Improving dairy herd performance with data integration: A guide to continuously monitor feed efficiency, income over feed cost, and other key metrics

T. E. Silva[1,*], S. R. Wangen[2], M. Li[1], F. Zhang[1], Y. Gong[1] and V. E. Cabrera[1]
[1]*Department of Animal and Dairy Sciences, University of Wisconsin-Madison, WI 53706, USA*
[2]*American Family Insurance Data Science Institute and Wisconsin Institute for Discovery, University of Wisconsin-Madison, Madison, WI 53715, USA*
[*]Corresponding author: Tadeu Eder da Silva, tdasilva2@wisc.edu

Abstract

Feed efficiency (FE) in milking cows is squarely linked to profitability due to its positive correlation with income over feed cost (IOFC). In turn, IOFC is an indicator that ties nutrition to profitability of dairy operations. However, monitoring these key performance indicators (KPI) on the farm continuously and consistently can be difficult as it necessitates merging data from various sources. Thereby, data integration through the application of data engineering techniques becomes mandatory to achieve it. We have been successfully developing a robust system that integrates data from disparate on and off-farm data sources and provides FE, IOFC, and other KPIs on a continuous basis. Thus, the objective of this work is to describe the process involved in developing such a system and to demonstrate how it helps farmers to make decisions in a more timely manner. The calculations require integration of data from: 1) farm's feeding system; 2) the milking parlor system; and 3) milking processor report to correct the milk yield to a standard value of energy and milk prices. All data collected is parsed from their native file formats and loaded into a data warehouse. The FE, IOFC, and other KPIs are made available in a user-friendly report to the farmer. Thus, this work demonstrates that it is possible to constantly integrate disparate sources of data and add value to data in dairy farms, helping farmers to get improved decisions. Furthermore, it demonstrates that the active involvement of the farmer in the process makes its results more robust and meaningful.

Keywords: data engineering, decision-making, feed intake, income over feed cost, milk yield

Introduction

Feed efficiency (FE) is used at the field level as a proxy to measure the relative ability of milking cows to convert feed nutrients consumed into milk (Bach et al., 2020). Despite all FE limitations as a performance metric (Connor, 2015), higher FE in milking cows is squarely linked to higher profitability because it is positively correlated with income over feed cost (IOFC). In turn, IOFC is an indirect indicator that ties nutrition to the profitability of dairy operations, hence supporting short-term decisions (Atzori et al., 2021). Therefore, monitoring both FE and IOFC of milking cows helps the farmer and other decision-makers make better decisions at the farm gate. Moreover, higher FE can be associated with a reduced environmental footprint from dairies (Løvendahl1 et al., 2018).

Mathematically, FE is calculated by dividing milk yield (kg or lbs) by the feed intake (dry matter basis; kg or lbs). However, on-farm FE monitoring is quite challenging, as it requires integrating data from different sources, which frequently prevents farmers with limited resources from obtaining this key performance indicator (KPI) in a more timely manner. In this way, FE calculation requires integration of data from: 1) farm's feeding system, which provides the feed delivery information (at pen level); 2) the milking parlor system, which provides the milk yield records (at pen or animal levels); and 3) milking processor report to normalize the milk yield to a standard value of energy (energy corrected milk; ECM) based on the average milk fat and protein contents. Additionally, to perform IOFC calculations, the farm-specific milk price settlement rules and prices are needed, in addition to the cost of the diet offered to lactating cows.

Thereby, data integration, decoding, cleaning, and homogenization through the application of data engineering techniques become mandatory to make data streams from contemporary dairy operations available consistently (Wangen et al., 2021). Along these lines, we have been successfully developing a robust and flexible system that accesses and integrates data from disparate on and off-farm data sources and provides FE and IOFC calculations (in addition to other KPIs such as total milk income, average milk yield per cow, milk composition, etc.) on a continuous basis.

Beyond data engineering techniques, another key component is the active involvement of the farmer or other decision-makers in the whole process (collaborative development), which supports the system greater chances of success with regard to outputs that are generated (data quality, business rules, etc.) and ultimately its use to support the on-farm decision-making process (Baldin et al., 2021; Rose et al., 2016).

Thus, the objectives of this paper are 1) to demonstrate the feasibility of constantly integrating disparate sources of data and adding value to it in dairy farms, helping farmers to get better insights and improved decisions; and 2) to demonstrate the importance of having the farmer involved in the whole process of data integration and communication by developing a more robust, meaningful, and user-friendly data reporting system.

Materials and methods

Background and procedure

The information and data shown in this paper are part of the main deliverables obtained with the progress of the Dairy Brain project. The Dairy Brain project is an initiative supported by the USDA that aims to demonstrate how multiple data sources can be connected and how, based on those, it is possible to deliver on-farm advanced analytics to generate value and innovate the decision-making process (Cabrera et al., 2020; https://DairyBrain.wisc.edu).

It is worth emphasizing that what is presented in this paper is the active partnership with a farmer enrolled in the Dairy Brain project (i.e., a single farm). What is presented here are preliminary results subject to future changes and improvements. The identity of the farmer is kept anonymous as a matter of protecting the privacy and sensitive information.

Throughout the process we tried to follow some principles of the Agile methodology for software development, such as focusing on our customer's satisfaction (i.e., the farmer), total involvement and engagement, based on frequent conversations, with continuous deliveries, prioritizing simplicity and relevance (Balaji and Murugaiyan, 2012).

Data sources and integration

The data sources required to calculate the FE and IOFC are milking parlor records (daily milk yield), feed system records (total feed delivered, feed refusals, and feed costs), milking processor (milk composition and milk settlement rules), and official milk and milk components prices.

In order to calculate the FE, we normalized the milk yield for a standard value of protein (3.2%) and fat (3.5%), according to the equation 1:

$$ECM = [0.327 \times Milk\ Yield\ (lbs)] + [12.95 \times Fat\ (lbs)] + [7.65 \times Protein\ (lbs)] \quad (1)$$

Where: ECM = energy-corrected milk.

Units in calculations and reports follow what dairy farmers use in the US. For example, milk yield data and some other measurements taken on the farm (i.e., feed intake) are presented in pounds (lbs: 1 lbs = 0.456 kg) and milk price is presented in $ per hundredweight; $/cwt (1 cwt = 100 lbs). It is important to mention that the feed efficiency calculations involved only the feed intake of milking cows, therefore not accounting for the intake of dry cows.

In addition to the data sources mentioned above, to calculate the IOFC, it is necessary to obtain the average price of milk sold ($/cwt). The average price of milk is calculated according to the milk class to which it belongs (i.e., class III), its component prices (i.e., protein, butterfat, other solids, and somatic cell count (SCC)), producer price differential (PPD) for its pool (i.e., upper Midwest), and the processing company price settlement rules (i.e., premiums and penalties). Component prices and PPD are collected from an Application Programming Interface (API) which is made available by the USDA through the Agricultural Marketing Service on a monthly basis (www.ams.usda.gov/resources/marketing-order-statistics). However, those prices are usually not published on the API until the end of the next month. In order to get a timely assessment of the current daily or weekly IOFC, we calculated the predicted milk price according to the historical PPD and component prices. Once those prices are published on the API, we update the IOFC according to the settled prices. Thus, IOFC is calculated as the equation 2:

$$IOFC\ (\$/cow) = Milk\ Yield\ (lbs) \times Milk\ Price\ (\$/cwt)\ /\ 100 - Feed\ cost\ (\$/cow) \qquad (2)$$

The IOFC calculations in this work do not include the cost of feeding dry cows.

As can be seen in Figure 1, data from the milking parlor (milking parlor backups – Postgres dump files) and the feeding system (feeding system backups – comma-separated values format) are automatically generated and collected on-site, packaged, and periodically uploaded to the Dairy Brain Agricultural Data Hub (AgDH; Wangen et al., 2021) via a minimal software installed on a farm computer. On the other hand, the data with the amounts of milk produced and sold, as well as its composition (e.g., fat, protein, and other solid contents), SCC, and milk urea nitrogen (MUN) are daily emailed by the processor (Excel spreadsheet files) directly to the AgDH.

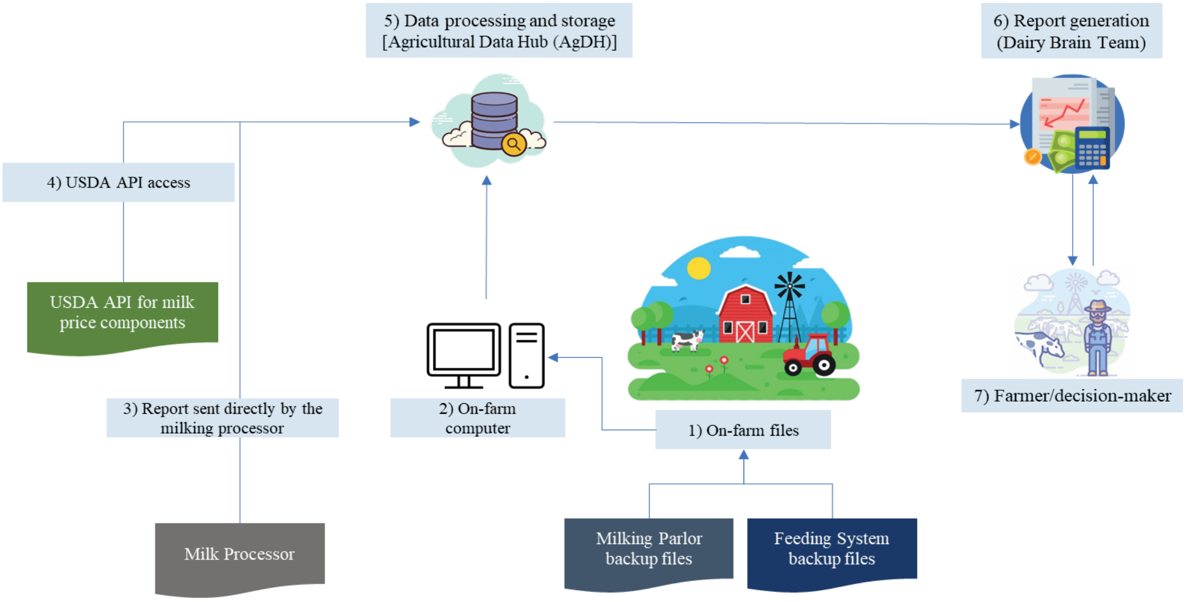

Figure 1: Schematic describing the process to make on-farm data sources available (milking and feed data (1)) via data collection [on-farm computer (2) and processor report (3)], USDA API access (milk price components (4)); data processing and storage (Wangen et al., 2021) (AgDH; 5); report development (farmer report generation (6); and farmer decision-making (7).

Report development and interaction with the farmer

Before starting the actual discussions with the farmer and kicking off the report development containing all indicators and data visualizations, we set up an on-farm computer with all scripts required to collect the data we need. For this, we visited the farm and installed the computer that would run all the scripts and receive the data to be transferred to the Dairy Brain AgDH. Once verified the data was flowing smoothly and without unexpected interruptions, we visited the farmer to align how the report development process would be, emphasizing how significant the farmer's involvement and engagement would be for this development.

We all agreed to have monthly meetings to discuss what has been developed and what needs to be developed. For instance, we discussed as a team what KPIs to include, how to validate them, and deciding what is the best manner to display them. Furthermore, as the integration process involved other stakeholders (i.e., milking processor), the active participation of the farmer facilitated our interaction with them to obtain faster results.

Additionally, we discussed the frequency of updating the report that would be made available to the farmer online. Although part of the data is collected on a daily basis (i.e., milking parlor and feeding system records), we decided to update the report on a weekly basis due to the fact that the farmer thought this timeframe would suffice the decision-making support at that moment.

Results and discussion

Report organization and sections

The report with all KPIs and other visualizations, which is updated and made available to the farmer weekly has 5 sections: Overview, milk data, feed data, feed efficiency, and IOFC. Each section has a specific purpose, with different ways of displaying the data. The overview section aims to demonstrate in a more summarized way the farm's KPI's at a glance, bringing both a snapshot of economic and performance indicators (Figure 2). This part of the report allows the farmer to get a general overview of the overall farm's KPIs and identify any disturbances that may be subject to correction. As can be seen in Figure 2, in addition to FE and IOFC, other indicators related to the farm economics (milk price and feed cost) and productive performance are also displayed such as energy corrected milk production, feed intake, milk composition, etc.

On the other hand, the other sections aim to explore more in detail each indicator and its respective breakdowns. In these other sections, there are time trends in the KPI's (or in the indicators that compose them), which could be used to understand any disturbance observed in the overview section and take some corrective action in a timely manner to reduce economic losses. Furthermore, these sections can be used as a way to investigate the impacts of changes in management practices (e.g., nutritional and feed management) over time and verify their effectiveness. Due to space limitations in this paper, we will not explore in depth all the details of all other sections.

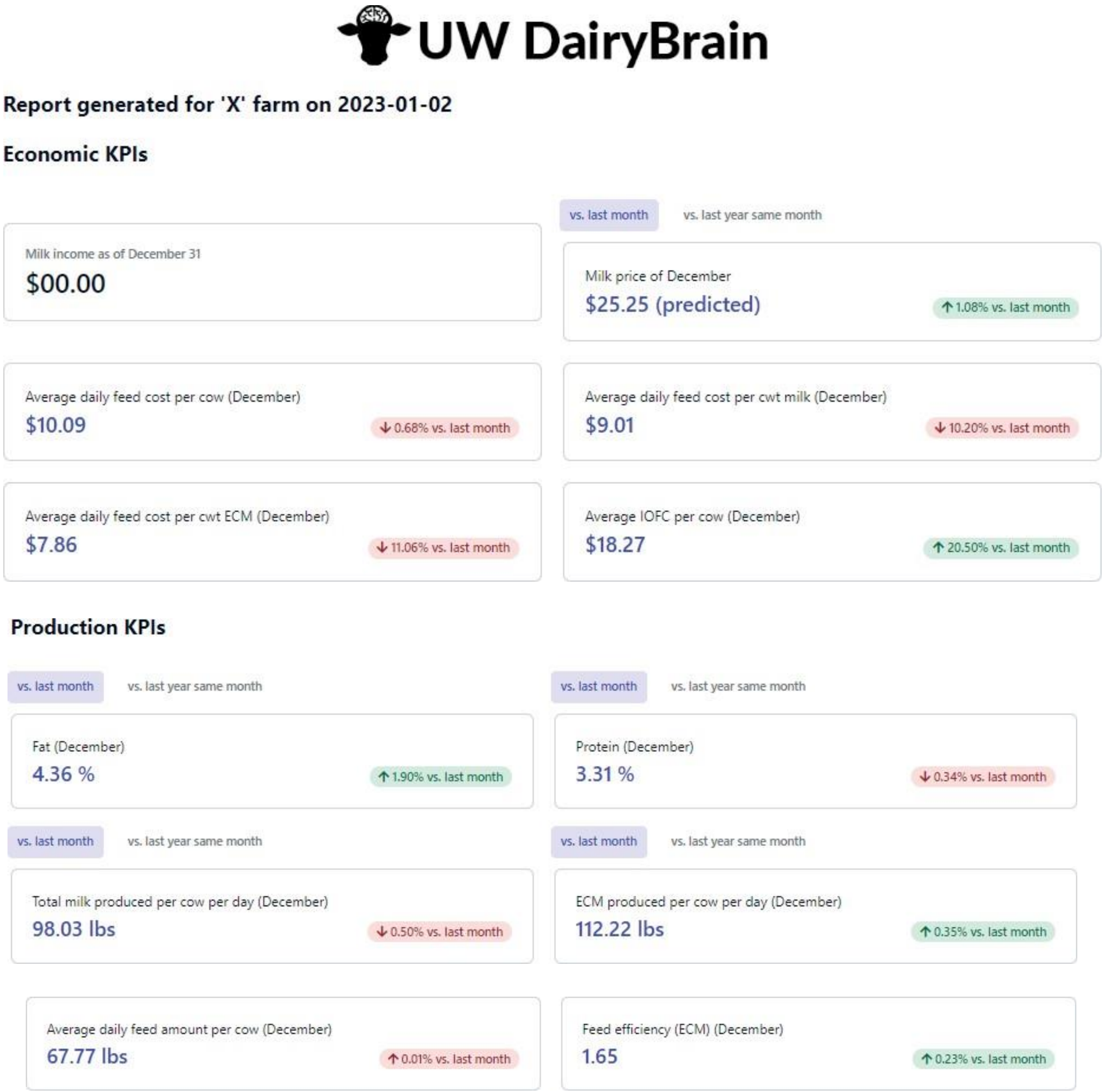

Figure 2: Screenshot of the report's overview section. Each card displays the average values of each indicator for the month and also displays a comparison with the previous month or alternatively the same month of the previous year.

Figure 3 is an example of the analyses that can be performed by the farmer in the other sections. In this particular case, we can observe variations in 3 indicators throughout the year: observed IOFC, IOFC with fixed milk price, and feed costs. The inclusion of the IOFC with a fixed milk price was a request from the farmer since the farmer uses it as a way of understanding whether the variations observed in the IOFC throughout the year could be due to changes in the cost of feed. Furthermore, the farmer has a good idea of what the farm's IOFC would be by selling it at a standard milk price. For instance, as can be observed in Figure 3, from the month of July/22, the observed IOFC began to fall, and at the same time, the feed cost began to rise. However, from October/22 on, the IOFC rose again, even with feed costs at higher levels (IOFC with milk price fixed stable). This leads us to infer that these late IOFC increases could be attributed to the increase in milk prices since October/22.

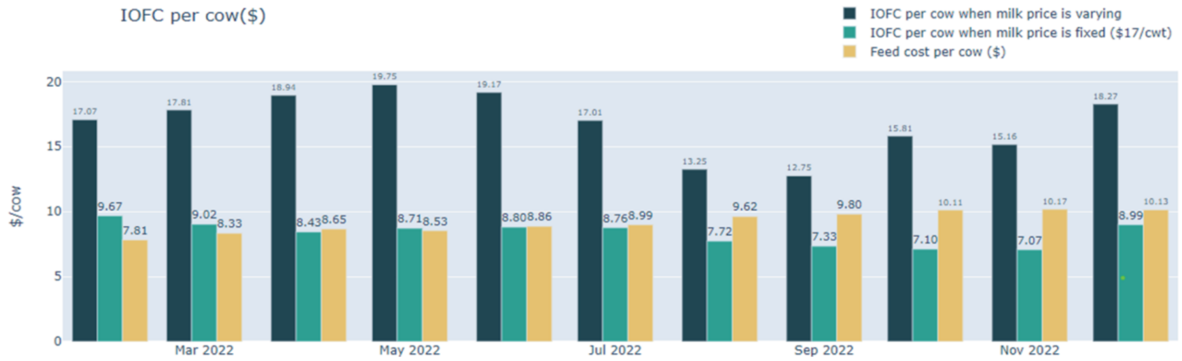

Figure 3: Screenshot of the IOFC report section. Dark green, green and yellow bars represent the IOFC observed as is (milk price and feed costs varying), IOFC with fixed milk price (17 $/cwt), and feed costs per cow per day ($), respectively.

A very frequent complaint pointed out by dairy farmers is that sometimes they have the data they need, but this is very dispersed and diluted in countless other types of reports (e.g., pdf, screenshots, spreadsheets, etc.), which makes the task of getting information very difficult. Thus, it should be noted that one of the major gaps filled here with this proof-of-concept work was the fact that we continuously integrate a lot of relevant information for the farmer in just one place. Therefore, it helps the decision-making process because not only valuable information is generated due to the data integration process, but also because it saves time for the producer in the search for information.

Insights and decision-making changes with the report

One of the main potential advantages of monitoring indicators on a more continuous basis is the possibility of observing trends in the data from the moment that new on-farm management practices are adopted. In this sense, the producer reported to us that, together with the farm nutritionist, they had changed some nutritional strategies for the year 2022 in order to reduce the drop in milk protein during the summer (it was not revealed to us what specifically was changed).

Figure 4: Screenshot of the monthly average milk protein content (%) from 2019 to 2022 from the single farm in the study.

As can be seen in Figure 4, in all 4 years plotted there is a seasonal drop in milk protein content in the summer months. However, in 2022, the intensity of this expected drop was much lower when compared to 2021, 20220, and 2019. This indicated that the practice carried out in 2022 seemed to work well and potentially could be repeated next year.

Team experience and professional development

Another important aspect that should be highlighted in this interaction between the Dairy Brain team and the farmer is the opportunity that is offered to graduate students to work directly with dairy farmers. From the point of view of professional training, this implies better preparation of students for a possible career as extension workers. This becomes even more important given that many animal agriculture students come from urban areas and therefore not having much experience characterized by growing up on a farm (Daigle, 2016). In this way, they can practice their communication skills, decoding information, understanding, and being problem-solving oriented. The student-farmer interaction connects the student to the world of work, allowing them to learn by solving real-life problems and leveraging their more generic and practical skills, which is in line with current education approaches such as problem-based learning (Abbey et al., 2017).

In this regard, all responsibility regarding the development, maintenance, and improvements of the report was given to a Ph.D. student (co-author Yijing Gong). Furthermore, the student was responsible for leading meetings with the farmer, demonstrating all progress and developments in detail in the report. The student not only practiced her hard skills such as programming and statistical analysis, but also acquired many soft skills such as teamwork, communication, and presentation skills. For example, lessons were learned about how to set each other's expectations, and how to best display information to make it easiest for the farmer to get sufficient information and become more efficient in the iterative process of developing the integrated tool.

Conclusions

Although the challenges are substantial with different layers of complexity, this work demonstrates that it is possible to constantly integrate disparate sources of data and add value to data in dairy farms, helping farmers to get better insights and improved decisions. Furthermore, it demonstrates how important the active involvement and engagement of the farmer in the process as a whole is, making its results more meaningful and useful.

Acknowledgments

This project was supported by the Food and Agriculture Cyberinformatics and Tools grant no. 2019–68017–29935/project accession no. 1019780 from the USDA National Institute of Food and Agriculture.

References

Abbey, L., Dowsett, E., and Sullivan, J. (2017) Use of problem-based learning in the teaching and learning of horticultural production. *Journal of Agricultural Education and Extension* 23, 61-78.

Atzori, A.S., Valsecchi, C., Manca, E., Masoero, F., Cannas, A., and Gallo, A. (2021) Assessment of feed and economic efficiency of dairy farms based on multivariate aggregation of partial indicators measured on field. *Journal of Dairy Science* 104, 12679-12692.

Bach, A., Terré, M., and Vidal, M. (2020) Symposium review: Decomposing efficiency of milk production and maximizing profit. *Journal of Dairy Science* 103, 5709-5725.

Baldin, M., Breunig, T., Cue, R., De Vries, A., Doornink, M., Drevenak, J., Fourdraine, R., George, R., Goodling, R., and Greenfield, R. (2021) Integrated decision support systems (idss) for dairy farming: A discussion on how to improve their sustained adoption. *Animals* 11(7), 2025.

Daigle, C.L. (2016) In Search of the Urban Cowboy: The Need to Incorporate Animal Husbandry into the United States Higher Education Curriculum and Its Implications for Production Animal Welfare. *Frontiers Veterinary Science*, 3, 84.
Connor, E.E. (2015) Invited review: Improving feed efficiency in dairy production: challenges and possibilities. *Animal* 9, 395-408.
Løvendahl, P., Difford, G.F., Li, B., Chagunda, M.G.G., Huhtanen, P., Lidauer, M.H., Lassen, J., and P. Lund. (2018) Review: Selecting for improved feed efficiency and reduced methane emissions in dairy cattle. *Animal* 12, s336-s349.
Rose, D.C., Sutherland, W.J., Parker, C., Lobley, M., Winter, M., Morris, C., Twining, S., Ffoulkes, C., Amano, T., and Dicks, L.V. (2016) Decision support tools for agriculture: Towards effective design and delivery.
Agricultural Systems 149, 165-174.
Balaji, S., and Murugaiyan, M.S. (2012) Waterfall vs. V-Model vs. Agile: A comparative study on SDLC, *International Journal of Business and Management*, 226-230.
Wangen, S.R., Zhang, F., Fadul-Pacheco, L., Silva, T.E., and Cabrera, V.E. (2021) Improving farm decisions: The application of data engineering techniques to manage data streams from contemporary dairy operations. *Livestock Science*, 250, 104602.