Reviewing non-intrusive load monitoring using a pilot study of an IoT device to disaggregate energy usage

Reviewing Non-Intrusive Load Monitoring using a Pilot Study of an IoT Device to Disaggregate Energy Usage

Matthew McCrory[1], Adele H Marshall[1,2], Aleksandar Novakovic[1,2] and

Geoffrey Collins[3]

[1] School of Mathematics and Physics, Queen's University Belfast, Belfast, United Kingdom
[2] Joint Research Centre in AI for Health and Wellness, Ontario Tech University, Oshawa, Canada
[3] Northern Ireland Water, Belfast, Northern Ireland, United Kingdom
{mmccrory12, a.h.marshall, a.novakovic}@qub.ac.uk, Geoffrey.Collins@niwater.com

Abstract Non-intrusive load monitoring (NILM) disaggregates energy consumption data collected from a single measurement point into appliance-level data. This process facilitates energy savings. Most studies treat NILM as a residential task with few considering its application in industry. By chronologically reviewing existing literature this paper presents a review of the latest research in NILM, focusing on its potential employment within a utility company, Northern Ireland Water. A practical example of NILM is also provided using data collected by a pilot IoT device where the benefits of NILM are exhibited via a cost analysis. Results from the literature review show deep learning models to be the most recent preferred disaggregation approach. Furthermore, the standardization of evaluation metrics is deemed essential to facilitate the comparison of different disaggregation models. Finally, the NILM tool kit is outlined as a useful platform for Northern Ireland Water to practically implement NILM.

Keywords: Non-intrusive load monitoring, energy consumption, IoT, pilot study.

1 Introduction

Northern Ireland Water (NIW) has the largest electricity bill in Northern Ireland, totaling around £30m per year in 2021. This expense is due to the energy intensive procedures implemented to clean and pump water. Astonishingly NIW spends the same amount on electricity as a combined average of almost 40,000 UK homes. Consequently, this is driving innovative approaches to save energy through new technologies. This paper investigates the method of non-intrusive load monitoring (NILM) [1], an energy load disaggregation system, by exploring its suitability, development, and application using a piloted IoT device.

Using only the aggregated data from a single measurement point, NILM can both identify the operational state and predict the power consumption of individual appliances. An effective NILM system could potentially identify the most energy intensive appliances in NIW operations without the cost of installing individual smart meters. NIW could subsequently evaluate each element of the water treatment process and make informed decisions on the appliances value relative to its energy consumption.

Figure 1 exhibits a household NILM system. Fluctuations in power arise following the change in operational state of each appliance.

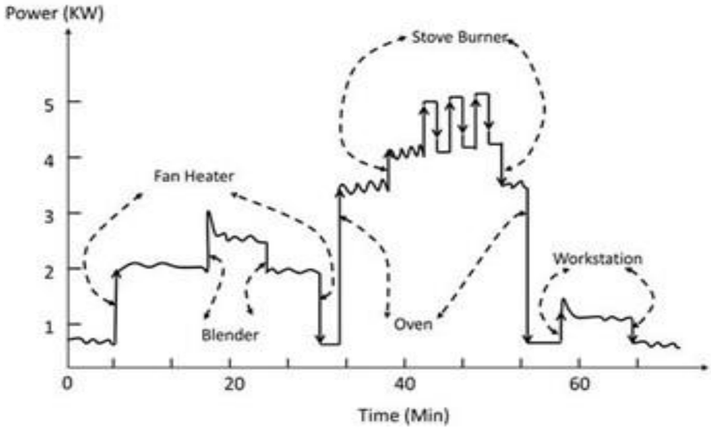

Fig. 1. Residential application of NILM (taken from [2]).

The unique power consumption of each appliance is referred to as the 'load signature' which is used to group appliances into one of four types (Table 1).

Table 1. Appliance Types.

Type	*Description*	*Examples*
TYPE I	Two states.	Toaster
TYPE II	Multiple number of states.	Washing machine
TYPE III	No fixed number of states as consumption varies constantly.	Dimmer-lights
TYPE IV	Constant state.	Smoke alarm

The goal of this paper is explore the literature to identity the cutting edge research in NILM with the aim of making recommendations for its potential implementation by NIW. A practical example using a piloted IoT device provides data that is utilized in cost analysis to assist the quantification of the practical benefits of NILM.

The remainder of this paper is arranged as follows. Section 2 provides a literature review of the data, learning approaches, and evaluation metrics involved in addressing the research questions. Section 3 gives a brief description of the pilot IoT device, the results of which are presented in Section 4. Section 5 discusses the potential for the NILM implementation in the NIW setting and Section 6 concludes the paper.

2 Literature Review

NILM was introduced by George Hart in 1992 [1]. Since then, advancements in deep learning, metering equipment, costs and internet of things (IoT) devices has caused a surge in recent research. Given the availability of data, early NILM studies utilised low-frequency residential datasets [3-4]. These datasets continue to dominate modern literature, however, the creation of industrial datasets [5-7] has seen more studies consider NILM within industry [6-7]. The increased sophistication of deep learning models has also influenced NILM literature with recent studies moving away from machine learning approaches in favour of deep neural networks, namely, convolutional and recurrent neural networks (CNN and RNN).

Amongst existing literature, Angelis et al. [8] are one of the few authors to provide an extensive evaluation of the current state-of-the-art for NILM. Their paper utilises findings from over 200 studies to emphasize the evolution of NILM datasets, learning approaches and evaluation metrics. Results showed the prevalence of residential studies as well as the clear gap in performance between deep learning and machine learning models, with the former providing state-of-the-art disaggregation results. Similarly, Pereira and Nunes [9] provide a thorough review of the main datasets, metrics, and tools for evaluating the performance of NILM systems and technologies. Results indicated several barriers that make performance evaluation challenging including missing data and limited labelling. This paper also acknowledges a lack of energy estimation studies and how research breakthroughs are only viable if the NILM research community integrate contributions in a common framework.

Aside from these evaluation-type papers most NILM studies focus on the implementation of a NILM system, often including a variety of machine and deep learning disaggregation models in a bid to find the most effective model.

2.1 NILM Data

NILM Datasets. Rapid advances in IoT device capabilities and metering equipment has led to a recent spike in the number of publicly available NILM datasets. Despite this, most datasets primarily use residential data with only three using industrial data (Table 2). Datasets typically include electrical features at both an aggregate and appliance level with the latter considered 'ground-truth' data, used for NILM evaluation.

Table 2. Breakdown of the three industrial datasets. Features include, Power (P), Reactive Power (Q), Apparent Power (S), Voltage (V), Current (I).

Name	Year	Number of appliances	Features	Time Period (days)	Aggregate sampling	Appliance sampling
ASF	2015	4	P	68	1 Hz	1 Hz
IMD	2018	8	P,Q,S,V,I	111	1 Hz	1 Hz
HIPE	2018	10	P,Q,S,V,I	92	1/5 Hz	1/5 Hz

Data Collection. IoT devices and electric meters collect data at either a high (1kHz and above) or low (below 1kHz) frequency [6]. Whilst higher frequencies allow for more easily distinguishable load signatures, existing studies gave good results for low frequency datasets also [10]. Shin et al. [11] however claim it is possible for a sampling rate to be too low suggesting a rate of at least 1-3Hz to avoid load signatures being 'destroyed'. Data for NILM tasks is not always available and is often time consuming to collect. Atypical for a time series problem, the creation of synthetic data is an unlikely solution [12-13]. Most recently, Quy et al. [13] used four different interpolation techniques to augment high frequency NILM data from low frequency datasets.

Feature Extraction. The goal of the feature extraction process is to derive a signature (feature vector) that can be used to uniquely identify each appliance. The features included are classified as either steady-state or transient (Figure 2). Steady-state features exist when an appliance is operating with a steady-state (usually when a low sampling rate is used). Active Power is conclusively the most popular steady feature however, Bonfigli et al. [14] insist Reactive Power is equally as important with its inclusion found to improve the NILM algorithms predictive accuracy. Transient features are extracted when the signal evolves in an unpredictable way, typically when a high sampling rate is used for example in transient power and transient voltage noise.

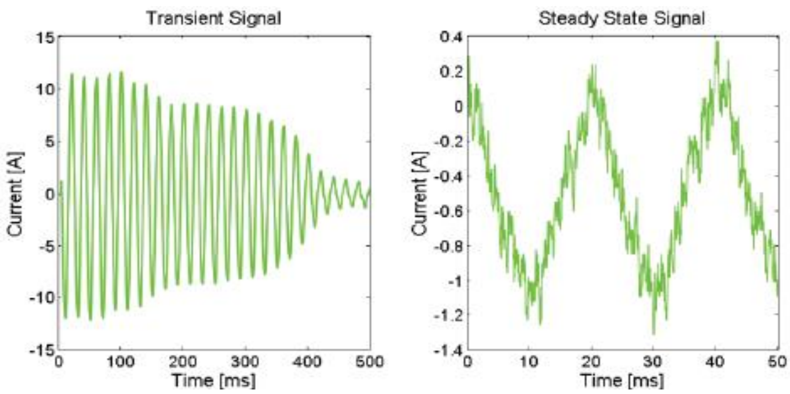

Fig. 2. Transient and steady-state signals (taken from [15]).

Studies have achieved good results using both feature types. Reddy et al. [16] recommend that a hybrid approach of both feature types will lead to an improved appliance signature. Results from this study saw the 'feature fusion' algorithm outperform the steady-state and transient feature algorithms by 9% and 15% respectively.

There is no 'optimal' number of features proposed. Both [17-18] used three features whilst two of the three industrial studies used Active Power as a solitary feature.

2.2 Machine learning approaches

NILM is formulated as a supervised learning task with a goal of predicting the power consumption or state (ON/OFF) of each appliance. Similarly, NILM is considered a classification or regression problem with each study using a machine or deep learning approach as the foundation of their model. Machine learning was the most popular approach for early studies, with the Hidden Markov Model (HMM) featuring prominently [3-4]. More recently, deep learning has moved to the forefront of NILM literature with CNNs and RNNs regarded as the state-of-the-art learning approaches.

NILMTK. Developed by Batra et al. [19] NILMTK is an open-source tool kit that includes a range of data sets, pre-processing algorithms, benchmark disaggregation algorithms and accuracy metrics. The purpose of NILMTK is to enable the comparison of different learning approaches, it was also the first platform to enable the comparison of multiple approaches across multiple datasets. Since its development, the NILMTK has been used to facilitate state-of-the-art results in various studies [9,20].

HMM. Studies using HMMs define several consumption states for each appliance, with each state receiving an individual probabilistic distribution. Despite its presence diminishing amongst studies, Hosseini et al. [21] found the recognition accuracy of the Factorial-Semi-HMM model (a variation of the HMM) to surpass that of RNNs, showcasing the model's competitiveness against a deep learning approach.

Deep Learning. Deep learning was first introduced to NILM by Kelly and Knottenbelt [22] who employed sequence-to-sequence learning (s2s). This approach trains a deep neural network to map between an input and an output sequence whilst removing the power consumption contribution of all but the target appliance. Note, the input sequence represents a window of aggregate power that applies to the target appliance and the output sequence is a prediction of the target appliances power consumption. Zhang et al. [23] evolved this approach, employing a unique sequence-to-point (s2p) model to predict the midpoint of the output signal as opposed to the entire sequence. Since then, s2s and s2p learning have become staple output approaches for NILM, most often utilising a CNN. The popularity of CNNs emanates from the model's resistance to noise in the input data as well as its ability to extract complex features that are independent from time via its convolutional layers. Studies have achieved high performance using both 1D and 2D convolutional layers. Notably, Barber et al. [24] applied four pruning algorithms on the s2p CNN seen in [23], reducing the number of weights by 87% whilst maintaining a state-of-the-art performance. Yang et al. [25] utilised 2D convolutional layers to conduct a current-to-image conversion, representing the characteristics of each appliance as the CNN's input.

RNNs are best known for their ability to handle time-series data. Additionally, variants LSTM and GRU's capacity to deal with the vanishing gradient problem make it an attractive approach for NILM tasks. Kim et al. [26] were first to explore state detection with a RNN, since then, the model has been adapted. Krystalakos et al. [27] take a GRU approach with a goal of improving disaggregation accuracy and reducing computational complexity. Similarly, Quek et al. [28] designed two LSTM networks for appliance classification. Alongside their results both the CNN and RNN's compatibility with the NILMTK has cemented their presence amongst NILM literature.

Industrial Studies. Martins et al. [6] and Kalinke et al. [7] use a combination of machine and deep learning approaches. Their goal to disaggregate the total load by each appliance using the HIPE, IMDELD and REDD datasets [7] found the deep learning models to record the lowest error rate in 11 out of 12 instances. Similarly, [6] disaggregate the total load by each appliance and found the WaveNILM model overcame the Factorial-HMM model in most evaluation metrics for every appliance. Holmegaard and Kjaergaard [29] instead compare three machine learning approaches, Combinatorial Optimization, FHMM and FHMM with day specific training, the latter providing the best outcome. Ultimately, whilst certain learning approaches appear more attractive than others, their performance is judged by the evaluation metrics.

2.3 Evaluation metrics

In NILM studies, there is no standardized evaluation metric. The most popular metrics are those used within regression and classification problems. This includes f-score, recall, accuracy, precision, true positives, true negatives, false positives, false negatives, root mean squared error (RMSE) and mean absolute error (MAE). Studies have also employed specific load disaggregation metrics that include the energy-based f-score [30], the total energy correctly assigned (TECA) [3] and the energy accuracy [8].

An important separation is offered by Mayhorn et al. [30] where metrics are categorised as either Event Detection (ED) or Energy Estimation (EE). The former regards NILM as a classification problem, assessing how the model performs in identifying an appliance's operational state. The latter treats NILM as a regression problem, approximating the amount of energy each appliance consumes at each timestamp. The authors insist ED and EE are equally significant. Furthermore, the inclusion of both metrics is considered essential to any NILM study as it allows for the most comprehensive assessment of a model's performance.

When considering the most appropriate evaluation metric Angelis et al. [8] insist studies should acknowledge both the operational power level and the frequency of usage for each appliance. This is because of the impact the operational power has on certain EE metrics like the RMSE and MAE. Attention is drawn to how a MAE comparison between two appliances operating at different wattage levels will not provide any significant conclusions given that the appliance operating on a higher level will ordinarily have a greater MAE. Consequently, RMSE and MAE are regarded as inappropriate evaluation metrics, despite their popularity amongst NILM literature. Despite comprehensively reviewing the theory, a full understanding of NILM isn't complete without considering its practical implementation.

3 Methods

The proposed IoT device (Figure 3) was built to record the frequency of electric meter pulses per minute. The resistor is responsible for regulating the flow of electrical current whilst the photoresistor registers each pulse from the electric meter. This resistor was connected to the photodiode using the appropriate pins. The antenna provided an internet connection, and the battery powered the device ready for piloting.

Before implementation of the device, the ambient light level in the garage (location of the electric meter) was identified. Knowledge of this allowed for the device to be coded to recognise a change in the light level as the electric meter pulsed. The device was positioned accordingly and the pulse counting started. The number of pulses observed each minute were posted to Ubidots [31], an IoT platform used to store and visualize data. Periodically, different appliances in the house were switched on, causing a change in the number of pulses observed by the device. Using Ubidots, the different events (appliances changing state) were visualised.

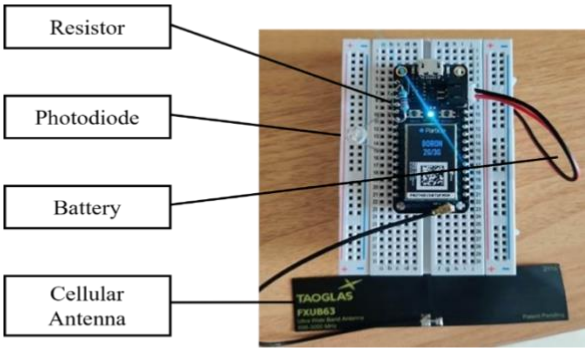

Fig. 3. Wiring diagram of IoT device.

The intention was then to establish the load signatures of a sample of appliances as well as the consumption patterns of different appliance types and a cost analysis for a sample of appliances.

An experiment was conducted to identify each respective load signature of each appliance by periodically switched them on and off again over a period of 21 hours. The study did record for 24 hours with three of having no additional appliances used allowing the baseline ambient pulse to be established.

Using Equation (1) a prediction was made as to how many times the electric meter would pulse. The ambient pulse level relates to the low power consuming appliances operating in the background at baseline.

$$\text{Pulses per min} = 60 / (3600/\text{Appliance Wattage}) + \text{Ambient Level} \quad (1)$$

Equation (2) was then used to convert the number of pulses to kWh, which is a measure of how much energy is being used per hour.

$$\text{kWh} = (\text{number of pulses}) / 1000 \quad (2)$$

The cost for each appliance can then be calculated by multiplying the cost per unit of energy per kWh by the number of KWh energy consumed from equation (2).

4 Results

Pulse Experiment. The results of the pulse experiment (Table 3) were recorded over a one minute period. Pulses ranged from 1-5 per second with the expectation that if there is 1 pulse per second, then there should be 60 pulses recorded per minute, 2 pulses per second should yield 120 pulses recorded per minute and so on, hence formulating the 'Pulses expected' column in Table 3. The 'Pulses recorded' are those observed in the experiment with the final column in the table representing the difference between the expected and observed. The results show the device performed well up to 4 pulses per second with the performance dropping off at 5 pulses per second, thus the device's performance is satisfactory.

Table 3. Experiment results.

Pulses per second	Pulses expected	Pulses recorded	Difference
1	60	60	0
2	120	118	2
3	180	176	4
4	240	234	6
5	300	284	16

Identifying Appliance Load Signatures. Given the known Wattage of the various appliances, it was possible to determine an expected number of pulses per appliance using equation (1). The results (Table 4) showed a high level of accuracy where the predictions were close to the actual. A visualization of the actual load signature for each appliance is illustrated in Figure 4.

Table 4. Pulses prediction vs actual.

Appliance	Wattage	Predicted	Actual
Toaster	1500W	29	28
Kettle	2500W–3000W	42-52	49
Hairdryer	2000W	38	37

Observing Appliance Behaviour. Figure 5 depicts how the consumption pattern (kWh) varies across different appliance types over the observation time. For example, the shower (type I) has two states, which is either OFF or consuming 0.082kWh. Conversely, the dishwasher is a type II appliance. Across its 30 minute cycle it consumes between 0.003kWh and 0.084kWh. Despite appearing to behave sporadically, the dishwasher has a fixed number of states making it easier for the disaggregation algorithm to learn its load signature in training and disaggregate it effectively. An example of a type III appliance could not be included as they are uncommon within a household setting. The main takeaway however is that the energy consumption will be different each time the appliance is used, making disaggregation difficult.

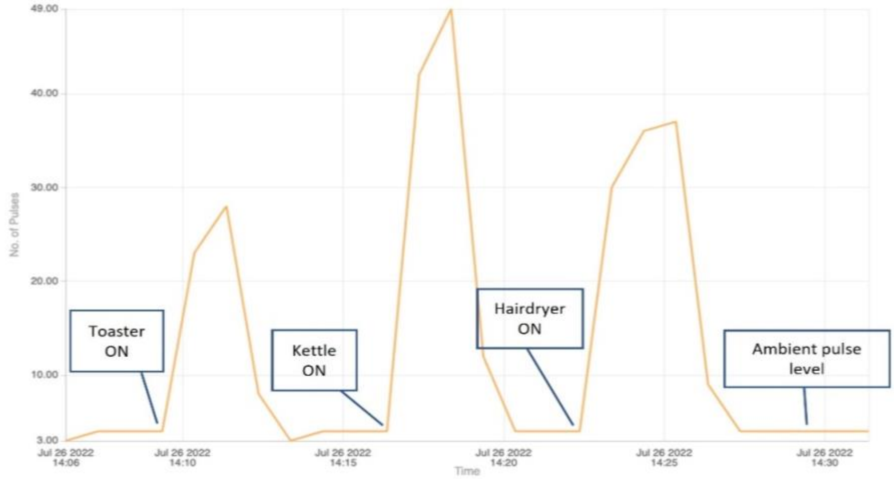

Fig. 4. Load signatures for toaster, kettle and hairdrye

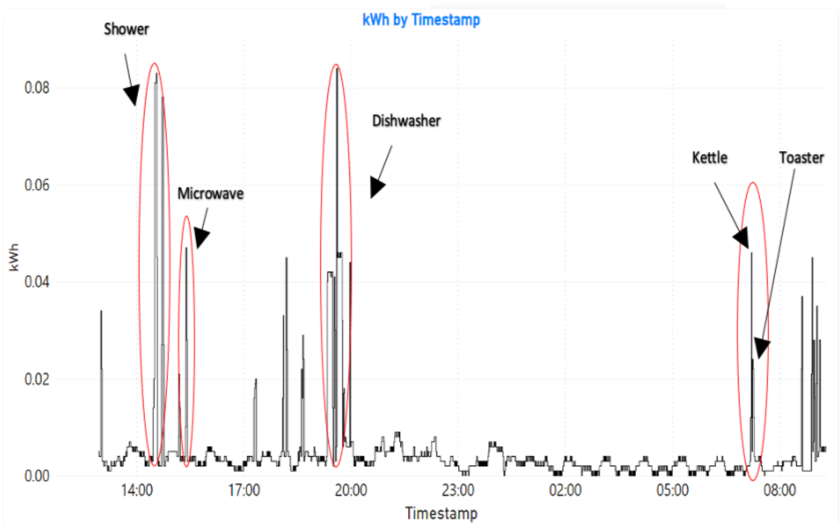

Fig. 5. Aggregate kWh over 21-hour period.

Cost Analysis. The cost of running each appliance is a product of the appliance's total energy consumption over its runtime and the tariff [32] imposed by the electricity provider. Examples using the data in Figure 5 are illustrated in Table 5. Whilst the dishwasher appears the most expensive, it is worth noting that costs are heavily influenced by each appliance's runtime.

Table 5. Cost analysis.

Appliance (Type)	kWh	Tariff (£ per kWh)	Cost (£)
Shower (I)	0.328	0.2952	0.097
Dishwasher (II)	1.013		0.299
Microwave (I)	0.073		0.022
Kettle (I)	0.075		0.022
Toaster (I)	0.058		0.017

5 Discussion

5.1 Data

To date there are only three recognised industrial NILM datasets. This poses a problem for NIW as a lack of available data has intrinsically led to a lack of research in NILM within industry. Until industrial NILM studies become more abundant, our understanding of both the benefits and challenges will be limited. However, it is still possible to gather valuable insights from residential datasets but with the caveat that there are characteristics of industrial datasets that are not transferrable. For example, within the HIPE dataset, Kalinke et al. [7] acknowledge a dependency between different industrial appliances like the screen printer and the soldering oven, these dependencies are less common in residential data. Additionally, unlike household appliances, industrial machinery tends to be of Type III classification given the continuous nature of its energy demand. Type III is also regarded as the most difficult appliance type to disaggregate and likely holds a strong presence amongst NIW's machinery.

Existing literature indicates NIW can achieve optimal results collecting data at a high or low frequency. With a low sampling rate NIW must be aware of the difficulty in detecting low energy consuming devices considering that switching events are not prominent. Despite this, steady state features arising from a low sampling rate, namely Active Power, continue to deliver state-of-the-art results, particularly within industrial studies. Conversely, a high sampling rate brings the benefits of transient features which are useful for distinguishing appliances operating simultaneously. However, this requires using more expensive hardware to collect the data. The data is also complex and requires a high computation time for pre-processing. The best approach may be to sample at a rate that captures both steady and transient features as seen in [16].

Just as there is no recommended sampling rate, NIW can choose any number and combination of features for their dataset. In line with other industrial studies [6-7] it may be wise to begin with Active Power as the solitary feature and then follow with a combination of steady state and statistical features given the effectiveness of statistical features in disaggregating Type I and II appliances [17].

5.2 Machine learning approaches

Machine learning approaches continue to feature amongst studies most likely due to their simpler implementation and lower computational complexity. Whilst the results in Hosseini et al. [21] indicate the HMM is a competitive disaggregation model, HMMs are known to be inefficient should the number of disaggregated appliances increase [33]. This has led to HMMs only being included as one of several benchmark models.

Zhang et al. [23] had one of the most impactful contributions to NILM literature highlighting the flexibility of a neural network architecture, a characteristic that is less prevalent within machine and statistical learning approaches. In this study, five convolutional layers were included, allowing the model to extract the most meaningful features to perform disaggregation. Additionally, the CNNs ability to include both 1D and 2D convolutions make it an attractive approach for NILM.

The RNN's functionality within NILM derives from its ability to capture the sequential information of the input data whilst retaining a memory of what it has processed in previous steps, two features that are non-existent amongst machine learning models, yet extremely effective for NILM. Given the efficiency of the convolutional layer as a feature extractor and the LSTM/GRUs applicability for sequential data it is unsurprising that authors have designed their models to include each of these components, it is advisable for NIW to do the same.

Both the RNN and CNN are two of eight different algorithms compatible with NILMTK. Like Kalinke et al. [7] NIW can train each algorithm using the implementation at NILMTK and then evaluate the performance of each of these algorithms using their own appliance level data. Ultimately, NILMTK gives NIW a guided introduction to the NILM process alongside the tools to implement it practically. NIW must however be aware of the time and resources necessary to train each model given the extensive number of parameters and their complexity for understanding.

From NIW's perspective, the HMM can provide adequate disaggregation results with a simpler implementation and at a lower cost. However, modern literature dictates that deep learning models provide promising results. Ultimately, NIW must decide whether less computationally intensive models like the HMM are worth pursuing at the expense of higher accuracy results. Equally NIW must have the resources and understanding to employ complex deep learning models like the CNN and RNN.

5.3 Practical discussion

Results from Section 4 provided both the load signatures and a cost analysis for a sample of household appliances as well as drawing attention to how the consumption pattern varies amongst different appliance types.

From NIW's perspective, identifying the load signatures (Figure 4) will be less straightforward. This is due to industrial machinery operating simultaneously, creating more overlap amongst the aggregate data. Another consideration for NIW is the sampling rate used. Data in Figure 5 is collected every minute, this is a low sampling rate however it provides enough data for a signature to notably develop. Too low of a sampling rate such as aggregate data collected every 30 minutes, would make appliance disaggregation difficult as it may fail to capture an appliance's runtime. Conversely, a higher rate such as every second, would make disaggregation much easier.

Providing a cost analysis for individual appliances is one of the key advantages of NILM. However, the cost of an appliance can be calculated using Equations (1) and (2) so it is important to highlight the refinements that NILM has to offer. NILM identifies the state, runtime, and consumption for each machine. This facilitates a detailed cost breakdown that would not be possible using just these two equations. Secondly, NILM acknowledges type III machinery which features prominently in industry. The amount of energy consumed by a type III machine is purely circumstantial, limiting the ability to predict its cost. For NIW this may relate to a water pump that's runtime is fluctuating in response to a change in rainfall. An effective NILM algorithm can detect this fluctuation, providing an accurate breakdown of the cost.

6 Conclusion and recommendations

Rising energy tariffs have made NILM imperative for NIW. This paper considers NILM from a theoretical and practical perspective, with a particular focus on the use case for NIW and the 'state-of-the-art' for each element of the NILM process. Despite a lack of industrial studies, results showed that deep learning models are dominating the literature and providing state-of-the-art results. Furthermore, the practicality of NILM was demonstrated via a cost analysis using data collected with an IoT device.

It is apparent that more industrial studies are required to fully understand the behaviour of type III appliances. Future studies should also consider NILM as both a classification and regression task as well as adopting the energy f-score for load disaggregation and TPs/TNs/FPs/FNs for event detection. This will help standardise the evaluation process whilst facilitating the comparison of different NILM models. Furthermore, to overcome the extensive training time of deep neural networks future studies should consider lightweight disaggregation models.

For NIW the next step should be to explore the data collection process at both an aggregate and appliance level. This will require deploying high frequency energy meters on sites where NIW have already existing sub-metering installed. The frequency of data collection may require alteration for the most optimal level. Meanwhile the readings from the sub-meters can be used to characterize equipment and provide ground truth data. In the first instance, the NILMTK will be used to gain a practical understanding of NILM using pre-existing industrial datasets, with the view to incorporating NIW data once it is available. The NILM techniques may then be applied to the aggregated data from NIW's high frequency meters and the results compared with the known data from sub-metering. This will prove particularly valuable to the business when considering application to sites where there is a concentration of energy-intensive assets such as the water treatment works and wastewater treatment works.

References

1. Hart, G.W.: Nonintrusive appliance load monitoring. In: Proceedings of the IEEE, vol. 80, no. 12, pp.1870-1891 (1992).
2. Le, T., Kim, H.: Non-Intrusive Load Monitoring Based on Novel Transient Signal Household Appliances with Low Sampling Rate. Energies 11(12), 3409 (2018).
3. Kolter, J.Z., Johnson, M.J.: REDD: A public data set for energy disaggregation research. In: Workshop on data mining applications in sustainability (SIGKDD), vol. 25, pp. 59-62, San Diego (2011).
4. Parson, O., Ghosh, S., Weal, M., Rogers, A.: Non-intrusive load monitoring using prior models of general appliance types. In: Proceedings of the 26th AAAI Conference on Artificial Intelligence, 26(1), pp. 356-362, Toronto, Canada (2012).
5. Bischof, S., Trittenbach, H., Vollmer, M., Werle, D., Blank, T., Böhm, K.: HIPE: An energy-status-data set from industrial production. In: Proceedings of the 9th International Conference on Future Energy Systems, pp. 599-603 (2018).
6. Martins, P.B., Gomes, J.G., Nascimento, V.B., de Freitas, A.R.: Application of a deep learning generative model to load disaggregation for industrial machinery power consumption monitoring. In: IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm), pp. 1-6 (2018).
7. Kalinke, F., Bielski, P., Singh, S., Fouché, E., Böhm, K.: An evaluation of nilm approaches on industrial energy-consumption data. In: Proceedings of the 12th ACM International Conference on Future Energy Systems, pp. 239-243 (2021).
8. Angelis, G.F., Timplalexis, C., Krinidis, S., Ioannidis, D., Tzovaras, D.: NILM Applications: Literature review of learning approaches, recent developments and challenges. Energy and Buildings 261, p.111951 (2022).
9. Pereira, L., Nunes, N.: Performance evaluation in non-intrusive load monitoring: Datasets, metrics, and tools—A review. Wiley Interdisciplinary Reviews: data mining and knowledge discovery, 8(6), p.e1265 (2018).
10. Kalogridis, G., Efthymiou, C., Denic, S.Z., Lewis, T.A., Cepeda, R.: Privacy for smart meters: Towards undetectable appliance load signatures. In: 1st IEEE International Conference on Smart Grid Comms, pp. 232-237 (2010).
11. Shin, C., Rho, S., Lee, H., Rhee, W.: Data requirements for applying machine learning to energy disaggregation. Energies, 12(9), p.1696 (2019).
12. Valenti, M., Bonfigli, R., Principi, E., Squartini, S.: Exploiting the reactive power in deep neural models for non-intrusive load monitoring. In: International Joint Conference on Neural Networks (IJCNN) pp. 1-8 (2018).
13. Quy, T.L., Zerr, S., Ntoutsi, E., Nejdl, W.: Data augmentation for dealing with low sampling rates in NILM. arXiv preprint arXiv:2104.02055 (2021).
14. Bonfigli, R., Principi, E., Fagiani, M., Severini, M., Squartini, S., Piazza, F.: Non-intrusive load monitoring by using active and reactive power in additive Factorial Hidden Markov Models. Applied Energy, 208, pp.1590-1607 (2017).
15. Meehan P, et al.: Temporal and frequency analysis of power signatures for common households. In: Proceedings of the ICT and Energy Efficiency and Workshop on Information Theory and Security (CIICT) pp. 22-27 (2012).
16. Reddy, R., Garg, V., Pudi, V.: A feature fusion technique for improved non-intrusive load monitoring. Energy Informatics, 3(1), pp.1-15 (2020).
17. Schirmer, P.A., Mporas, I.: Statistical and electrical features evaluation for electrical appliances energy disaggregation. Sustainability, 11(11), p.3222 (2019).

18. Tabatabaei, S.M., Dick, S., Xu, W. Toward non-intrusive load monitoring via multi-label classification. IEEE Transactions on Smart Grid, 8(1), pp.26-40 (2016).
19. Batra, N., et al.: NILMTK: An open source toolkit for non-intrusive load monitoring. In: Proceedings of the 5th International Conference on Future Energy Systems, pp.265-276 (2014).
20. Batra, N., et al.: Towards reproducible state-of-the-art energy disaggregation. In: Proceedings of the 6th ACM International Conference on Systems for Energy-efficient Buildings, Cities, and Transportation, pp.193-202 (2019).
21. Hosseini, S.S., et al.: Non-intrusive load monitoring through home energy management systems: A comprehensive review. Renewable and Sustainable Energy Reviews, 79, pp.1266-1274 (2017).
22. Kelly, J., Knottenbelt, W.: Neural nilm: Deep neural networks applied to energy disaggregation. In: Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-efficient Built Environments, pp.55-64 (2015).
23. Zhang, C., Zhong, M., Wang, Z., Goddard, N., Sutton, C.: Sequence-to-point learning with neural networks for non-intrusive load monitoring. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 32, no. 1 (2018).
24. Barber, J., Cuayáhuitl, H., Zhong, M., Luan, W.: Lightweight non-intrusive load monitoring employing pruned sequence-to-point learning. In: Proceedings of the 5th International Workshop on Non-Intrusive Load Monitoring, pp. 11-15 (2020).
25. Yang, D., Gao, X., Kong, L., Pang, Y., Zhou, B.: An event-driven convolutional neural architecture for non-intrusive load monitoring of residential appliance. IEEE Transactions on Consumer Electronics, 66(2), pp.173-182 (2020).
26. Kim, J., Le, T.T.H., Kim, H.: Nonintrusive load monitoring based on advanced deep learning and novel signature. Computational intelligence and neuroscience (2017).
27. Krystalakos, O., Nalmpantis, C., Vrakas, D.: Sliding window approach for online energy disaggregation using artificial neural networks. In: Proceedings of the 10th Hellenic Conference on Artificial Intelligence, pp.1-6 (2018).
28. Quek, Y.T., Woo, W.L., Logenthiran, T.: Load disaggregation using one-directional convolutional stacked long short-term memory recurrent neural network. IEEE Systems Journal, 14(1), pp.1395-1404 (2019).
29. Holmegaard, E., Kjærgaard, M.B.: NILM in an industrial setting: A load characterization and algorithm evaluation. In: IEEE International Conference on Smart Computing (SMARTCOMP), pp. 1-8, (2016).
30. Mayhorn, E.T., Sullivan, G.P., Petersen, J.M., Butner, R.S., Johnson, E.M.: Load disaggregation technologies: real world and laboratory performance. Pacific Northwest National Laboratory (PNNL), Richland, WA (US), Tech. Rep. PNNL-SA-116560 (2016).
31. https://ubidots.com/docs/ [Accessed 15 August 2022]
32. https://www.ofgem.gov.uk/check-if-energy-price-cap-affects-you [Accessed 15 August 2022]
33. Lanzisera, S., Dawson-Haggerty, S., Cheung, H.I., Taneja, J., Culler, D., Brown, R.: Methods for detailed energy data collection of miscellaneous and electronic loads in a commercial office building. Building and Environment, 65, pp.170-177 (2013).