changes, is such a subjective component although tested by sensitivity analyses. Well aware of the shortcomings in the quality of data and the choice of measure of resilience in the present study, we could not find any support for our hypothesis, as the variance of daily milk volume per cow did not differ consistently between production system, across various data cleaning procedures, stratification strategies or models built.

Evaluating disease monitoring

The possibility to use data collected on a daily basis to provide real time information about herd health is appealing. It would provide an opportunity to detect and possibly mitigate the consequences of disease outbreaks at herd and regional levels at an early stage. The possibilities of syndromic surveilance in order to detect viral diseases such as Bluetongue and Schmallenberg in dairy herds has previously been described using milk recording data (Madouasse et al., 2013; Veldhuis et al., 2016), while milk robot data has been used for early detection of mastitis (Huybrechts et al., 2014), but to the knowledge of the authors exploring the possibilities of milk delivery data has not yet been done. Early detection of outbreaks with endemic viruses, such as Bovine respiratory syncytial virus and Bovine Corona virus, would be desirable in order to reduce the impact of, and to control disease outbreaks. However, despite thorough cleaning, the milk shipment data were rather noisy and not likely to be useful for such disease outbreak detection.

Acknowledgements

The authors would like to thank the participating farmers and Växa Sverige for good cooperation.

This research received funding from the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning, grant no: 220-2010-1875.

References

Ahlman, T., et al., 2011. Culling reasons in organic and conventional dairy herds and genotype by environment interaction for longevity. J. Dairy Sci. 94 (3), 1568–1575. http://dx.doi.org/10.3168/jds.2010-3483.

Arts, D., De Keizer, N., Scheffer, G.-J., 2002. Defining and improving data quality in medical registries: a literature review, case study, and generic framework. J. Am. Med. Inf. Assoc. JAMIA 9, 600–611. http://dx.doi.org/10.1197/jamia.M1087.

Charfeddine, N., Pérez-Cabal, M.A., 2017. Effect of claw disorders on milk production, fertility, and longevity, and their economic impact in Spanish Holstein cows. J. Dairy Sci. 100 (1), 653–665. http://dx.doi.org/10.3168/jds.2016-11434.

Cumming, G.S., et al., 2005. An exploratory framework for the empirical measurement of resilience. Ecosystems 8 (8), 975–987. http://dx.doi.org/10.1007/s10021-005-0129-z.

Döring, T.F., et al., 2014. Resilience as a universal criterion of health. J. Sci. Food Agric. 95 (3), 455–465. http://dx.doi.org/10.1002/jsfa.6539.
Elgersma, G.G., et al., 2018. Fluctuations in milk yield are heritable and can be used as a resilience indicator to breed healthy cows. J. Dairy Sci. 101 (2), 1240–1250.
Fall, N., Forslund, K., Emanuelson, U., 2008. Reproductive performance, general health, and longevity of dairy cows at a Swedish research farm with both organic and conventional production. Livest. Sci. 118 (1–2), 11–19.
Felleki, M., et al., 2012. Estimation of breeding values for mean and dispersion, their variance and correlation using double hierarchical generalized linear models. Genet. Res. 94 (6), 307–317. http://dx.doi.org/10.1017/S0016672312000766.
Folke, Carl, 2006. Resilience: the emergence of a perspective for social–ecological systems analyses. Global Environ. Change 16, 253–267.
Holling, C.S., 1973. Resilience and stability of ecological systems. Annu. Rev. Ecol. Syst. 4 (1), 1–23. http://dx.doi.org/10.1146/annurev.es.04.110173.000245.
Huybrechts, T., et al., 2014. Early warnings from automatic milk yield monitoring with online synergistic control. J. Dairy Sci. 97 (6), 3371–3381. http://dx.doi.org/10.3168/jds.2013-6913.
IFOAM, 2018. The Principle of Health. Available at:. (Accessed: 7 March 2018). https://www.ifoam.bio/en/principles-organic-agriculture/principle-health.
Madouasse, A., et al., 2013. Evaluation of a continuous indicator for syndromic surveillance through simulation. application to vector Borne disease emergence detection in cattle using milk yield. PLoS One 8 (9). http://dx.doi.org/10.1371/journal.pone.0073726.
Sorensen, H.T., Sabroe, S., Olsen, J., 1996. A framework for evaluation of secondary data sources for epidemiological research. Int. J. Epidemiol. 25 (2), 435–442. http://dx.doi.org/10.1093/ije/25.2.435.
Sverige, Växa, 2018. Husdjursstatistik 2018 - Cattle Statistics. Available at:. Date accessed 2018-03-14. https://www.vxa.se/fakta/styrning-och-rutiner/mer-om-mjolk/.
Toftaker, I., et al., 2017. A cohort study of the effect of winter dysentery on herd-level milk production. J. Dairy Sci. 100 (8), 6483–6493. http://dx.doi.org/10.3168/jds.2017-12605.
Veldhuis, A., et al., 2016. Application of syndromic surveillance on routinely collected cattle reproduction and milk production data for the early detection of outbreaks of Bluetongue and Schmallenberg viruses. Prev. Vet. Med. 124, 15–24. http://dx.doi.org/10.1016/j.prevetmed.2015.12.006.
Wolff, C., et al., 2015. Bovine respiratory syncytial virus and bovine coronavirus in Swedish organic and conventional dairy herds. Acta Vet. Scand. 57 (1), 2. http://dx.doi.org/10.1186/s13028-014-0091-x.
Wüthrich, M., et al., 2016. A case-control study to estimate the effects of acute clinical infection with the Schmallenberg virus on milk yield, fertility and veterinary costs in Swiss dairy herds. Prev. Vet. Med. 126, 54–65. http://dx.doi.org/10.1016/j.prevetmed.2016.01.021.