Quantification of dairy farm energy consumption to support the transition to sustainable farming

1st Tamara Todic
Electronic and Electrical Engineering
University of Strathclyde
Glasgow, UK
tamara.todic@strath.ac.uk

2nd Lina Stankovic
Electronic and Electrical Engineering
University of Strathclyde
Glasgow, UK
lina.stankovic@strath.ac.uk

3rd Vladimir Stankovic
Electronic and Electrical Engineering
University of Strathclyde
Glasgow, UK
vladimir.stankovic@strath.ac.uk

4th Jiufeng Shi
Data Science
Discovergy GmbH
Heidelberg, Germany
jfs@discovergy.com

***Abstract*—As the need for using energy-efficient machinery escalates, energy consumption estimation plays an important role in decision support and planning in the agri-sector. Within the present research study, energy consumption in dairy farms was examined. A deep learning-based load disaggregation approach was used to develop data-driven models to quantify individual energy consumption of milk production-related devices of dairy farms, from a single aggregate measurement. According to the experiments conducted on three dairy farms in Germany, load disaggregation from a single aggregate meter is a viable, cheaper alternative to submetering multiple pieces of equipment to accurately quantify electricity consumption at scale in dairy farms in order to provide the decision support needed to inform measures for tackling climate change.**

***Index Terms*—NILM, industrial NILM, energy disaggregation, sustainable farming**

I. Introduction

The agri-sector, especially dairy farming, is undergoing close scrutiny as one of the key contributors to climate change. According to the United Nation's Food and Agriculture Organisation (FAO), the global consumption of dairy has increased and the industry's environmental footprint as well. Whilst methane is considered the key contributor to greenhouse gas emissions, energy consumption to drive machinery, using a combination of diesel and other fuels and electricity, cannot be neglected. Furthermore, with increasing financial pressure on the industry, and increased demand, there is a drive on electrification of machinery that are energy-efficient, including automatic milking via milking robots. Therefore quantification of energy consumption provides a useful measure for decision support and planning in the agri-sector.

Traditionally, prediction of energy consumption in dairy farms is carried out via modelling, which requires a large number of variables such as water temperature and milking times, currently not readily available at scale [1]. More recently, machine learning has been used for energy consumption prediction since it requires coarser parameters for empirical modelling, whilst still achieving acceptable prediction accuracy. For example, Support Vector Machine was used to predict the annual electricity consumption of 16 Irish dairy farms [2], where electricity consumption was measured, at 15 minute intervals, for a year totalling 20,314 kWh. Annual analysis of electricity consumption of automatic milking in [3] highlighted that milking robot, air compressors and milk cooling were the biggest electricity consumers accounting for, respectively, 33%, 26%, and 18%, of total electricity consumed in the 7 farms being metered, where 111 kWh/cow/year is consumed just for milking, which means that milking 20 cows consumes as much electricity as an average UK mid terrace house or flat.

In the above studies, all the components of the automatic milking process still need to be metered separately, which is expensive to install and maintain, and not scalable. With advances in smart metering, this paper leverages on virtual submetering, via machine learning based load disaggregation, of the electricity consuming processes on a dairy farm from a single meter. This, purely software-based load disaggregation, is also referred to as Non Intrusive Load Monitoring (NILM). In the literature, NILM methodologies are meant to identify when individual loads are used (classification problem) and quantify energy use of the individual pieces of equipment contributing to the aggregate load (regression problem). Commercial NILM has mostly focused on high frequency sampling of the order of kHz, in order to estimate consumption of individual loads in factories. There are fewer studies that focus on cheaper and low frequency sampling (less than 1 Hz), that is akin to national smart meter rollouts. To the best of our knowledge, there has only been one attempt at leveraging on NILM classification, without quantification of energy consumption of milk production equipment, for dairy farms [4], [5].

The paper is organised as follows. First, a background on

This research was funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 955422.

low frequency NILM and its relevance for dairy farms is presented. Next, we analyse a dataset of 3 dairy farms followed by the experimental setup in Section IV. Results are presented in Section V before we conclude in Section VI.

II. Background on NILM for dairy farms

The state of the art in terms of methodology and datasets for low-frequency NILM based on deep neural networks is summarised in a recent review [6], and shows that there has been significant progress in designing NILM methods for residential buildings and that data sampling intervals below 10 seconds, large field of view, usage of GAN losses, and post-processing, provide the most accurate results (in terms of classification and estimation accuracy). However, a recent review for NILM in commercial buildings [7], indicates that progress in commercial NILM, at any sampling rate, is slow because of the following challenges arising in non-residential buildings. In the commercial setting, there is usually large number of loads, with widely varying wattage and duty cycle, many of which tend to operate continuously. Additionally, there are limited datasets with electrical measurements in non-residential buildings, which further hampers progress. In order to progress NILM-based load disaggregation in non-residential buildings, the recommendation was therefore to use a hybrid of aggregate measurements and submetering, improved labelling of equipment to produce useful datasets and additional contextual knowledge such as type of equipment and occupancy patterns.

Load disaggregation attempts in the agri-sector are even further limited, impeded by the aforementioned factors, but steps are being taken to remedy the situation. Energy demand of a poultry feed factory in Brazil was disaggregated using low-frequency measurements in [8]. The resulting IMDELD dataset was made open access, including recordings from 8 types of industrial equipment. Two methods for load disaggregation were discussed: Factorial Hidden Markov Models (FHMM) and Deep Learning-based model (WaveNILM). WaveNILM outperformed FHMM with lower Signal Aggregated Error (SAE) (0.1 ± 0.2). Identification of on/off states of the machines resulted in classification accuracy of F_1-Score of 0.93 ± 0.07 for WaveNILM and 0.79 ± 0.12 for FHMM. In [4], two different approaches for NILM classification are compared, one using LSTM model, and the other using One-Directional Convolution Layer-Bidirectional GRU Recurrent Neural Network, to detect milk cooling and vacuum pump, for two dairy farms in Germany. Unfortunately, the dataset was not made public. The two approaches were compared using precision and F1-score. The 1DConv-BRNN model had better performance - a precision value of 0.86 and 0.69 for milk cooler and vacuum pump, respectively, compared to the LSTM-based model that produced a precision values of 0.81 and 0.56 for the two devices, respectively. Both models found classification of the vacuum pump more challenging, which is because it was only used for a few minutes per day, and hence there was not enough data to train the models. The proposed models were used only to classify the loads, and no energy disaggregation was performed; therefore no information was provided about ability to estimate energy consumption. Similarly, [5] investigates the suitability of deep learning based algorithms for NILM classification on dairy farms, exploring combination of one-dimensional convolutional neural network and bi-direction recurrent neural network, a combination of one-dimensional convolutional neural network and LSTM layers and a neural network composed of multiple LSTM layers. Firstly, only active power measurements are considered as input, then active and reactive power measurements combined as a 2-channel input. On/off state detection is performed on data acquired from four dairy farms in Germany. The four appliances that were disaggregated were: milk cooling, vacuum pump, milk pump and cleaning automatic machine. Multi-feature models were only trained for vacuum pump and cleaning automatic machine, due to unavailability of reactive power measurements for the other two appliances. The recordings were not continuous - various appliances were available on different dates at different farms. Classification performance of the models considered varied for different farms and dates, but using reactive power as an input feature improved performance in general. Hence the recommendation was to use different deep learning models for different loads.

TABLE I: Active power statistics in Watts [W].

Metered label	mean [W]	std [W^2]	max [W]
Location 18			
Vacuum pump	1102.51	1213.46	7174.089
Compressor	1657.61	1759.94	10400.775
Milking robot	1800.75	1783.49	11700.839
Cleaning	1405.21	1429.06	3343.873
Aggregate	5786.24	5128.80	29790.972
Location 20			
Milking robot	268.44	455.31	1718.91
Aggregate	955.20	1415.29	6374.792
Location 21			
Water treatment	1643.16	1808.47	7625.831
Pipe cooler pump	2423.58	2233.27	10939.672
Water pump	3030.16	2510.94	11942.944
Dunging milking robot	128.00	131.79	1183.782
Aggregate	7224.92	6447.26	29800.336

From the above review, we can conclude that there is limited progress on NILM for non-residential, commercial settings, and especially less so for dairy farms where only load identification was investigated and useful quantification of energy consumption of milk production equipment was not looked into. We address this gap in this paper by analysing the type of equipment present (relatively larger and more complex loads than present in residential loads) in dairy farms and leveraging on advances in low-frequency deep-learning based regression NILM models, we quantify the energy consumption of individual pieces of milk production equipment to support endeavours towards energy efficient farming.

III. Understanding energy consumption in dairy farms

The dataset used in this paper was obtained from German dairy farms, which were continuously metered from 01.02.2020 until 05.03.2021. The recorded data contain active

power measurements in Watts, in 3 phases, and were provided at sampling intervals of 1 second. Measurements were available for 21 locations. At each location, a number of loads were sub-metered, and were identified from their labels. Loads directly related to milk production include milking robots, compressor, vacuum pump, as well as water pump, pipe cooler pump and water treatments for milking robots. These loads were located only in 3 locations (Locations 18, 20 and 21), which we selected for disaggregation. Measurements for these loads were present on only one phase, hence the other two phases were discarded. Aggregate data is only available for location 20. For locations 18 and 21, aggregates were created by summing all the metered loads at each location, including the ones that are not directly related to milk production. The loads available at these 3 locations, as well as the statistics - mean value, standard deviation and maximum of the active power for those devices, as well as aggregate at each location, in Watts, are shown in Table I. These indicate a large variability of these loads and difficulty to perform standard prediction. More importantly, it can also be seen that, even the milk robots, common to all 3 locations, have very different operating power possibly because of the differences in the devices themselves, or due to the fact that farmers may be metering (labelling) milking robot as the robot alone or as a combination of the robot and accompanying devices.

Figure 1 shows monthly consumption of power for milking robots at locations 18 in blue, 20 in red and 21 in brown colour, presented in kWh. It can be seen that the milking robot from location 18 consumes the largest amount of power among all the robots, and the robot from 21 the smallest. It was expected that there would be seasonal patterns and that the consumption would be largest during spring and summer, which is not the case, except maybe for the robot from location 21. From the recorded data, we also observe that the robot at location 18 was always on, unlike the milking robots at the other two locations. Robots at locations 20 and 21 consume significantly lower amount of power when on, compared to the robot at 18. All the 3 robots are working during the night.

Figure 2 shows consumption by device in MWh at location 18. Milking robot consumes the largest amount of power among all the devices - 30.21%, while compressor, cleaning and vacuum pump consume 27.85%, 23.41% and 18.53%, respectively. These figures are inline with the observations of [3] for Irish dairy farms. Total annual consumption at this location is 57.60 MWh. As shown in Figure 3, at location 21, the smallest part of active power is consumed by milking robot - only 1.68%, while 42.02% of total active power is consumed by water pump, 33.52% by pipe cooler and 22.79% by water treatment. Total annual consumption at this location is 64.39 MWh.

IV. Disaggregation: Experimental setup

A. NILM Regression model

We leverage on one of the best performing low frequency deep learning based regression networks as highlighted in [6], namely the WaveNet-based algorithm of [9]. Good performance was demonstrated in [9] on household appliances, which are very distinct to those observed in milk production loads. One of the model's major strengths is that it has a large field of view, produces concatenated and processed outputs from multiple layers in the network, each with different field of view, which makes this model available to recognise patterns on multiple scales. This is beneficial since duration of active use times of loads vary significantly. In order to replicate the architecture of [9] to develop individual models for milk production equipment, measurements from the dataset were resampled so that the new sampling interval is 10 sec. After resampling, there are a total of 3,438,721 samples of measurements per device during the entire recording period.

The default model parameters were kept, as they were reported as best performing in [9]. The input and output sequences are of the same length - 127 samples. They are obtained from the data time series by sliding a window with a step of the size 100 samples. The model uses 32 skip channels and the batch size used is 128. The loss function used for training was mean absolute error (MAE). The training was always performed via 300,000 iterations with decaying learning rate. Since the devices were switched on most of the time, on/off classification was not performed.

Fig. 1: Monthly power consumption from milking robots.

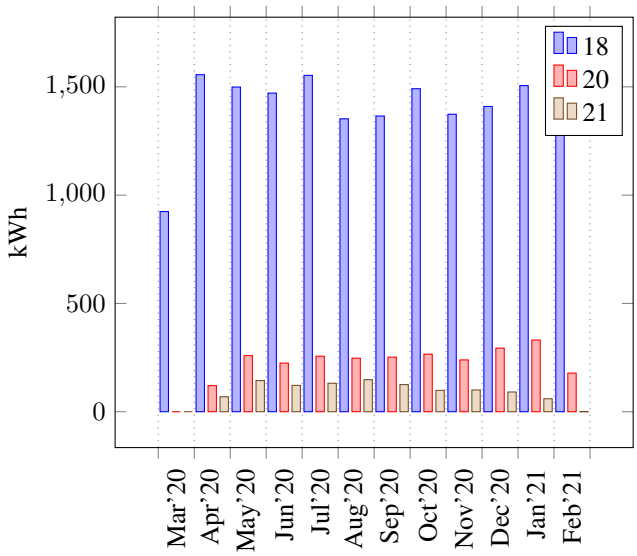


Fig. 2: Annual consumption by device at location 18 in [MWh].

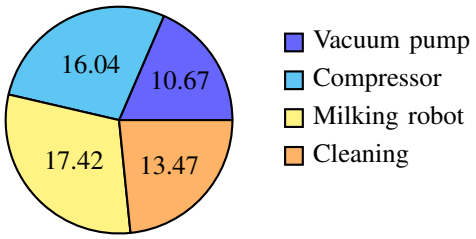
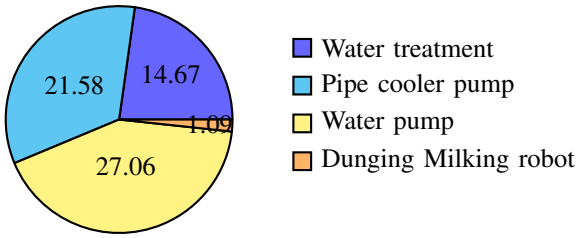


Fig. 3: Annual consumption by device at location 21 in [MWh].



B. Train/Validation/Test split

Since the devices present at the three locations are different, the data from each of them was split into train, test and validation folds, and trainings were run using the data from only one location at a time. The data was split as follows.

For all three locations, training time was from 01.02.2020 00:00:00 to 31.12.2020 23:59:59, validation from 01.01.2021 00:00:00 to 30.01.2021 23:59:59 and test from 31.01.2021 00:00:00 to 04.03.2021 23:59:59. The measurements were split such that all devices of interest are used in train, validation and test parts of recordings. This leads to 84.17% of data used for training, 7.54% for validation and 8.29% for testing.

C. Performance evaluation metrics

For the evaluation of model performance, common metric that is used for regression-based energy disaggregation tasks is MAE:

$$\frac{\sum_{i=1}^{N} |y_i - x_i|}{N}, \tag{1}$$

where N is the length of the sequence used, y_i is the predicted power consumption for sample i and x_i is the target, i.e., the ground-truth device consumption value from submetering. This metric is used during the process of training, but since it is not very informative for regression-based disaggregation problem, other metrics have also been considered in literature.

The match rate (MR) metric evaluates the performance based on the overlap of true and estimated energy. Its range is between 0 and 1 - the closer to 1, the better the performance, and is defined as:

$$\frac{\sum_{i=1}^{N} min\{x_i, y_i\}}{\sum_{i=1}^{N} max\{x_i, y_i\}}, \tag{2}$$

where N is the length of the sequence used, x_i is the target and y_i predicted power consumption.

D. Transferability evaluation

Since milking robot is the only load that is present at all three locations, transferability of models from one location to another was tested only for this device. The models were trained and validated on one location and tested on the other unseen locations. The inference was done only on the test part of data from all locations, so that performance can be compared with the models that were trained and tested at the same location.

V. Experimental results & Discussion

Table II presents the performance metrics, MAE and MR, that represent the accuracy of estimated energy consumption of the listed milk production equipment. Besides the Dunging milk robot, all other individual equipment in the dairy farms could be disaggregated accurately with above 95% match rate. Dunging milking robot is hard to disaggregate, because it consumes a much smaller amount of power compared to the aggregate signal, so the model does not effectively discern and learn its signature. This can be seen in Figure 5 - the pattern of milking robot is easily recognizable in the aggregate signal at location 18, slightly more difficult at location 20 and hardly at location 21. A similar observation was made by [7], where small loads cannot be detected when operating at the same time with large-load equipment. It can also be seen from Table II that using only MAE as performance metric can be misleading as MAE for high loads (e.g., vacuum pump and compressor) tends to be high despite very high MR indicating successful disaggregation. Similarly, a low MR for Dunging milking robot indicate poor disaggregation result, which is not reflected by MAE. Figure 4 shows the actual and predicted power consumed by devices at locations 18 and 21 during the entire test period. At location 18, consumption for cleaning equipment is slightly overestimated, while for the other devices, consumption is slightly underestimated. At location 21, pipe cooler and milking robot are slightly overestimated while water treatment and water pump are slightly underestimated. These are inline with the MR values of Table II, except the Dunging milking robot. Figure 5 also confirms that it would be futile to use MAE metric alone as a performance measure - looking at its values in Table II would lead to conclusion that the worst performance would be for the robot at location 18, and the best at 21, which is in contrast with actual results, while MR values gave better insight into actual results.

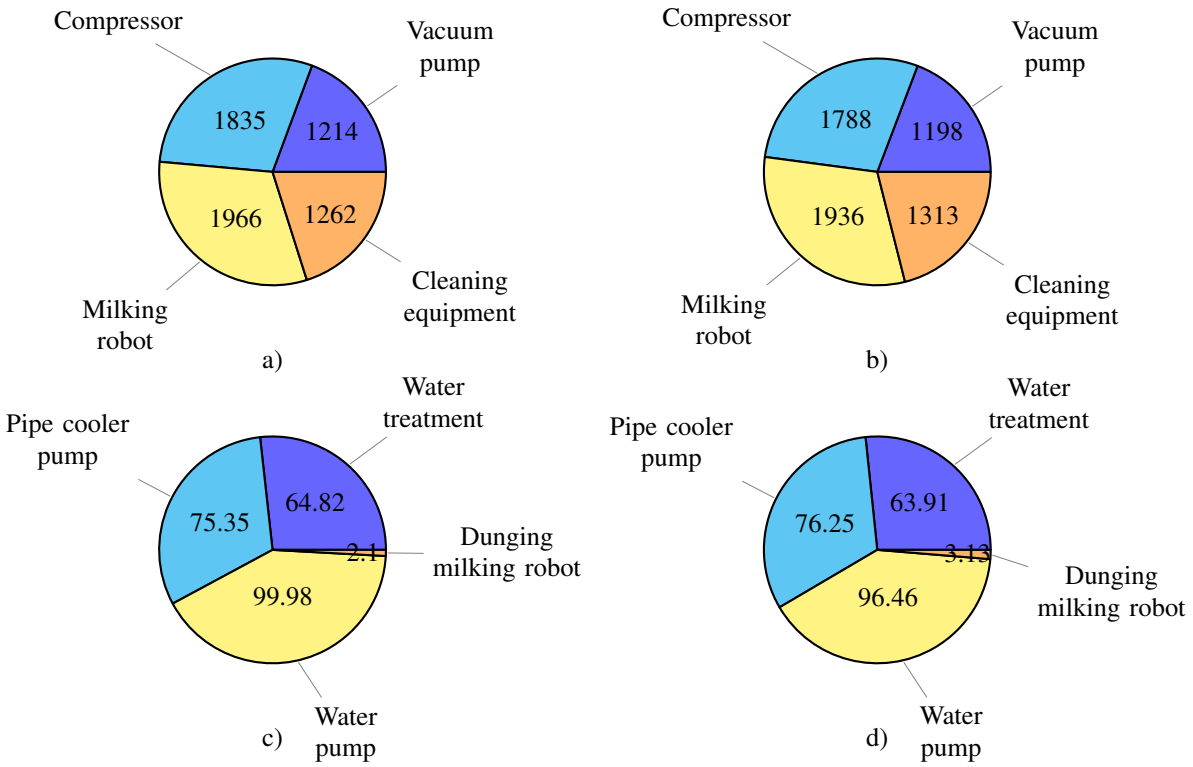
Overall, the NILM architecture of [9], originally developed for residential buildings with very different load profiles, is therefore capable of accurately separating various milk production related equipment present at dairy farms, which can be seen by low MAE and close to 100% MR values.

TABLE II: Disaggregation results: Training and testing on the same location data.

Machine	MAE [W]	MR [%]
Location 18		
Vacuum pump	44.94	97.0942
Compressor	68.12	97.0673
Milking robot	65.83	97.3630
Cleaning equipment	78.45	95.2953
Location 20		
Milking robot	11.24	95.8923
Location 21		
Water treatment	1.30	98.4282
Pipe cooler pump	1.28	97.7119
Water pump	4.53	96.4370
Dunging milking robot	3.22	35.0197

Table III presents the match rate when the milking robot model is trained on one location and tested on an unseen

Fig. 4: Consumption by device during the test period in kWh. a) Actual consumption at location 18. b) Predicted consumption at location 18. c) Actual consumption at location 21. d) Predicted consumption at location 21.



milking robot in another location. As expected, due to large differences in wattage of the milking robots as shown in Table I and Figure 5, transferability of learnt models from one farm to another is not possible.

TABLE III: Transferability test results: MR results when training on one location and testing on another.

		Train location		
		18	20	21
Test location	18	97.3630	55.9016	24.0567
	20	70.2020	95.8923	24.5175
	21	43.4069	2.8097	35.0197

VI. Conclusion

This paper addresses the gap in the industrial non intrusive load monitoring (NILM) and agri-farm sector in scalable, data-driven, accurate quantification of energy consumption in milk production equipment. Leveraging on a deep learning based architecture, originally designed for NILM for residential buildings, we develop individual regression models for a range of high consuming automatic milking devices. The models are validated for three dairy farms in Germany, which have distinct equipment wattage and total energy consumption. Excellent performance is obtained when models are trained on the same farm data and tested on unseen data from the same farm. Transferability to other farms is an issue since labelling of individual pieces of equipment varies, and devices themselves have distinct power consumption signatures at different farms or devices. Hence, the need for load disaggregation via one smart meter, monitoring all devices in the farm (or a more expensive unified standard of submetering) cannot be overstated for scalable data-driven analysis of energy consumption in farms.

References

[1] P. Shine, J. Upton, P. Sefeedpari, and M. D. Murphy, "Energy consumption on dairy farms: A review of monitoring, prediction modelling, and analyses," *Energies*, vol. 13, no. 5, p. 1288, 2020.

[2] P. Shine, T. Scully, J. Upton, and M. Murphy, "Annual electricity consumption prediction and future expansion analysis on dairy farms using a support vector machine," *Applied Energy*, vol. 250, pp. 1110–1119, 2019.

[3] J. Shortall, B. O'Brien, R. D. Sleator, and J. Upton, "Daily and seasonal trends of electricity and water use on pasture-based automatic milking dairy farms," *Journal of dairy science*, vol. 101, no. 2, pp. 1565–1578, 2018.

[4] A. Yadav, A. Sinha, A. Saidi, C. Trinkl, and W. Zörner, "Nilm based energy disaggregation algorithm for dairy farms," in *The 5th International Workshop on Non-Intrusive Load Monitoring (NILM'20)*, (Virtual Event, Japan. ACM, New York, NY, USA), pp. 16–19, November 18, 2020.

[5] A. Sinha, "Multi-feature based development of a power-disaggregation algorithm for dairy farms," Master's thesis, Technische Hochschule Ingolstadt, 2021.

[6] P. Huber, A. Calatroni, A. Rumsch, and A. Paice, "Review on deep neural networks applied to low-frequency NILM," *Energies*, vol. 14, no. 9, p. 2390, 2021.

[7] A. Meier and D. Cautley, "Practical limits to the use of non-intrusive load monitoring in commercial buildings," *Energy and Buildings*, vol. 251, p. 111308, 2021.

[8] P. Bandeira de Mello Martins, "Non-intrusive industrial load monitoring on a factory in brazil," Master's thesis, Universidade Federal do Rio de Janeiro, 2020.

[9] J. Jiang, Q. Kong, M. D. Plumbley, N. Gilbert, M. Hoogendoorn, and D. M. Roijers, "Deep learning-based energy disaggregation and on/off detection of household appliances," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 15, no. 3, pp. 1–21, 2021.

Fig. 5: Aggregates, actual and predicted active power signatures in Watts for milking robots at all 3 locations.