Sensor data for animal health and welfare: present perspectives and future applications

K. Schodl[1], M. Burke[2], R. van der Linde[2] and C. Egger-Danner[1]

[1]ZuchtData EDV-Dienstleistungen GmbH, Vienna, Austria
[2]ICAR, Utrecht, The Netherlands

Abstract

Sensor technologies measuring individual animal behaviour and physiological parameters are increasingly used in dairy farms to improve fertility and health management. These technologies produce a large amount of high-resolution data at individual cow level and thus interest in using these data exists beyond herd management. In this study, which was conducted within ICAR's Brian Wickham Young Persons Exchange Program (BWPEX) five representatives from ICAR member organisations and research institutions were interviewed to gain more insights into benefits and challenges of the use of sensor data beyond its intended purpose. The topics addressed in the interview were about

1. The greatest potential of using sensor data in general and for the interview partner's organisation specifically,
2. How sensor data is currently used in the interview partner's organisation and planned to be used in the future,
3. Which challenges exist and how they can be overcome,
4. How sensor data can be used for animal health and welfare improvement and for breeding, and
5. How important sensor data will be for the dairy industry in the future.

All interview partners attributed great potential to the use of sensor data beyond herd management and were interested in using it also in their organisations. However, several challenges were identified and although ideas on how to overcome them exist, it was concluded that the development of third-party applications or other products based on sensor data is not ready yet. Some aspects of how the data may contribute to enhancement of animal health and welfare and in a breeding context were mentioned and there was consensus that these data will play an important role for dairy industry in the future.

Keywords: cow, sensor data, animal health and welfare, breeding, interview.

Introduction

Dairy farms increasingly use new technologies such as automatic milking systems (AMS) and wearable sensor devices measuring behaviour such as activity or rumination and physiological parameters (e.g. rumen temperature) in dairy cows. Manufacturers offer these technologies in combination with software programmes for certain management purposes such as notifications for oestrus or calving detection or for health monitoring. These notifications are based on algorithms, which identify for example changes in movement patterns and relate them to a potential heat event.

However, the large amount of high-resolution data at individual cow level offers a huge potential beyond its intended purpose and may be used in research, for breeding or by milk recording organisations. ICAR is working on several aspects of the use of sensor data to investigate their potential. Currently, guidelines on the validation of sensor systems are developed by the ICAR Measuring, Recording and Sampling Devices Sub-Committee and on the use of sensor data by the ICAR Functional Trait Working Group together with the IDF Standing Committee on Animal Health and Welfare. Furthermore, the ICAR Animal Data Exchange Working Group deals with technical issues and requirements for interfaces between sensor companies and milk recording or breeding organisations, who want to obtain the data. There may still be some legal and technical challenges, which have to be overcome, but nonetheless it is important to think about use cases and added benefit of these data. Thus, the aim of this study was to disclose potential fields of applications within ICAR member organisations focusing on the aim of animal health and welfare improvement.

Material and methods

This study was conducted within ICAR's Brian Wickham Young Person's Exchange Program (BWYPEX), which supports young researchers, who work on topics important to ICAR and its member organisations, in building a network and gaining experience through visiting different member organisations in various countries. Guided interviews were conducted with five persons related to ICAR, its member organisations or research institutions between March and May 2023. Two persons were working in research institutions, one person in a breeding organisation, one person in an artificial insemination (A.I.) company, and one person for ICAR. Interviews were recorded, transcribed, and subsequently coded based on five categories, some including subcategories, which are listed in Table 1. The categories were deducted based on the interview guide. For content analysis, interview transcripts were printed, and colour coded to assign them to the single categories. Finally, each category was analysed on its own using the text snippets.

Table 1 Categories for analysis of guided interviews

Category	Subcategory
Greatest potential of using sensor data in general and specifically for the interview partner□s organisation	
Use of sensor data in the interview partner□s organisation	Status quo and purpose Plans for the future
Challenges	Identified challenges Ways to overcome them
Use of sensor data for animal health and welfare and breeding	
Future perspectives of sensor data for the dairy sector	Expectation for the future and possible developments Importance of sensor data for the dairy industry in the future

Potential applications of sensor data

Two major aspects were mentioned. First, integrating the data with other (historical) farm and cow-specific data, may increase the value of these data for farm management. Whereas these devices create a large amount of data and offer decision support in certain areas (e.g. heat detection) the real benefit only emerges when integrating it with historical information and other farm data. This way predictions for diseases, behaviour, or milk yield may improve. Based on that, smart insights for farm management may be created and if data are used across farms, this may enable benchmarking applications. Starting from that and using various other sources, added value may also be generated beyond farm level. Additionally, DHI and/or milk recording organisations may benefit by adding value to existing services or even broaden their service portfolio for farmers.

Furthermore, these data bear a great potential for use in genetic evaluation, which in most cases may be regarded as by-product of data recording programs. Given that it serves its initial purpose of herd management improvement, so farmers keep using it, data can be used for large scale phenotyping and trait definition. These may comprise completely new traits based on what the sensor is measuring or what can be predicted using integrated data sets on the one hand or the development of new proxies for complicated traits such as feed efficiency, resilience, or health traits on the other hand. Furthermore, phenotypes based on these sensor-derived data may be closer to the animal's physiology and thus improve genetic evaluation. Generally speaking, three main technologies can be used for the development of new traits - vision, accelerometer-data and mid-infrared. The latter, however, will not be able to do sensing around transition or in the dry period.

Use of sensor data in the different organisations

Current use of sensor data

On ongoing initiative aims at a validation for sensor systems, which may be regarded as a check or a guideline for several aspects of the sensor systems. It should help users to understand what kind of data changes happen between the sensor measurement and changing the measurement into an observation and how big the black box in between is. Eventually, the user should be able to understand what the system is able to do, and for which purpose the data can be used (e.g. for management or for genetic evaluation). From a scientific perspective, showing variation in the different behaviours and trying to scientifically define ranges for normal behaviour and for deviations from normality are currently interesting fields of research. Other aspects mentioned comprise how useful these data are for predictions on animal health and welfare and how they can be integrated with other data for this purpose. These basic understandings and predictions should also form the basis for the development of new traits for different purposes. For one, new technologies such as AMS may pose new demands to cows in terms of milking behaviour and thus, these can be recorded using these technologies. Furthermore, there is research being carried out on using the sensor data for fertility related genetic evaluation. The idea behind is that sensor data may be closer to the physiology of the animal and may thus yield more heritable traits than those based on breeding records.

However, more research is needed in terms of data editing and trait definition because in the current form they do not meet the quality standards. Furthermore, sensor data allow to characterize intensity of heat expression, which was shown to positively correlate with retainment of a successful pregnancy in embryo transfer recipient cows in one study. In addition to research on sensor data applications, also technology development is of interest and particularly the development of sensors based on computer vision is currently boosted. Its benefits are that it can offer a way to mimic what is otherwise visually assessed by an observer and that the sensor does not have to be attached to the animal itself, which may in some cases be considered as a painful intervention.

Use of sensor data planned in the future

Future plans address approaches for standardizing data across manufacturers and join forces among organisations in a collaborative effort to access sensor data. In this context, ICAR is expected to take the initiative to speak on behalf of its partner organisations. Furthermore, the definition of new traits and the possibility of introducing new evaluations (e.g. fertility or heat stress related) based on sensor data will be investigated. Further research will – if existing – build on preceding work revising data editing and amplifying the data set by for example increasing the number of herds. Implementation of these new traits, indices or evaluations depends not only on their performance and suitability for genetic evaluation, but also on their superiority compared to current ones in terms of higher heritabilities or lower costs for phenotyping, etc. Further ideas are to develop (genetic) models for an on-farm use, so the data does not have to leave the farm anymore. This approach is inspired by the use of phone data by companies such as Google, who passed from training their models in the cloud to directly training them on the user's phone. However, these initiatives are still at the very beginning, and it is difficult to receive funding for it.

Challenges

Data security and ownership

Before thinking about any further use of data from sensor systems these data must be available to the person or institution interested in working with it. First, there is the legal side of data ownership. Who is the owner of the data and what are the conditions for using the data? This is not as straightforward as it may seem. Farmers using the sensor systems are the supposed owner of the data generated on their farms and thus, data cannot be shared without the farmer's consent. Furthermore, data must be treated confidentially – it must be clear who is using the data for what. Farmers may be hesitant in sharing these data due to the fear of misuse, unauthorized sharing with third parties or the emergence of disadvantages for them. However, even with farmers' consent access to the data is still depending on the agreement of the manufacturer companies, who may be reluctant to share these data. Most of them are large companies acting at a global scale, which may sometimes imply the lack of a clear policy on how to handle data sharing with interested third parties or changes in existing policies due to takeovers by other companies. Another scenario is that manufacturers charge a fee for data provision, which may present a challenge for research initiatives and business opportunities with unknown outcomes. In many countries, companies doing genetic evaluation used to have access to a lot of data at no or low cost (e.g. milk recording or classification data) because they were recorded at national levels with a lot of subsidies and genetic evaluation was often considered a by-product of these (herd management) data. However, this may change when it comes to sensor data from private companies and breeding organisations may have to pay for these data in the future.

Validity and quality of sensor data

Assuming that access to data is granted, further challenges arise. Parameters and values generated by the sensor are not clearly defined and information on how the raw measurements of a sensor (e.g. an accelerometer) are changed into the output parameter are mostly not accessible due to intellectual property. This is even more true for alarms, which are generated based on the measurements without disclosure of thresholds or algorithms. However, for deriving traits for breeding from these data it may be important to get access to this information to better understand what exactly is being measured. Furthermore, information on the accuracy of alerts generated by the system is also lacking. For the sensor parameters themselves it is not clear how accurate the data is and if the output parameter is measuring what it claims to measure. Hence, further use of the data may require validations tools or regular calibrations of the sensor. This opens up another important issue concerning reference values and what can be considered normal, let alone which expression is desirable – more or less active cows for example. The sensor information is often very generic, and the challenge is to associate these non-specific measurements to anything related to e.g. welfare. Irrespective of the final purpose of the sensor data application – be it herd management, breeding, or welfare assessment – there is need for a solid reference value and gold standard. This relates to the sensor parameter itself in terms of what exactly the output parameter is measuring but also to the definition of the trait or disease, which should be predicted by the sensor values. Having a clear understanding of the predicted phenotype is particularly important if this information is further used for genetic or genomic prediction models. However, often the real diagnosis is hard to get or difficult to detect (e.g. silent estrus) and additionally frequencies are too low to obtain enough data and reliable results. On top of these cross-sectional data, there is need for longitudinal data to assess repeatability over time.

Besides those general requirements of data validity and accuracy there is another important aspect of data quality – measurement errors and outliers. The system may break down or the sensor may run out of battery generating faulty data or no data at all. In commercial farms there are also a lot of management related sources for missing or erroneous data. The replacement of an empty battery or a lost device on a cow will be depending on available time of the farmer or the urgency of the sensor to work because for example the cow is going to be up for breeding soon and thus the heat alarm system is needed. However, if the cow was already bred then changing the sensor may not be first priority to the farmer.

Furthermore, these systems are intentionally created as a management tool and thus alerts, or other information is optimized to serve this purpose rather than to correspond to the correct physiological trait. Taking heat alarms as an example, they may be intentionally prolonged indicating the time window for a successful breeding rather than the physiological duration of an oestrus. Moreover, an alert may be generated based on changes in the sensor parameters although the indicated event should be impossible (e.g. heat alerts during pregnancy). Other disturbances than heat or calving events or diseases such as social interactions between animals or the use of synchronization protocols, respectively can influence cow behaviour or physiological states and thus create system alerts. Thus, correctly identifying these irregularities in the data is another challenge for any further application and in addition to adequate skills in data science it requires a lot of domain specific knowledge.

One important step towards solving these issues and creating added value is the integration of the sensor data with other farm (management) records and historical data. However, this entails several other challenges. First, integrating these animal-individual data, which are available at daily or even hourly resolution, with additional data and using them in smart applications requires a lot of space for data storage as well as computational capacity. Furthermore, correctly matching the sensor data with other animals-specific records can also be quite challenging, particularly regarding the correct animal identification. Whereas this may be easier to solve for wearable devices, which can be assigned to an individual animal, this is more challenging for installed systems working for example with computer vision techniques. While these are well performing in assessing lameness in cows, the positive identification of the correct animal is much harder.

Standardisation between sensor systems

Even if those requirements were met for individual sensor systems, there is still the issue of lacking standardisation between systems of different manufacturers. Parameters may be called the same and intentionally measuring the same thing (e.g. rumination or activity), but they differ in measurement, definition and algorithm between devices or even between devices by the same manufacturer. Some parameters are expressed in time units (e.g. minutes of rumination per hour or during 24 hours) and thus their values seem more comprehensible whereas for example activity is often expressed as a dimensionless value, which cannot be related to any known scale or unit. With the intention of using these data for breeding or other purposes, which involve the use of data across many farms, standardization or comprehensive definition of those parameters or traits presents an important issue. Even though models may be able to correct for some of the variation this may still not be enough to harmonize these measurements.

Much experience with these data lies within research organisations and while there is more research needed this does not necessarily always require new data. Revising already existing data from new perspectives is one very important task, which may also need more resources in terms of manpower and financing of research activities. However, it is very difficult to get this kind of research funded because it may lack novelty and other important grant criteria. Moreover, not many people stay in this field of research for a longer period of time to continuously work on these topics.

Sensor data in research

To define normal variation, either phenotypic or genetic, data has to be explored in a neutral way. Although researchers' intention is to be objective in their work, underlying values, mindsets, or presumptions may bias this neutrality. Moreover, data science and domain knowledge need to be combined to yield best results. Domain knowledge is needed before any modelling as well as afterwards and results may have little value for application without it. This may also comprise the documentation of domain knowledge in a way, which is interpretable by an algorithm, so it can feed into the modelling process and make the 'black box' of this process more comprehensive.

Usually, ICAR member organisations do not have the capacity for undertaking this research by themselves, which makes the involvement of research organisations and universities even more important. This requires an open dialogue between all the involved parties, to ensure that developed traits are relevant for implementation and economically interesting. When it comes to traits and trait definition finding a gold standard for reference is a key element. Revising scientific literature on different aspects of the trait in question and relating it to the sensor measurements is a starting point. Furthermore, using reference herds with a lot of detailed phenotypes and sensor data and subsequently validating the results at larger scale may present a good strategy. In case of traits with a low frequency it is helpful to increase the data set and the number of herds, if this is possible. The challenges relating to data quality may best be addressed in a collaborative effort, which may comprise an open documentation of data cleaning and editing when publishing results. If researchers share their experiences with various types of sensor data, follow-up or other research may benefit a lot and may be able to start later in the process instead of 'reinventing the wheel'. Combining these efforts may also help in terms of research funding.

However, none of this research will be possible without access to sensor data and in particular to sensor data from commercial farms. Issues around data ownership suggest that the whole setting is very complex and there are different interests at stake. ICAR as an umbrella organisation is expected to take the initiative to speak on behalf of its members and try to negotiate with sensor manufacturers. Communicating a clear purpose to manufacturers may help in these negotiations.

Use of sensor data for (breeding for) animal health and welfare

One of the most obvious positive impacts on animal health and welfare may lie in the potential to early detect cows with potential health issues, which allows for prompt intervention. Even if farmers are very attentive and monitor animal health closely, sensors may detect problems before they are visible to the farmer. Early detection and intervention allows for the chance of reducing suffering of the animal as well as costs due to diseases. Especially farms with large herds may profit from these alerts so they can filter out animals, which may need treatment and pay more attention to those specifically.

From a research perspective these sensor data may help us to understand more about the normal behaviour of cows by exploring (normal) variations due to parity, breeds, age, etc. or diurnal patterns. Similar to lactations curves, we may deduce patterns of rumination or activity over a whole lactation period. Understanding these patterns and knowing when for example rumination time may be higher or lower is important to differentiate physiological from pathological states. If these relationships are well understood, data-driven assessments can be used beyond herd or farm level to assess and report health and welfare across farms, farm types, regions, or climatic areas. When it comes to welfare assessment, these data may help to monitor welfare continuously instead of just at certain points in time. Benchmarking tools together with extension services may function as a tool for health and welfare improvement on farms.

In terms of breeding, these data can be used to define proxies for health traits, which may be closer to the animal's physiology. This could enable genetic selection for more resilient animals or animals, which are more robust towards certain diseases.

Future perspectives for sensor data in the dairy sector

The overall impression was that sensor data will be highly important for the dairy sector in the future. Farm management can be improved along with the opportunity for objective monitoring of animal welfare. However, the usefulness of sensor technology for the farmer must be priority, otherwise they will stop using it. Furthermore, the farmer as a mutual client presents an important link between the sensor company and any third party using the data and developing applications for farmers. At this point however, the information from these technologies is not yet ready for the development of routine applications and organisations should be careful not to promise solutions too soon. Or, as one of the interview partners put it: “You can only ring the bell once. And if that bell isn't positive, you cannot undo it.”

Acknowledgements

The first author would like to thank ICAR's Brian Wickham Young Persons Exchange Program for the opportunity to visit ICAR and partner organisations and conducting this research. Furthermore, I thank my employer ZuchtData for allowing me to take part in this program and ICAR and ICBF for hosting me. Finally, I want to thank my interview partners for taking the time for the interviews and their valuable contributions to this work.